

Viorel Badescu
Editor

MARS

Prospective
Energy
and Material
Resources



Springer

Mars

Prospective Energy and Material Resources

Viorel Badescu (Ed.)

Mars

Prospective Energy and Material
Resources



Springer

Prof. Viorel Badescu, Candida Oancea Institute
Polytechnic University of Bucharest
Spl. Independentei 313, Bucharest
060042
Romania
E-mail: badescu@theta.termo.pub.ro

ISBN 978-3-642-03628-6

e-ISBN 978-3-642-03629-3

DOI 10.1007/978-3-642-03629-3

Library of Congress Control Number: 2009934313

© 2009 Springer-Verlag Berlin Heidelberg

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Data supplied by the authors

Production & Cover Design: Scientific Publishing Services Pvt. Ltd., Chennai, India

Printed in acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

This book is dedicated to those who believe that any sensible decision should be made by taking into account the world as it will be.

Foreword

Mars, the Red Planet, fourth planet from the Sun, forever linked with 19th and 20th Century fantasy of a bellicose, intelligent Martian civilization. The romance and excitement of that fiction remains today, even as technologically sophisticated robotic orbiters, landers, and rovers seek to unveil Mars' secrets; but so far, they have yet to find evidence of life. The aura of excitement, though, is justified for another reason: Mars is a very special place. It is the only planetary surface in the Solar System where humans, once free from the bounds of Earth, might hope to establish habitable, self-sufficient colonies.

Endowed with an insatiable drive, focused motivation, and a keen sense of exploration and adventure, humans will undergo the extremes of physical hardship and danger to push the envelope, to do what has not yet been done. Because of their very nature, there is little doubt that humans will in fact conquer Mars. But even earth-bound extremes, such those experienced by the early polar explorers, may seem like a walk in the park compared to future experiences on Mars.

Mars is a small planet, with a radius about half that of Earth, but it is a place of extremes. The temperature can fall to a low of -140°C (-220°F) in polar winter, but seem an almost pleasant 20°C (68°F) in tropic summer. The atmosphere is thin. Even at the surface, Mars' atmospheric pressure is comparable to Earth's at an altitude of 35 km (22 miles), but with an extreme difference in composition. The atmosphere of Mars is 95.7% carbon dioxide with just 0.2% oxygen. There is no magnetic field to divert the solar wind, although much is stopped by interactions in the upper atmosphere. And, the surface is windy, very windy; dust storms are common occurrences which can upon occasion envelop the entire planet. But water does in fact occur, and that's what makes the establishment of habitable, self-sufficient colonies on Mars feasible.

The once seemingly impossible dream is on the way to becoming reality. The first steps have been taken; the robots are exploring Mars. Now, the planning begins in earnest. Future self-sufficient colonies will need energy and material

resources. The planning for those needs, already begun, is the theme and focus of this book: Mars: Prospective Energy and Material Resources.

We have to welcome this initiative of Springer and the highly appreciated effort of Professor Viorel Badescu to collect twenty-six chapters, written by selected experts in the field.

June 2009

J. Marvin Herndon
San Diego, USA

Preface

A manned mission to Mars has been planned since Wernher von Braun's times (1912 – 1977). Von Braun succeeded in placing astronauts on the Moon on 20 July 1969. However, the next logical step in his legacy, a manned mission to Mars was postponed for various reasons related to the financial and strategic considerations of the main national and international space agencies. Now, von Braun's vision finally seems to be turned into reality. An advanced development project called "A Manned Expedition to Mars" was drafted in 2005 as part of the Russian Federal Space Program. The very optimistic time target might be 2017-2018 (the nearest "ballistic window" for an Earth-Mars mission) but pessimists consider it unrealistic, above all for technological reasons. In September 2007, Michael Griffin, the NASA administrator, hinted that NASA may be able to launch a human mission to Mars by 2037. Also, the European Space Agency has the Aurora Program, the long-term vision of sending a human mission to Mars by 2030.

A safe energy supply is a key issue, especially when considering a manned mission where Environmental Control and Life Support devices have to work safely throughout the whole mission endurance timeframe. The future Mars surface exploration stations will be based on In-Situ Resource Utilisation (ISRU) technology, which means that indigenous materials will be used to produce various utilities, including rocket propellants and human life-supporting consumables. This reduces dramatically the weight that has to be raised from the Earth. Indeed, among other advantages, an ISRU plant equipped with a light stock of hydrogen can be used to produce a multiple of propellant mass using in situ atmospheric carbon and oxygen. This process makes a sufficient supply of energy viable, since the conversion of hydrogen into methanol, using indigenous carbon dioxide for example, is a very endothermic process.

The manned mission is seen as a first step towards a Mars surface exploration base-station and, later, establishing permanent Mars settlements. The location and use of Mars's natural resources is vital to enable cost-effective long-duration human exploration and exploitation missions to the Red Planet as well as subsequent human colonization. Planet resources include various crust-lodged materials, a low-pressure natural atmosphere, assorted forms of utilizable energy, lower gravity than Earth's, and ground placement advantages relative to human operability and living standards will be researched and developed fully. Power resources may

include solar and wind energy, the importation of nuclear reactors and the harvesting of geothermal potential. Mars's surface material and energy disequilibrium is, nowadays, strictly a response mainly to internal radiogenic heat and external solar radiation. Soon, large-scale human exploration, exploitation and settlement activities—extraction of water, food production, tapping of energy resources and waste disposal will add an anthropogenic forcing additional to tectonic, climatic and space debris bombardment. In fact, eventually, Mars may become humanity's Second Biosphere homeland—a new branch of human 21st Century civilization could become established there permanently.

But, meantime, an inventory and proper social assessment of Mars's prospective energy and material resources for human use is required. This book investigates the possibilities and limitations of various systems supplying manned bases on the surface of Mars with energy and other vital resources. The book collects together recent proposals and innovative options and solutions. It is a useful source of condensed information for specialists involved in current and impending Mars-related activities and a good starting point for young researchers.

The book is structured along logical lines of progressive thought and may be conceptually divided into three parts. The first part contains thirteen chapters and deals with energy sources and resources on Mars. After the introductory Chap. 1, presenting the history of power systems for Mars surface exploration, a group of four chapters (2 to 5) refers to solar energy utilization. Systems based on photovoltaic and photothermal conversion are presented. Chapter 6 is devoted to power systems based on fuel cells. The next four chapters focus on the utilization of nuclear energy. The first in this series (Chap. 7) presents the evolution of nuclear power generation for Mars. The series continues with Chap. 8, where distributed power sources are considered in the perspective of Mars colonization. Chapter 9 proposes advanced nuclear compact structures while Chap. 10 develops new solutions for nuclear energy and flights on Mars in the middle-term future. The utilization of geothermal energy on Mars is discussed in Chap. 11. The last chapters in the first part (12 and 13) present a comparative analysis of power systems architectures and economic aspects of energy utilization on Mars, respectively, after human settlements are established.

The second part of the book contains eight chapters and deals with technological aspects related to material resources existing on Mars. Chapter 14 proposes Tumbleweed, a new paradigm for surveying Mars surface for in-situ resources while Chap. 15 presents methods of drilling and excavation for construction and in-situ resource utilization. Chapter 16 gives a broad perspective of using local resources while the next four chapters show details concerning the utilization of rock and soil (Chap. 17), on-situ resources availability for space agriculture (Chap. 18), utilization of regolith for manufacturing construction material (Chap. 19) and the utilization of Martian iron and carbon dioxide (Chap. 20). The last chapter in this part (21) presents a Mars aqueous processing system for producing iron, high-grade metal oxides and oxygen from Mars soils.

The third part of the book contains five chapters and deals with material resources utilization after Mars colonization. Chapter 22 presents how the gravitational potential of Mars may be used in relation with human physical

growth. Chapters 23 and 24 deal with artificial environments on Mars and self-sustaining colonies, respectively. The last two chapters refer to climate engineering using orbiting solar reflectors (Chap. 25) and large scale water transportation on Mars (Chap. 26).

All chapters are highly innovative but the ideas they present are in different stages of elaboration. This is obviously related to the expected distance in the future of their implementation. Some chapters (such as 1-3, 5-7, 14-19, 21) refer to technologies ready to use in the near future. Other chapters mainly refer to later stages of Mars exploration, including Mars colonization (e.g. Chaps. 4, 8-13, 20, 22-26). However, this is a rather artificial separation and parts of some chapters could be easily classified as belonging to the other category. More details about the twenty-six chapters of the book are given below.

Chapter 1 by Simon Fraser discusses the possibilities and the limitations of available energy resources and power generation options for Mars surface applications, and evaluates the power systems of previous and present Mars surface exploration systems with respect to future applications. The surface of Mars is a challenging environment with respect to possible power system options for robotic and manned surface exploration. The low ambient temperatures require passive and/or active thermal protection and control systems with many of the power system elements used in state-of-the-art spaceflight applications. The solar radiation available on the surface of Mars is weak compared to the irradiance available in Earth orbit or in lunar applications. The weak solar radiation is further reduced by dust storms and dust deposition on photovoltaic and/or solar thermal panels. Wind energy utilization is theoretically possible, but strongly limited due to the thin atmosphere. In addition, setup and installation of large wind turbines in remote or fully automated operation is a very challenging task. Landing of nuclear and non-nuclear energy source is another option already proven in previous missions, and to be evaluated for the larger and more challenging missions to come.

Chapter 2 by Viorel Badescu deals with available solar energy and weather forecasting on Mars surface. The chapter studies the optimum slope and orientation of a surface receiving solar radiation on the Mars surface (Northern surface). A model to evaluate direct, diffuse and global irradiance on arbitrary tilted surfaces is developed. Also, the model estimates direct, diffuse and global irradiances on different time intervals, such as a day, a season or a year. Any of these irradiances/irradiations is a maximum for a certain optimum tilt angle. Most part of the primary data is based on Viking Landers measurements. The recommended strategies to collect the maximum amount of solar radiation depend on the dust content in the atmosphere. During the "clear" days the best strategy is to keep the receiving surface perpendicular on Sun's rays. This strategy allows collecting the maximum direct solar radiation and it is recommended in the case of solar concentrators. A simpler procedure is to use a South oriented surface whose tilt angle is continuously adjusted in order to collect the maximum global irradiance. During the dust storms the dependence of the incident irradiance upon surface's tilt and orientation becomes less important and simpler strategies are more convenient (e.g. a horizontal receiving surface or a once-per-day adjustment of the tilt angle). Regression relations between the optical depth on Mars and common near-surface meteorological parameters

(such as pressure and temperature) are also derived in this chapter. The accuracy of all types of regressions is lower during dust storm season while during “clear” sky periods the accuracy is usually good. Atmospheric pressure is more recommended than ambient temperature as a regression parameter. The influence of the areocentric longitude and local solar time on regression’s accuracy is important mainly during the dust storm season. Generally, the regression accuracy is very good at low atmospheric pressure (i.e. on clear sky). The regression accuracy is better at VL2 site than at VL1 site, especially at higher atmospheric pressure. Optical depth is well correlated to the daily variation of surface pressure Δp and to the standard deviation of daily surface pressure $\sigma_{p,\text{day}}$. Stationarity of daily averaged surface pressure time series on Mars can be obtained (by differencing) for short time intervals only. Large pressure variance intervals are usually associated with the dust storm season. Medium and small pressure variance intervals can be found usually during clear sky periods. General auto-regressive integrated moving average (ARIMA) models were developed for daily average pressure forecasting at Viking Lander sites. A procedure to find the “appropriate” model for a given time series was developed. It is based on various degrees of differencing the original time series and a t-statistics-assisted estimation of the significance of the fitted coefficients. A method employing the cumulative spectrum of the residuals was used to check the models. ARIMA(3,2,3) seems to be the most “appropriate” model to forecasting the daily average surface pressure on Mars. Seasonal ARIMA models were used to forecast various weather parameters (as atmospheric pressure and temperature and wind speed) on short time intervals (less than one day).

Chapter 3 by Viorel Badescu deals with simulation of solar photovoltaic cells on the surface of Mars. PV arrays based on silicon and gallium arsenide solar cells are analyzed under the assumption of both ideal and present state-of-the-art technology. Meteorological measurements provided by the Viking Landers are used as inputs. They refer to ambient temperature, wind speed and the optical depth of the Martian atmosphere. The incoming global solar radiation flux (direct, diffuse and ground reflected) is computed by using a model presented in Chapter 2. A simple thermodynamic model of solar cell is developed. Simulations were performed for both Viking Landers 1 and 2 sites during all the four Martian seasons. Information about the optimum voltage and cell temperature is provided as well as details about the electric power supply and cell efficiency. Different sorts of stationary and tracking solar cells arrays are studied. The PV efficiency is higher than that obtained under terrestrial conditions due to the lower ambient temperature on Mars. One confirms that the low operating temperatures will reduce the advantage in efficiency of GaAs solar cells over silicon. The strategy to be chosen for maximum power supply mainly depends on the atmospheric optical depth. The results reported in this chapter should be seen as upper bounds for real solar cell performances since a number of disturbing phenomena, such as dust deposition on the surface of the panels, ageing and temperature-induced degradation, were neglected.

Chapter 4 by Pini Gurfil and Joseph Cory presents a novel design comprising an original use of photovoltaic technologies for generating energy on Mars. The authors suggest using a lighter-than-atmosphere helium-filled platform containing

embedded thin photovoltaic arrays (PVA) in order to extract electrical power from the incoming solar radiation. The goal is to develop an efficient, portable, land-independent, inexpensive and reliable energy source for future exploration of Mars. Three concepts are developed: (i) a simple, helium-filled spherical balloon containing a PVA on the outer skin surface; (ii) a specialized paraboloidal balloon with a transparent upper part and an opaque lower part containing paraboloid-shaped photovoltaic cells on the inner surface, thus increasing the solar energy conversion efficiency; and (iii) the same balloon wherein the inner collectors are replaced by flexible thin mirrors, and the upper transparent part contains a PVA at the focus of the inner paraboloid-shaped mirror, thus augmenting the incoming solar energy flux. In both the spherical and paraboloidal configurations, the power output is similar; however, the power-to-volume ratio of the paraboloidal design is higher. A detailed preliminary design is presented.

Chapter 5 by Viorel Badescu deals with weather influence on solar thermal power plants operation on Mars. A "dynamic" solar power plant (which consists of a solar collector - thermal engine combination) is proposed as an alternative for the more usual photovoltaic cells. Upper bounds for the efficiency of solar thermal power plants operating in the Martian environment are first evaluated. A general thermodynamic approach clearly shows which of the three theories usually quoted in literature gives the exergy of thermal radiation. Recent works reporting accurate upper bounds for the efficiency of thermal radiation energy conversion into work are subsequently used in this chapter. The results refer to thermal engines powered by direct or diffuse solar radiation on Mars. Diffuse solar radiation is modeled as diluted or multiply scattered thermal radiation. A more elaborated model uses an endoreversible Carnot cycle to describe solar engine operation. Two strategies to collect solar radiation are analyzed: a solar horizontal collector (H) and a solar collector whose tilt and orientation are continuously adjusted to keep the receiving surface perpendicular on Sun rays (P). Meteorological data measured at Viking Landers (VL) sites are used in computations. Results show that generally the influence of latitude on performance is important. In some situations the meteorological effects compensate the latitudinal effects and the output power is quite similar at both VL1 and VL2 sites. During a winter dust-storm day the maximum output power is much smaller than during autumn. High efficiency thermal engines should be used in combination with solar collectors kept perpendicular on Sun's rays. When a horizontal solar collector is considered, the dependence of the maximum output power on optimum solar efficiency seems to be quadratic at both VL1 and VL2 sites. When a collector perpendicular on Sun's rays is considered, this dependence is more complicate, but keeps the quadratic feature. No obvious difference exists between power plant performances in the two years of VL2 operation. A solar Stirling engine based on a horizontal selective flat-plate converter is analyzed in the last part of the chapter. All the computations were performed for a solar collection area similar in size with that of Mars Pathfinder's Sojourner. The solar efficiency at noon is as high as 18 %. The power provided by the engine is as high as 16 W during autumn and winter. These results suggest that under the Martian environment the performance of properly designed solar Stirling engines is comparable with that of PV cell power systems.

Chapter 6 by Simon Fraser begins by showing that reducing mass and volume without sacrificing safety, reliability and robustness is a major challenge in power system design for space flight applications. Already in the past, fuel cells provided a very good overall package and were thus chosen for onboard power generation applications in the Gemini and Apollo capsules as well as in the Space Shuttle Orbiter. Significant advances have been achieved in fuel cell technology since the early 1990s, effectively replacing the liquid alkaline fuel cells currently still used aboard the Space Shuttle Orbiter by new cell designs based on polymeric and ceramic membranes. Current fuel cell technology, although still not being economically competitive in most terrestrial applications, is currently being considered as a very attractive power system option for a wide range of applications in Mars exploration. This includes high-performance power generation in stationary and mobile applications from the sub-Watt up to the Megawatt scale. This also includes electrical energy storage applications by combining a fuel cell and an electrolyser in a closed-loop system. In addition, fuel cells fit perfectly into state-of-the-art In-Situ Resource Utilisation scenarios as they can be used to convert hydrogen, carbon monoxide or hydrocarbon fuels into electricity very efficiently. Fuel cells are therefore promising candidates for a wide range of applications in future Mars surface exploration. Possibilities and limitations of this innovative power generation technology are critically reviewed.

Chapter 7 by Liviu Popa-Simil shows that the nuclear power systems for space applications have a 60-year long history and their advance was important to the success of various space missions having as final effect the explosion of knowledge about Mars from the last decade. Nuclear power is vital to successful medium and long-term missions in space, Mars representing the first benchmark in space exploration with manned missions. The evolution of nuclear power for Mars is strongly correlated with the entire planet development of the science, technology and politics. The main nuclear power sources are relying on radioactive isotopes and on nuclear reactors. The radioisotope based power sources have a simple robust construction, but low conversion efficiency, delivering continuous power. The nuclear reactor power sources have larger complexity, accommodate the power production with the demand, and have better efficiency and scalable power. The specific power per mass ratio favors the nuclear reactors over the isotope and solar power sources for large power applications over several kW. Over the time many nuclear reactor designs and many tests and studies have been performed covering a large variety of conversion processes, nuclear structures, materials and heat and electricity management concepts briefly presented. After all these 60 years of research, the competition is not closed and a race among various solutions is still in place. Heat pipe based reactor named HOMER seems to have the best chances for future applications. One very important development requirement is safety, and many safety concepts and designs have been developed. Another very important nuclear issue for the mission success is the shielding optimization and the users' radiation dose exposure. The message of the chapter is that in spite this very large complexity and diversity in technologic solution, the nuclear reactors are in a mature stage ready to successfully support a future Mars mission.

Chapter 8 by George H. Miley, Xiaoling Yang and Eric Rice shows that one of the fundamental needs for Mars colonization is an abundant source of energy. Two types of nuclear power plants, the IEC fusion torch and small LENR power units, are proposed to serve as initial power sources for Mars terraforming. All of the elements needed to construct these two power units exist on Mars, allowing the development of basic power units directly on the Red Planet. The IEC device would provide small central units in the 500 kWe - 1 MWe level. The LENR units would serve as small portable power sources ranging from watts up to kilowatts. All units would be designed to minimize radiation emission and radioactive waste. To reach the first step of Mars colonization and create a habitable environment, the IEC torch would be employed for producing super greenhouse gases to warm up the planet, and also for producing oxygen and nitrogen to form breathable air. All these gases are realized by decomposing and recombining the existing material on Mars. The first IEC fusion torch, the small LENR power units and some robotics may be brought from Earth. Robotics will do iron mining, to be supplied to the IEC torch for steel manufacture and components for self-reproducing. The small LENR power cells may also serve as a supplement to supply the starting power for the IEC devices. Thus this combination of power sources and processors can cover a wide range of basic needs for colonization of Mars.

Chapter 9 by Liviu Popa-Simil shows that after more than one half of century of space activity, putting a human crew on Mars remains a challenging, many requirements mission. The nuclear power systems gained a large diversity over time and represent an important factor of success providing the mission with sufficient available power. The specific mass /power ratio related to the difficulty to place in orbit heavy objects made the dominance of reactors cooled with light liquid metals as Li, Na, K or He and using dynamic conversion system based on Stirling or Rankine cycles, being mainly derivatives from space propulsion systems. A strong candidate for powering the near future missions is "HOMER" and its derivative "SAFE" reactors cumulating the necessary safety and operability features. The actual development of nuclear power considers the use of micro-nano-hetero-structures as potential candidates for advanced power systems. Most of these new proposals are in conceptual stage with promising theoretical and computer simulations developments, and only in part, backed by experimental results. It is expected that the micro-hetero structures to provide cleaner, durable nuclear reactors, generically called generation 5, and changes in the fuel cycle and waste management. The pulsed conversion system is proposed to harvest the thermal energy as electricity. The nano-hetero structures have great chances to drive the developments of direct nuclear energy conversion into electricity generation 6 compact nuclear reactors. Nano-cluster structures are seen usable in nuclear fuel breeding and isotopes transmutation. The use of nano-structures may provide light radiation shielding and electrically controlled direct conversion nuclear reactors, very light compatible with most of the mobile applications. The future belongs to complex hybrid fission fusion systems and antimatter stimulated fusion, with no radioactive waste and harming radiation. The need for developing advanced nuclear power systems is seen as a condition for the mankind of colonizing Mars.

Chapter 10 by Alexander Bolonkin has two parts. The first part describes production of electricity by a special lightweight thermonuclear reactor; in the second part the author considers a new form of electromagnetic space propulsion suitable for the rarefied Mars atmosphere. The future Mars colony will be a major user of electrical energy. No known native fuel stocks exist (though geological methane might be a possibility) but even if the fuel need not be made, the oxygen must be—there is almost no oxygen in the atmosphere of Mars. To power major installations during the polar night, a Mars electric plant may only be nuclear or thermonuclear. Today's projected designs for thermonuclear reactors have the weight of thousands of tons. Even with a revolution in cheap spaceflight—and achieved fusion—it would be very expensive to deliver to Mars the delicate yet massive components of a conventional design thermonuclear reactor—and then to assemble them there with scarce labor pools amounting to a handful of colonists. The author offers a new very low mass thermonuclear reflector reactor. The remarkable part of this new style of reactor is a triple net AB reflector, which confines the high temperature plasma. The plasma loses part of its energy when it contacts with the triple net but this loss can be compensated for by permanent additional plasma heating. When plasma is rarefied (has small density) the heat flow to the AB reflector is not large and temperature of the triple reflector net is lower than 2000 - 3000 K, survivable for tungsten. The offered AB-reactor has significantly less power than the current developed power reactors with magnetic or inertial confinement (hundreds-thousands of kW). But it is enough for Mars, many vehicles, ships and space apparatus. It requires no air, and it may be realized in the near term. The author made numerous innovations in the design of this reactor, researched its theory, developed methods of computation, and made the computation of typical project. Main advantage of this reactor is cheapness and ease of realization. In past research the author proposed a new engine which produces a large thrust without throwing away large amounts of reaction mass (unlike the conventional rocket engine). This propulsion system uses the rare atmosphere and low gravity of Mars. In the chapter, the author develops the theory of this kind of impulse engine and computes a sample project, which shows the large-scale possibilities opened by this new “AB-Space Propulsion”. The AB-Space Propulsion gets the energy from ground-mounted power stations; a planet's electrical station can transfer electricity up to 1000 millions (and more) of kilometers by plasma ‘wires’ streaming through empty space. The author shows that AB-Space Propulsion can produce thrust of 10 tons (and more). That can accelerate a space ship to some thousands of kilometers/second. AB-Space Propulsion has a staggering specific impulse owing to the very small mass expended. The AB-Space Propulsion reacts not by expulsion of its own mass (unlike rocket engine) but against the mass of Mars (located perhaps a thousand of millions of kilometers away) through the magnetic field of its plasma cable. For creating this plasma cable the AB-Space Propulsion spends only some kg of hydrogen.

Chapter 11 by Paul Morgan shows that the exploitation of terrestrial geothermal energy has grown significantly since about 1990, leading to the improvement of technologies for direct use and electricity generation. During the same time period, although no direct geothermal measurements have yet been made on Mars,

orbital imaging spacecraft and landers have greatly improved the resolution of Mars' surface mapping, and, together with modeling studies have increased the understanding of the thermal structure and evolution of Mars. General conclusions from these studies is that Mars has water in the near subsurface and has been geologically active in terms of faulting and volcanic activity in the recent geological past, perhaps as recent as 100 million years or younger, and has the potential for future volcanic activity. These conclusions give Mars good potential for geothermal resources. Mars has a very tenuous carbon dioxide atmosphere, which at an average of 7 mbar is less than one-hundredth the pressure of Earth's atmosphere. Its average surface temperature is -63°C. An "air"-cooled binary power plant system would be the most suitable technology for electricity generation on Mars. Energy from geothermal fluids could also be extracted for direct use, such as space heating.

Chapter 12 by Mathew Silver, Wilfried Hofstetter, Chase Cooper and Jeffrey Hoffman shows that power system design critically constrains human missions to the surface of Mars. A power system can be represented as a heterogeneous network of interacting elements, constituting both generation technologies and consumption technologies. Yet, most reference mission studies to date have focused on individual power systems or compared core technologies without considering their architectural relationships to non-core generating technologies and consumption technologies. This chapter provides a comparative analysis of power systems from an architectural perspective, for human Mars exploration. Thirteen competing architectures based around three core-generation technologies are enumerated and evaluated. The results demonstrate that, when compared at the architectural level, conclusions are not necessarily the same as at a point-design level. In particular, given the assumptions, solar-based architectures may be superior to alternative architectures based on nuclear systems, even for long-duration missions. These specific conclusions depend on assumptions with respect to power system performance and mission requirements. The authors argue that as our understanding of the requirements for Mars missions becomes more detailed, and specific technologies evolve to greater maturity, it becomes increasingly important to broaden the design analysis beyond point systems and beyond point missions. Comparative architectural analysis as presented is one way to do this rigorously. The preceding conclusion becomes even more pertinent if one considers trends in power system technologies. The architectural comparison is therefore complemented by a survey of existing power technologies and an analysis of emerging trends in cutting edge systems outside of aerospace including ultra-capacitors and bio-electric technologies.

Chapter 13 by Joseph E. Palaia IV, Mark S. Homnick, Frank Crossman, Alexander Stimpson and John Truett provides a snapshot of the way in which energy will effect human operations and infrastructure development on the Martian surface. The intent is to improve understanding of the context in which energy will play a central role, and to provide a powerful toolset for energy system comparison and long-term strategic planning. The chapter begins by discussing the significant role that energy has played in the history of human habitation and industry, observing the correlation between standard of living and energy

consumption. Next, one discusses the perceived role energy will play in the development of the Martian frontier, focusing on energy demand. Then the authors examine how this understanding can help with early decision making and long-term strategic planning, which includes the proposal to establish energy as the basic metric by which economic value on Mars is defined. A worked example, the design of a Martian greenhouse, illustrates this approach. Also an approach for approximating the energy “cost” of human effort is proposed. Trade between different Mars surface locales is predicted to develop, and such trade will be significantly influenced by the energy costs of transportation. Mars surface trade, as well as trade between Mars and other solar system domains is discussed. Finally, a simple economic model to illustrate the positive impact on the Martian economy of an early Mars industrial center is presented. While serving as an initial beachhead in the Martian frontier, it will provide key logistical support for further development and serve as a focus for trade.

Chapter 14 by Kimberly R. Kuhlman, Alberto Behar, Jack Jones, Penelope Boston, Jeffrey Antol, Gregory Hajos, Warren Kelliher, Max Coleman, Ronald Crawford, Lynn Rothschild, Martin Buehler, Greg Bearman and Daniel W. Wilson shows that long-range surveys of *in situ* resources on the surface of Mars could be readily accomplished with a fleet of Tumbleweeds – spherical vehicles capable of using the readily available Martian wind to traverse the surface of Mars with minimal power, while optimizing their capabilities to perform a variety of measurements over relatively large swaths of terrain. These low-cost vehicles fill the niche between orbital reconnaissance and landed rovers capable of much more localized study. Tumbleweeds can be configured to enable operation in varying terrains and accommodate a wide range of instrument packages making them suitable for autonomous surveys for *in situ* natural resources. Tumbleweeds are lightweight and relatively inexpensive, making them very attractive for multiple deployments or piggybacking on larger missions. A variety of Tumbleweed structures have been envisioned, including inflatable concepts constructed and tested by NASA’s Jet Propulsion Laboratory and deployable structures, including a Box Kite, Dandelion and Tumble-cup, which have been developed and tested by NASA’s Langley Research Center. A group of Tumbleweeds, each with unique analytical or long-range communication capability, could survey a particular region in a random manner. When something interesting is detected by one Tumbleweed, a swarm intelligence-based algorithm could be activated directing the others to proceed to the same general area and conduct additional sensing. For surveys of *in situ* resources, a variety of miniaturized instrumentation could be deployed including surface mounted soil moisture sensors, ground penetrating radar, gas sensors, a miniature X-ray fluorescence spectrometer and a multispectral imager.

Chapter 15 by Kris Zacny and Yoseph Bar-Cohen deals with drilling and excavation for construction and in-situ resource utilization on Mars. Mining for resources starts first with reconnaissance of the subsurface. This is initially accomplished using non-contact geophysical methods that include electro-magnetic, gravity and seismic. However, no mine will be established unless a more accurate ground-truth is performed. A more accurate resource evaluation takes place via

coring and drilling in order to obtain either intact core samples or rock chips for analysis. Once the subsurface has been evaluated, only then the mining equipment is brought in. There are a number of different mining approaches depending on the required output, formation strength and type of resource. However, environmental conditions also play a role. For example equipment that has to work in Canadian Arctic has to be modified to withstand extremely cold conditions. Mars, is a much colder planet than the Arctic, has a very thin atmosphere, a third of the Earth's gravity, and it is millions of kilometers away. All these conditions make mining and excavation on Mars orders of magnitude more challenging. This chapter initially focuses on these conditions on Mars that affect mining operations. A number of Mars drilling approaches are given, with a few examples of existing systems developed just for this planet. Various mining and excavation methodologies are outlined. These include conventional open pit mining and somewhat unconventional water mining using water-ice sublimation. The chapter ends with an example of a "mine of the future" that is currently being developed in a remote location on Earth and will no doubt serve as a guide for the future mine on Mars.

Chapter 16 by Donald Rapp shows that in situ resource utilization (ISRU) on the Moon or Mars is an approach for converting indigenous resources into various products (primarily propellants and life support consumables) that are needed for a space mission. By utilizing indigenous resources, the amount of materiel that must be brought from Earth may be reduced, thus reducing the Initial Mass in Low Earth Orbit (IMLEO.) IMLEO is typically used as a measure of the mission scope and cost. Lunar ISRU suffers from the need for high-temperature processing or a lack of abundant accessible feedstocks, depending on the process. Several approaches are under study but the practicality of these processes is doubtful. In addition, NASA lunar mission planning appears to include approaches and requirements that may reduce or even obviate the potential for ISRU to provide mission benefits. By contrast, Mars ISRU processes are simple to implement and they utilize abundant resources. The potential mission impact is significant. The key to practical Mars ISRU is acquiring water as a feedstock from near-surface deposits that have been discovered from orbit with a neutron spectrometer. The author states that regrettably, several years ago, NASA stopped funding Mars ISRU (which is practical) and is presently funding lunar ISRU (which appears to be impractical). Furthermore, the author observes that NASA's Mars Program does not seem to have any intention of exploiting these water resources for leveraging Mars surface missions.

Chapter 17 by Yuuki Yazawa, Takashi Mikouchi and Hiroshi Takeda shows that the study of Martian meteorites and spacecraft explorations to Mars have provided valuable data on geochemistry and mineralogy of the surface materials. Although almost all Martian meteorites are young igneous rocks, small amounts of alteration products are present, suggesting aqueous alteration by interactions with solutions near the surface. Recent Mars missions further revealed wide distributions of altered materials. They discovered phyllosilicates distributed in the oldest terrains. Sulfates were also detected, suggesting a subsequent acidic environment. Hematite dominated in the areas younger than 3.5 Ga indicating slow superficial weathering without water. Understanding the distribution and nature of these

surface materials is important when utilizing them as available resources and energy sources. Plagioclase feldspar is one of most abundant minerals on Mars, which can be used to produce silicon, aluminum, calcium and oxygen by dissolution. Dissolution experiments by organic acids showed that especially fulvic acid will help with Fe and other micronutrients, by forming organo-metallic complex. Calcium ion left in the solution after precipitation of aluminum, can be used to fix carbon dioxide. Hematite can be an important iron ore if we can extract iron metal. Strong methane release was recently detected in some areas, which can be used as fuel directly or as fuel to generate electricity. Because the main target of Martian exploration is astrobiology research, one should limit the use of microbial means as small as possible. Composts produced from human activities should be kept in small containers, and the produced fertilizers will be applied for soils in a large dome after killing bacteria.

Chapter 18 by Masamichi Yamashita, Hirofumi Hashimoto and Hidenori Wada refers to On-Site Resources Availability for Space Agriculture on Mars. Once the manned Mars mission is extended in the scale of crew number and the period of habitation, closure of the materials recycling becomes indispensable. Space agriculture is the concept of bio-regenerative life support based on the functions of an agro-ecosystem to produce foods, and revitalize air and water. On-site resource availability for both materials and energy is the key issue to realize “more than 100% closure” for increasing life support capability at reduced transport mass from Earth. Nutrition and crop efficiency are major factors to make a decision on crop plant species in space agriculture. The best combination of plant species in this regard is rice, soybean, sweet potato and green-yellow vegetable. Sodium, animal-origin vitamins and fat are supplemented in mineral form and animal origin foods, respectively. Insects and fish are appropriate candidates to fill this need. Planting trees is intended to harvest excess oxygen, and feed their leaves to insect larvae. Creation of amenable environment by wooden interior is another advantage of space agriculture for habitation on Mars. Tri-culture of rice, *Azolla* and loach fish in rice paddies is promising to fill nutritional needs for animal origin foods, and fix nitrogen for fertilizing plants by symbiotic cyanobacteria living in an aquatic fern, *Azolla*. Safe and quick composting human waste to fertilizer can be made by the hyperthermophilic, 80°C to 100°C, aerobic bacterial system. Sodium and potassium management, to solve mismatch between human excreta and plants, is made by the use of either marine algae or halophytes. Scenario of phased development of space agriculture on Mars is discussed in detail including formation of agricultural soil from Martian regolith. Energy resource, mainly depending on solar light, is essential to drive the materials recycling and induce local climate and heat and mass transport in the greenhouse. Terrestrial agriculture is facing severe problems, which resemble the conditions for Martian agriculture. The success of Martian agriculture may help renewing impaired terrestrial agriculture.

Chapter 19 by Yoji Ishikawa refers to the utilization of regolith as a manufacturing construction material on Mars. Humans on Mars require a habitat and air-tightness, volume of living area, radiation protection and cost are the factors to be considered for design and construction of habitat. Especially, In-Situ Resource Utilization (ISRU) will be required to lower cost for construction of the habitat.

Regolith, which is abundant on the Martian surface, can be utilized as it is or as a raw material to manufacture concrete, glass, metal and brick. As one of such manufacturing processes using regolith, a cold press method is proposed as a simple and effective method to make brick by just pressing a mixture of regolith and water without heating. The cold press method uses the adhesive forces of the soil particles. The major players of this method are Smectite type clay mineral and water.蒙脱石矿物是黏土矿物的一种，是这种方法的主要成分。蒙脱石矿物由硅酸盐-铝硅酸盐层组成，具有较大的阳离子交换容量。当与适量的水混合并施加适当的机械压力时，蒙脱石的粘性性质会变成一种将颗粒结合在一起的结合力，使压成的块状坚固而结实。The manufacturing process of using simulated Martian regolith is experimentally examined and the feasibility on Mars is assured. It is concluded that the Martian brick can be used as construction material and shielding block to protect a Martian habitat from radiation. The Martian brick will be one of the plausible materials for constructing habitat on Mars in a future.

Chapter 20 by J. Marvin Herndon shows that in planning for permanent settlements on Mars, the relatively great abundance of iron on its surface has given rise to discussions about steelmaking. The low tensile strength of rock without means for reinforcement places serious construction limitations on structural and building height above ground as well as depth below the ground. Steel, not only for construction, but for farming, mining, transportation, and manufacturing, seems at first glance an obvious choice for Martian industrialization. The author's purpose in this chapter is not to review in detail the extensive, extant literature, but rather, from the perspective of a multi-disciplinary scientist with commercial mining experience, to share some thoughts pertaining to the utilization of two of the major Martian natural resources, iron and carbon dioxide, which are potentially related with respect to proposed metallurgical processes.

Chapter 21 by Mark Berggren, Robert Zubrin, Cherie Wilson, Heather Rose and Stacy Carrera describes the Mars Aqueous Processing System (MAPS), which is a novel technology for producing iron, high-grade metal oxides, and oxygen from Mars soils. The closed-loop process employs sulfuric acid extraction of soil followed by selective recovery of constituents by crystallization or by precipitation after pH adjustment using magnesium hydroxide base. Process reagents can be prepared on Mars from magnesium sulfate salts and water. Metallic iron is produced by hydrogen reduction of iron oxide concentrate, and oxygen is recovered as a byproduct via electrolysis of water produced by reduction. Alumina, magnesia, and calcia products (at greater than 94 percent purity) are useful refractory materials for manufacture of insulation, furnace liners, casting molds, and structural components. MAPS products are also excellent feeds for further refining to produce steel, ceramics, glass, and light metals. Soil extraction residues are elevated in silica content and are suitable for preparation of thermally-fused materials and glass as well as water-based structural materials such as salt- or magnesia-bonded concrete. A subset of the process in which reagents are brought from Earth and recycled with minimal losses is applicable for production of iron, oxygen, and metal oxides on Mars. Engineering designs, including mass and energy balances,

initial equipment and reagent mass, reagent make-up requirements, and estimates of consumables were prepared based on results of laboratory experiments conducted with soil simulants. High process leverage and quick mass breakeven time were projected from the engineering designs. MAPS will enable fabrication of complex structural components for habitats, machinery, spare parts, and other materials to support human self-sufficiency on Mars.

Chapter 22 by Enrique Canessa refers to fundamentals of Human Physical Growth on Mars. A theory based on Newton's law of motion, first introduced for studying body mass index (BMI) on Earth, is here extended to analyze an equivalent physical growth of humans and other organic organisms on Mars. By considering the human body as a dynamic system of variable mass (and volume), analytical biometric relations are deduced for the dependence on age of human height, weight and combinations of them such as the weight-for-height and BMI. The human growth velocity is also derived which also depends on common parameters that could relate to environment, genetics and natural processes known on Earth. Given the likelihood of human colonization in at least few Martian locations with an appropriate ecosystem, physics estimates of growth velocities on Earth and their "predicted" counterparts on Mars are given by taking into account the Martian reduced gravity field. The limiting case of physical growth in a weightless environment is also discussed. Similarly than those on Earth, growth velocities on Mars are found to oscillate down as a function of age to vanish in the adolescence. However, on Mars this may become almost constant for several years prior to puberty. The main consequence of an equivalent human growth and development on Mars is a 60% reduction on the magnitude of the growth velocity when assuming analogous biological parameters. The human height on Mars is also "predicted" to be substantially shortened. Humans on Mars will grow much shorter in stature and will have a growth velocity much slower than on Earth for the whole cycle of life. Nevertheless, to be raised smaller on a weaker gravity field could be an advantage in terms of energy consumption and food needs. Any "flora et fauna" on Mars is foreseen to be similarly influenced.

Chapter 23 by Alexander Bolonkin contains two parts. On a planet without good oxygen atmosphere and solar energy (such as Mars), sustaining human life is very difficult and dangerous, especially during the short sunlit periods of winter when low temperature prevails. To counter these environmental stresses, in the first part of the chapter the author offers an innovative artificial "Evergreen" dome, an inflated hemisphere with interiors continuously providing a terrestrial climate. The "Evergreen" dome theory is developed, substantiated by computations that show it is possible for current technology to construct and heat large enclosed volumes inexpensively. Specifically, a satisfactory result is reached by using magnetically suspended (at high altitude) sunlight reflectors and a special double thin film as a skin, which concentrates solar energy inside the dome while, at the same time, markedly decreasing the heat loss to exterior space. Offered design may be employed for settlements on the Mars and its satellites. In the second part of the chapter one suggests and researches a new revolutionary idea for a closed-loop irrigation method. The author proposes to cover a given site by a thin closed film with controlled heat conductivity and clarity located at an altitude of 5 – 300 m.

The film is supported at altitude by additional atmospheric overpressure and connected to the Martian ground by thin cables. Authors show that this closed dome allows full control of the weather (the day is always fine, the rain (or at least dripping roof condensation) is only at night, no strong winds). The dome (having control of the clarity of film and heat conductivity) converts the cold Martian regions to prosperous regions with a temperate climate. This is a realistic and cheap method of evaporation-economical irrigation and virtual weather control.

Chapter 24 by James Powell, George Maise, Jesse Powell and John Paniagua describes a strategy for Mars exploration and colonization. At its heart is ALPH, a compact, lightweight nuclear powered factory unit that would land on the North Polar Cap of Mars. Using the surface ice and the CO₂/N₂ Martian atmosphere, plus materials from trapped dust layers in the Polar Cap, ALPH would manufacture hundreds of tons of supplies per year, including liquid hydrogen, methane, and oxygen propellants, liquid water, breathable N₂/O₂ air, methanol fuels, plastic construction materials, metallic steel and aluminum, and food. A portion of the manufactured supplies would be stockpiled in sub-surface cavities to be available when astronauts landed to set up a base for the exploration of Mars. Another portion would be lifted by robotic shuttle spacecraft up to Mars orbit. From there, cargo spacecraft would transport the supplies to manned bases on the Moon. Transport of supplies from Mars, although requiring longer trip times, would significantly reduce delta V requirements, as compared to supplying the lunar base from Earth. ALPH unit could support a base of a dozen or more astronauts. As the size of the base grew and new bases were established, more ALPH units could be added, eventually becoming colonies. ALPH will also be able to create comfortable living quarters inside the Polar Ice Cap. Using heat from the nuclear reactor, one can create large cavities of desirable configurations by melting the subsurface ice. A cavity 10 meters below the ice surface provides the same protection from cosmic rays as does the Earth's atmosphere.

Chapter 25 by Colin McInnes investigates large-scale Mars terraforming using orbiting solar reflectors. In particular, a family of displaced planetary orbits is detailed which are suitable for use by solar reflectors as part of a large Mars terraforming effort. The orbits are configured to ensure that solar radiation is continually reflected onto the planetary surface to increase the total solar insolation. A particularly useful stable family of orbits has been identified which lie close to the planetary terminator. For comparison, static equilibria have been considered as potential locations for Mars solar reflectors. However, the reflector orientation required to ensure static equilibrium leads to significant losses of reflected solar radiation which is not intercepted by the planetary body. Static equilibria therefore have serious problems for use by Mars solar reflectors, while displaced orbits can be configured to be Sun-synchronous and passively stable leading to benefits for their use in large-scale terraforming.

Chapter 26 by Viorel Badescu, Dragos Isvoranu and Richard B. Cathcart focuses on intrinsic ecopoiesis and liquid water transportation on Mars. A simple time-dependent model to evaluate the ground surface temperature on the engineered Mars is developed. The daily minimum temperature is well below the water freezing point due to the low thermal inertia of Mars soil. Increasing the

ground surface thermal inertia seems to be necessary for Mars intrinsic ecopoiesis. This can be done either by removing the regolith layer covering the bedrock or by soil compression. A regional and seasonal terraforming of Mars is suggested. Mars' ecopoiesis offers the prospect of an arid exploited Mars formed by processes relevant to neopetrology that still requires domed cities and seasonal cropping and managed irrigation to sustain settled humans. Liquid water may be produced in large quantities in the polar regions, by using orbiting solar mirrors capable of locally melting the polar caps. The chapter examines the prospects for long-distance freshwater transfers by pipelines from the water-ice polar glaciers to the sunnier equatorial region of Mars. The authors closely track the early-20th Century footsteps of Charles Edward Housden (1855-1921). A case study is presented. It refers to a pipeline transporting water from the North polar cap to a 3 million person Martian colony in the equatorial region. The 4023-km long pipeline requires an input power of 58.4 MW to transport 680 kg/s fresh water. A number of 14 power stations are necessary to keep the transport safe. It is shown that hydraulics is cheaper on Mars than on Earth. Due to the difference in the gravitational acceleration, there is about 2.3 times less effort needed to pump water on Mars.

The book allows the reader to acquire a clear understanding of the scientific fundamentals behind specific technologies to be used on Mars in the future. The principal audience consists of researchers (engineers, physicists, biologists) involved or interested in space exploration in general and in Mars exploration in special. Also, the book may be useful for industry developers interested in joining national or international space programs. Finally, it may be used for undergraduate, postgraduate and doctoral teaching in faculties of engineering and natural sciences.

The Editor

Acknowledgments

A critical part of writing any book is the review process, and the authors and editor are very much obliged to the following researchers who patiently helped them read through subsequent chapters and who made valuable suggestions: *Dr. Alexander Blekherman* (New York, USA), *Prof. John Boland* (University of South Australia, Mawson Lakes, Australia), *Dr. Alexander Bolonkin* (CR & C, Brooklyn, USA), *Dr. Douwe Bruinsma* (ProtoneX Technology Corporation, Broomfield, CO, USA), *Prof. Noel Cameron* (Loughborough University, Leicestershire, UK), *Mr. Richard B. Cathcart* (Geographos, Burbank, CA, USA), *Dr. Larry Clark* (Lockheed-Martin, Denver, CO, USA), *Prof. George Cooper* (University of California, Berkeley, CA, USA), *Prof. Alexis De Vos* (Gent University, Belgium), *Dr. Andrew Dombard* (Washington University in St. Louis, USA), *Dr. Dennis A. Edgerley* (Test America, Inc., West Sacramento, CA, USA), *Dr. Michael Efroimsky* (US Naval Observatory, Washington, DC, USA), *Prof. Alex Ellery* (Carleton University, Ottawa, Canada), *Dr. Brian Enke* (Southwest Research Institute, Boulder, USA), *Dr. Martyn Fogg* (University of London, UK), *Dr. Simon D. Fraser* (Graz University of Technology, Austria), *Mr. Joseph Friedlander* (Shave Shomron, Israel), *Dr. Matthias Grott* (Institute of Planetary Research, Berlin, Germany), *Prof. Viktor Hacker* (Graz University of Technology, Austria), *Dr. Brian Haynes* (Los Alamos National Laboratory, NM, USA), *Prof. Barrie W. Jones* (The Open University, Milton Keynes, UK), *Dr. Kenneth Johnson* (Jet Propulsion Laboratory, Pasadena, USA), *Dr. Harry Kambezidis* (National Observatory of Athens, Greece), *Dr. Dimitri Kanevsky* (IBM T.J.Watson Research Center, Yorktown Heights NY, USA), *Dr. Robert Kok* (McGill University, Montreal, Canada), *Dr. Mark Krinker* (City University of New York, USA), *Dr. Maria D. Lane* (University of New Mexico, Albuquerque, USA), *Dr. Hans Ludewig* (Brookhaven National Laboratory, Upton, NY, USA), *Dr. Jesus Polo Martinez* (CIEMAT, Madrid, Spain), *Dr. Chris McKay* (NASA Ames, Moffett Field CA, USA), *Dr. Michael Monsberger* (Arsenal Research, Vienna, Austria), *Dr. Anthony C. Muscatello* (NASA John F. Kennedy Space Center FL, USA), *Dr. Mark Nelson* (Biospheres Foundation, Tucson, AZ, USA), *Prof. Masami Ohnishi* (Kansai University, Japan), *Dr. Nathan Owen-Going* (Going Consulting Inc., Ottawa, Canada), *Prof. Giorgio Pastore* (University of Trieste, Italy), *Prof. Marvin Rowe* (Texas A&M University, USA), *Dr. Vladimir Saly* (Slovak University of

Technology, Bratislava, Slovakia), *Dr. Max Schautz* (European Space Research and Technology Centre, Noordwijk, The Netherlands), *Prof. Enrico Sciubba* (University of Rome, Italy), *Dr. Stewart Sherrit* (Jet Propulsion Laboratory, Pasadena, USA), *Mr. Meyer Steinberg* (HCE, LLC, Oakton, VA, USA), *Dr. Brad Sutter* (NASA Johnson Space Center, Houston, TX, USA), *Dr. G. Jeffrey Taylor* (University of Hawai, Honolulu, USA), *Prof. Masamichi Yamashita* (Institute of Space and Astronautical Science, Kanagawa, Japan), *Dr. Mark Wieczorek* (Institut de Physique du Globe, Paris, France), *Dr. Paul Wooster* (SpaceX, Los Angeles, CA, USA), *Dr. Xiaoling Yang* (The University of Illinois at Urbana-Champaign, USA).

The editor, furthermore, owes a debt of gratitude to all authors. Collaborating with these stimulating colleagues has been a privilege and a very satisfying experience.

Contents

Chapter 1

Power System Options for Mars Surface Exploration: Past, Present and Future	1
<i>Simon D. Fraser</i>	

Chapter 2

Available Solar Energy and Weather Forecasting on Mars Surface	25
<i>Viorel Badescu</i>	

Chapter 3

Weather Influence on PV Solar Cells Operation on Mars	67
<i>Viorel Badescu</i>	

Chapter 4

Electric Power Generation on Mars Using Photovoltaic Helium Balloons	83
<i>Pini Gurfil, Joseph Cory</i>	

Chapter 5

Weather Influence on Solar Thermal Power Plants Operation on Mars	99
<i>Viorel Badescu</i>	

Chapter 6

Fuel Cell Power System Options for Mars Surface Mission Elements	139
<i>Simon D. Fraser</i>	

Chapter 7

- The Evolution of Nuclear Power Generation for Mars 175
Liviu Popa-Simil

Chapter 8

- Distributed Power Sources for Mars 213
George H. Miley, Xiaoling Yang, Eric Rice

Chapter 9

- Advanced Nuclear Compact Structures for Power
Generation on Mars 241
Liviu Popa-Simil

Chapter 10

- New Solutions for Nuclear Energy and Flights on Mars 287
Alexander A. Bolonkin

Chapter 11

- Geothermal Energy on Mars 331
Paul Morgan

Chapter 12

- Comparative Analysis of Power System Architectures: The
Case of Human Mars Surface Missions 351
Matthew Silver, Wilfried Hofstetter, Chase Cooper, Jeffrey Hoffman

Chapter 13

- Economics of Energy on Mars 369
*Joseph E. Palaia, Mark S. Homnick, Frank Crossman,
Alexander Stimpson, John Truett*

Chapter 14

- Tumbleweed: A New Paradigm for Surveying the Surface
of Mars for In-situ Resources 401
*Kimberly R. Kuhlman, Alberto Behar, Jack Jones, Penelope Boston,
Jeffrey Antol, Gregory Hajos, Warren Kelliher, Max Coleman,
Ronald Crawford, Lynn Rothschild, Martin Buehler, Greg Bearman,
Daniel W. Wilson*

Chapter 15

- Drilling and Excavation for Construction and In-Situ Resource Utilization** 431
Kris Zacny, Yoseph Bar-Cohen

Chapter 16

- In Situ Utilization of Indigenous Resources** 461
Donald Rapp

Chapter 17

- Available Resources and Energy Sources from Mars Rock and Soil** 483
Yuuki Yazawa, Takashi Mikouchi, Hiroshi Takeda

Chapter 18

- On-Site Resources Availability for Space Agriculture on Mars** 517
Masamichi Yamashita, Hirofumi Hashimoto, Hidenori Wada

Chapter 19

- Utilization of Regolith for Manufacturing Construction Material on Mars** 543
Yoji Ishikawa

Chapter 20

- Perspectives on the Utilization of Martian Iron and Carbon Dioxide** 551
J. Marvin Herndon

Chapter 21

- Mars Aqueous Processing System** 563
Mark Berggren, Robert Zubrin, Cherie Wilson, Heather Rose, Stacy Carrera

Chapter 22

- Fundamentals of Human Physical Growth on Mars** 587
Enrique Canessa

Chapter 23

- Artificial Environments on Mars** 599
Alexander A. Bolonkin

Chapter 24

- Self-sustaining Martian Colonies Utilizing the North Polar
Ice Cap and the Atmosphere** 627
James Powell, George Maise, Jesse Powell, John Paniagua

Chapter 25

- Mars Climate Engineering Using Orbiting Solar
Reflectors** 645
Colin R. McInnes

Chapter 26

- Ecopoiesis and Liquid Water Transportation on Mars** 661
Viorel Badescu, Dragos Isvoranu, Richard B. Cathcart

- Author Index** 683

- Index** 685

List of Contributors

Jeffrey Antol

NASA Langley Research Center, 11
Langley Blvd , Hampton, VA 23681,
USA
Jeffrey.Antol-1@nasa.gov

Viorel Badescu

Candida Oancea Institute, Polytechnic
University of Bucharest,
Spl. Independentei 313,
Bucharest 060042,
Romania
badescu@theta.termo.pub.ro

Yoseph Bar-Cohen

Jet Propulsion Laboratory, California
Institute of Technology, Oak Grove Dr,
Pasadena, CA 91109, USA
yosi@jpl.nasa.gov

Greg Bearman

Snapshot Spectra, 974 East Elizabeth
St., Pasadena, CA, 91104, USA
gregb@snapshotspectra.com

Alberto Behar

Jet Propulsion Laboratory,
California Institute of Technology, Oak
Grove Dr, Pasadena, CA 91109, USA
Aberto.E.Behar@jpl.nasa.gov

Mark Berggren

Pioneer Astronautics, 11111 W. 8th
Ave, Unit A, Lakewood, CO 80215
mberggren@pioneerastro.com

Alexander A. Bolonkin

CR & C, 1310 Avenue R, #6-F,
Brooklyn, NY 11229, USA
aBolonkin@juno.com,
aBolonkin@gmail.com

Penelope Boston

Earth & Environmental Science Dept.,
New Mexico Tech, 801 Leroy Place,
Socorro, New Mexico 87801, USA
pboston@nmt.edu

Martin Buehler

Decagon Devices, 2365 NE Hopkins
Court , Pullman, WA, USA 99163
mbuehler@decagon.com

Enrique Canessa

Abdus Salam International Centre for
Theoretical Physics, Strada Costiera 11,
34151 Trieste, Italy
canessae@ictp.it

Stacy Carrera

Pioneer Astronautics, 11111 W.
8th Ave, Unit A, Lakewood,
CO 80215
scarrera@pioneerastro.com

Richard B. Cathcart

Geographos, 1300 West Olive Avenue,
Suite M, Burbank,
California 91506-2225, USA
rbcathcart@gmail.com

Max Coleman

Jet Propulsion Laboratory, California
Institute of Technology, Oak Grove Dr,
Pasadena, CA 91109, USA
Max.Coleman@jpl.nasa.gov

Chase Cooper

Massachusetts Institute of Technology,
77 Massachusetts Av., Department of
Aeronautics and Astronautics,
Cambridge, MA 02139-4307, USA
chase@mit.edu

Joseph Cory

Geotectura, 17 Kiryat Seffer St.,
Haifa, 34676 Israel
info@geotectura.com,
geotectura@gmail.com

Ronald Crawford

Environmental Biotechnology Institute,
University of Idaho, Moscow, ID,
83844, USA
crawford@uidaho.edu

Frank Crossman

4Frontiers Corporation, 4315
Sanddollar Court, New Port Richey,
FL 34652, USA
fcrossman@
4FrontiersCorp.com

Simon D. Fraser

Christian Doppler Laboratory for Fuel
Cell Systems, Graz University of
Technology, Steyrergasse 21, 8010
Graz, Austria
simon.fraser@gmx.at,
simon.fraser@tugraz.at

Pini Gurfil

Faculty of Aerospace Engineering,
Technion - Israel Institute of
Technology, Technion City,
Haifa 32000, Israel.
pgurfil@technion.ac.il

Gregory Hajos

NASA Langley Research Center, 11
Langley Blvd , Hampton, VA 23681,
USA
Gregory.A.Hajos@nasa.gov

Hirofumi Hashimoto

Institute of Space and Astronautical
Science, Japan Aerospace Exploration
Agency (JAXA), Yoshinodai,
Sagamihara, Kanagawa 229-8510,
Japan
hashimoto@surc.isas.jaxa.jp

J. Marvin Herndon

Transdyne Corporation, 11044 Red
Rock Drive, San Diego, CA 92131
USA
mherndon@san.rr.com

Jeffrey Hoffman

Massachusetts Institute of Technology,
Department of Aeronautics and
Astronautics, 77 Massachusetts Av.,

Cambridge, MA 02139-4307, USA
jhoffma1@mit.edu

Wilfried Hofstetter
Massachusetts Institute of Technology,
Department of Aeronautics and
Astronautics, 77 Massachusetts Av.,
Cambridge, MA 02139-4307, USA
wilfried.hofstetter@gmail.com;
wk_hof@mit.edu

Mark S. Homnick
4Frontiers Corporation, 4315
Sanddollar Court, New Port Richey, FL
34652, USA
mhomnick@4FrontiersCorp.com

Yoji Ishikawa
Environmental Technology Department,
Obayashi Corporation, Shinagawa
Intercity Tower B, 2-15-2,
Konan,
Minato-ku, Tokyo 108-8502, Japan
ishikawa.yojo@
obayashi.co.jp

Dragos Isvoraru
Faculty of Mechanical Engineering,
Polytechnic University of Bucharest,
Spl. Independentei 313, Bucharest
060042, Romania
ddisvoranu@gmail.com

Jack Jones
Jet Propulsion Laboratory, California
Institute of Technology, Oak Grove Dr,
Pasadena, CA 91109, USA
Jack.A.Jones@jpl.nasa.gov

Warren Kelliher
NASA Langley Research Center, 11
Langley Blvd , Hampton, VA 23681,
USA
Waren.C.Kelliher@nasa.gov

Kimberly R. Kuhlman
Planetary Science Institute, 1700 Fort
Lowell Road, Suite 106, Tucson AZ
USA, 81719
kim@psi.edu

George Maise
Plus Ultra Technologies, Inc.,
Shoreham, N.Y. 11786, USA
maise@optonline.net

Colin R. McInnes
Advanced Space Concepts Laboratory,
Department of Mechanical
Engineering, University of Strathclyde,
Glasgow, G1 1XJ, Scotland, UK
colin.mcinnnes@strath.ac.uk

Takashi Mikouchi
Department of Earth and Planetary
Science, University of Tokyo, 7-3-1
Hongo, Bunkyo-ku, Tokyo 113-0033,
Japan
mikouchi@
eps.s.u-tokyo.ac.jp

George H. Miley
Department of Nuclear, Plasma and
Radiological Engineering,
University of Illinois, Urbana, IL 61801
USA
ghmiley@uiuc.edu

Paul Morgan
Department of Geology, Northern
Arizona University, Flagstaff, AZ
86011, USA
paul.morgan@nau.edu
and
Colorado Geological Survey
1313 Sherman St., Rm. 715, Denver,
CO 80203, USA
paul.morgan@nau.edu

Joseph E. Palaia IV
4Frontiers Corporation, 4315
Sanddollar Court, New Port Richey,
FL 34652, USA
joe@4FrontiersCorp.com

John Paniagua

Plus Ultra Technologies, Inc.,
Shoreham, N.Y. 11786, USA
jpania@yahoo.com

Liviu Popa-Simil

LAVM Co., 3213-C Walnut St. Los
Alamos, NM 87544-2092, USA
lpopas@ieee.org,
lpopas@yahoo.com

James Powell

Plus Ultra Technologies, Inc.,
Shoreham, N.Y. 11786, USA
jamesrpowell132@gmail.com

Jesse Powell

Plus Ultra Technologies, Inc.,
Shoreham, N.Y. 11786, USA
jpowell@coast.ucsd.edu

Donald Rapp

Viterbi School of Engineering,
University of Southern California,
1445 Indiana Avenue, South Pasadena,
CA 91030, USA
drdrapp@earthlink.net

Eric Rice

Orbital Technologies Corp., Madison,
WI 53717, USA
ricee@orbitec.com

Heather Rose

Pioneer Astronautics, 11111 W. 8th
Ave, Unit A, Lakewood, CO 80215
hrose@pioneerastro.com

Lynn Rothschild

NASA-Ames Research Center, Mail
Stop 239-20, Moffett Field,
CA 94035-1000, USA
Lynn.J.Rothschild@nasa.gov

Mathew Silver

IntAct Labs LLC, 247 Third Street,
Cambridge, MA, 02142;
Massachusetts Institute of Technology,
Engineering Systems Division, 77
Massachusetts Av., Cambridge,
MA 02139-4307, USA
silver@intactlabs.com

Alexander Stimpson

4Frontiers Corporation, 4315
Sanddollar Court, New Port Richey,
FL 34652, USA
astimpson@4FrontiersCorp.com

Hiroshi Takeda

Department of Life and Environmental
Sciences, Chiba Institute of
Technology, 2-17-1 Tsudanuma,
Narashino, Chiba 275-0016, Japan
takeda.hiroshi@it-chiba.ac.jp

John Truett

4Frontiers Corporation, 4315
Sanddollar Court, New Port Richey,
FL 34652, USA
jtruett@4FrontiersCorp.com

Hidegori Wada

Institute of Space and Astronautical
Science, Japan Aerospace Exploration
Agency (JAXA), Yoshinodai,
Sagamihara, Kanagawa 229-8510,
Japan
YFA09063@nifty.ne.jp

Cherie Wilson

Pioneer Astronautics, 11111 W. 8th
Ave, Unit A, Lakewood, CO 80215
cwilson@pioneerastro.com

Daniel W. Wilson

Jet Propulsion Laboratory,

California Institute of Technology, Oak
Grove Dr, Pasadena, CA 91109, USA
daniel.w.wilson@jpl.nasa.gov

Masamichi Yamashita
Institute of Space and Astronautical
Science, Japan Aerospace Exploration
Agency (JAXA), Yoshinodai,
Sagamihara, Kanagawa 229-8510,
Japan
yamashita@surc.isas.jaxa.jp

Xiaoling Yang
Department of Nuclear, Plasma and
Radiological Engineering,
University of Illinois, Urbana, IL 61801
USA
xlyang@illinois.edu

Yuuki Yazawa
Department of Life and Environmental
Sciences, Chiba Institute of
Technology, 2-17-1 Tsudanuma,
Narashino, Chiba 275-0016, Japan
yuuki.yazawa@it-chiba.ac.jp

Kris Zacny
Honeybee Robotics Spacecraft
Mechanisms Corporation, 460 W. 34th
Street / New York, NY 10001, USA
zacny@honeybeerobotics.com

Robert Zubrin
Pioneer Astronautics, 11111 W. 8th
Ave, Unit A, Lakewood, CO 80215
rzubrin@pioneerastro.com

Chapter 1

Power System Options for Mars Surface Exploration: Past, Present and Future

Simon D. Fraser

Graz University of Technology, Austria

1.1 Timeline of Mars Exploration

Within this chapter, the power systems developed and applied in Mars surface exploration applications are discussed in four different time domains:

- The early Mars exploration phase: 1960-1990
- The recent Mars exploration phase: 1990-2000
- The present Mars exploration phase: 2000-2010
- Future Mars exploration: 2010+

At first, an overview of the Mars missions conducted within the different time domains is given. All Mars missions launched within the investigated time domains will be presented; this includes fly-by probes, orbiters, and landers.

The focus of this chapter is on power system options for Mars surface exploration applications. Only the power systems of selected landers are therefore subsequently discussed in detail, presenting purpose, design, and limitations of the different power systems configurations. Power systems for probes designed for fly-by or orbiting missions are subject to completely different requirements and operating conditions. These power system are therefore not of relevance for Mars surface operations and are not discussed within this chapter.

1.2 The Early Mars Exploration Phase: 1960-1990

Mars was the first goal in interplanetary exploration when the Soviet Union launched Marsnik 1 in 1960. Unfortunately, Marsnik 1 was lost due to a launch failure, and so were many of the Soviet probes that followed Marsnik 1 in the early 1960s.

The first fly-by was therefore not achieved by one of the Soviet probes, but by the American Mariner 4 spacecraft in mid-1965. The first successfully orbit was achieved by Mariner 9 in 1971, and the first landing – although only being a partial success as communication ceased soon after the landing – was made by the Soviet Mars 3 Lander. The two missions commonly considered as the real start of scientific surface-based Mars exploration are the American Viking 1 & 2 probes, landed on Mars in 1976 and operated in the years following. Viking 1 & 2 were followed by the two Phobos missions of the Soviet Union more than a decade later. Both Phobos probes were not able to deploy their landers, though.

A timeline of the early Mars exploration phase is provided in Table 1.1.

Table 1.1. Timeline of the early Mars exploration phase

Launch Date	Mission Name*	Outcome
October 10, 1960	Marsnik 1 / Mars 1960A	Launch failure
October 14, 1960	Marsnik 2 / Mars 1960B	Launch failure
October 24, 1962	Sputnik 22 / Mars 1962A	Launch failure
November 1, 1962	Mars 1	Contact lost en route
November 4, 1962	Sputnik 24 / Mars 1962B	Failed to leave Earth orbit
June 4, 1964	Zond / Mars 1964A	Launch failure
November 5, 1964	Mariner 3	Failed to reach Mars
November 28, 1964	Mariner 4	Successful fly-by
November 30, 1964	Zond 2	Contact lost en route
February 25, 1969	Mariner 6	Successful fly-by
March 27, 1969	Mariner 7	Successful fly-by
March 27, 1969	Mars 1969A	Launch failure
April 2, 1969	Mars 1969B	Launch failure
May 5, 1971	Cosmos 419 / Mars 1971C	Launch failure
May 8, 1971	Mariner 8	Launch failure
May 19, 1971	Mars 2	Orbiter: success Lander: crashed
May 28, 1971	Mars 3	Orbiter successful Lander: partial failure
May 30, 1971	Mariner 9	Successful orbit
July 21, 1973	Mars 4	Entry into orbit failed, fly-by
July 25, 1973	Mars 5	Successfully entered orbit
August 5, 1973	Mars 6	Contact lost during landing
August 9, 1973	Mars 7	Landing probe missed planet
August 20, 1975	Viking 1	Orbiter: success Lander: success
September 9, 1975	Viking 2	Orbiter: success Lander: success
July 7, 1988	Phobos 1	Contact lost en route
July 12, 1988	Phobos 2	Orbiter: partial success Lander: failure

*Grey background indicates surface landing elements

As the two Viking Landers were the only systems successfully operated on the surface of Mars, only the power systems utilized with these two landers will be discussed in the following.

1.2.1 Viking Lander Power System

NASA's highly successful Viking program consisted of a pair of identical Orbiter/Lander systems launched in the mid-1970s. The launch mass of the Landers was just over 650 kg, including 85 kg of hydrazine propellant for terminal descent and landing and approximately 90 kg of scientific equipment.

After soft-landing on Mars in June/July 1976, the Landers were operated for more than 3.5 respectively 6 years (Viking 1/2). In this time, the Viking Landers performed scientific experiments studying biology, chemical composition, meteorology and physical properties of the Martian surface and atmosphere. Due to the long operational lifetime of the Viking Landers, the measured data covers more than a full Martian year and therefore provides a valuable insight into seasonal variations of the Martian weather. An image of a Viking Lander is shown in Fig. 1.1.

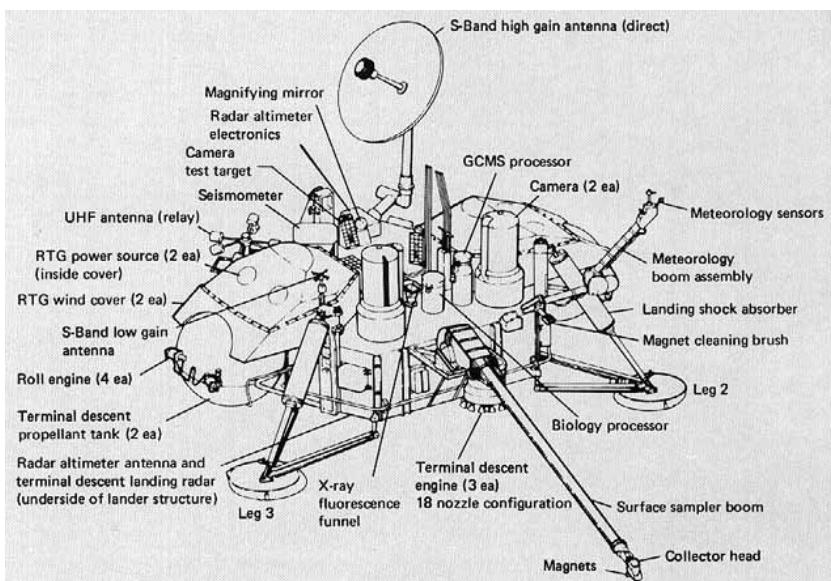


Fig. 1.1 Overview of the Viking Lander. The RTGs are located on the left hand side on top of the spherical propellant tank under a wind cover (Image courtesy of NASA).

Each of the two Viking Landers was powered by two radioisotope thermoelectric generators (RTGs) containing plutonium 238. The RTGs were installed into the Lander base on top of one of the two terminal descent propellant tanks. The RTGs were covered by a wind shield, as can be seen in Fig. 1.1.

Each of the two generators was 28 cm tall, 58 cm in diameter and had a mass of 13.6 kg. The continuous electrical power output of each generator was 30 W. Four wet-cell sealed nickel-cadmium (NiCd) rechargeable batteries (8 Ah, 28 V) were also installed to handle peak power loads in excess of the continuous RTG output power (NASA Viking 2009).

1.3 The Recent Mars Exploration Phase: 1990-2000

The 1990s saw a renewed interest in Mars exploration. After losing Mars Observer in 1993 just before arrival, NASA successfully returned to Mars with the Mars Global Surveyor probe in 1997 after a 20-year absence.

With the Russian Mars 96 being lost due to a launch failure, the Japanese Nozomi-B being lost en route due to electrical problems, the American Mars Climate Orbiter crash-landing due to the infamous metric-imperial mix-up, and the Mars Polar Lander being lost just before arrival, the 1997 Mars Pathfinder mission remains the only full success of the 1990 Mars exploration phase besides the Mars Global Surveyor.

A timeline of the 1990s Mars exploration phase is provided in Table 1.2.

Table 1.2 Timeline of the recent Mars exploration phase

Launch Date	Mission Name*	Outcome
September 25, 1992	Mars Observer	Contact lost en route
November 7, 1996	Mars Global Surveyor	Success
November 16, 1996	Mars 96	Launch failure
December 4, 1996	Mars Pathfinder	Success
July 3, 1998	Nozomi-B	Failure
December 11, 1998	Mars Climate Orbiter	Crashed onto Mars
January 3, 1999	Mars Polar Lander	Contact lost en route

*Grey background indicates surface landing elements

Power systems developed for the Mars Pathfinder Lander and the Microrover Sojourner as well as for the Mars Polar Lander were different from the approach previously applied with the Viking missions. Now, radioisotope units were only included for heating purposes and electrical power was rather provided by a combination of high-performance photovoltaic panels and primary or secondary battery modules. This change in power system design was not primarily due to advances in non-nuclear power system technology, but rather due to the general move towards applying faster, better, and cheaper missions introduced within the NASA Mars Robotic Exploration Program. In fact, the Mars Robotic Exploration Program was even precluded from using any nuclear power (Shirley 1999).

1.3.1 Mars 96 Mission Power Systems

The Mars 96 mission, launched by Russia in 1996, was a continuation of the long series of probes previously developed in the Soviet Union. Mars 96 actually was

an improved version of the two 1988 Phobos missions, which both had ultimately failed in operation. Unfortunately, the fourth stage of the launcher malfunctioned and Mars 96 was lost, too.

The Mars 96 mission consisted of an orbiter that carried two surface stations, the so-called Small Stations, and two surface Penetrators. The power systems of these two types of landing elements of the Mars 96 mission are briefly discussed in the following.

1.3.1.1 Mars 96 Small Station Power Systems

The two Small Stations were to soft-land on Mars and to conduct an ambitious scientific program. The Stations accommodated a wide array of instruments for investigating the descent phase and the Martian surface. The total mass of a landed Small Station was only 32 kg including airbags; each Station carried 12 kg of payload mass.

The power systems of the Small Stations were composed of two different battery modules and two RTGs. The main battery system was a rechargeable nickel-cadmium (NiCd) battery module. This battery module would have been fully charged prior to separation from the Orbiter. An additional Lithium-technology battery module was included for the decent phase.

Once landed on Mars, the primary power would have come from two RTGs. Two additional radioisotope heater units (RHUs) were installed to warm the Stations once landed on Mars. The radioisotope systems each provided a thermal power of 8.5 W and weighed 490 g (RTG) and 185 g (RHU). The nominal electrical output power of each RTG was 220 ± 10 mW. Thus, a total of 440 mW would have been available to power operation of the Small Station. In total, each small station carried some 120 g of plutonium 238 for heating and power generation purposes split up into the two RTGs and the two RHUs (Linkin et al. 1998).

The Small Stations would have been operated in standby mode most of the time. During standby mode, some 95% of the RTG power would have been available for charging the onboard storage battery module. The intermittent storage in the battery module would then provide sufficient energy for the active operation phases.

1.3.1.2 Mars 96 Penetrator Power System

Penetrating probes (*Penetrators*) were included with the Mars 96 mission as means of investigating the particularly interesting Martian subsurface region. The Penetrator probes were to strike the surface, and the forebody was to delve 5-6 m into the surface while the afterbody should remain on the surface connected to the forebody by wires.

The power system was installed into the afterbody of the probes and consisted of an RTG system with an electrical output power of 0.5 W and a Lithium-technology battery module with a capacity of 150 Wh. This combined power system had a mass of only 3.5 kg and thus only contributed a very small share to the total Penetrator mass of 125 kg (Surkov and Kremnev 1998).

1.3.2 Mars Pathfinder Mission Power Systems

The Mars Pathfinder mission is often considered as being the most important Mars surface mission since the mid-1970 Viking program. Although being designed as a low-cost mission providing a proof-of-concept to many technologies considered relevant for future Mars surface exploration, the Pathfinder mission nevertheless turned out to be a huge success not only with respect to the scientific output, but also with respect to the public response received.

The Pathfinder mission consisted of a stationary lander that carried a small six-wheeled microrover called Sojourner. Sojourner was the first in a series of rovers sent to Mars.

1.3.2.1 Mars Pathfinder Lander Power System

The Mars Pathfinder Lander carried scientific equipment, primarily the Imager, a magnetometer and a anemometer as well as atmospheric and meteorological sensors. In addition, the lander also relayed transmissions to and from the microrover Sojourner. Sojourner was fixed onto the Lander during transit and landing, and successfully rolled onto the surface of Mars on mission sol 2 (sol is a Martian day).

The Lander had three solar panels hinged to the main body carrying the Micro-rover. Unfolded, the total solar array area was 2.8 m^2 . The solar panels could supply up to 1,200 Wh of electrical energy per day on clear sky conditions at the beginning of lifetime. For nighttime operations, the lander also had a rechargeable silver zinc (AgZn) battery module with a beginning-of-life capacity of more than 40 Ah (NASA Pathfinder 1997).

1.3.2.2 Mars Pathfinder Microrover Sojourner Power System

Sojourner was the first rover to be successfully operated on the surface of Mars, providing scientists with a limited but nevertheless very welcome degree of freedom in closely analyzing rock and soil samples surrounding the stationary Pathfinder Lander. Surface mobility was achieved by remotely controlling the rover from the Earth utilizing an onboard camera-based autonomous control system for assistance. A wide array of scientific instruments was installed into the small rover, including, among others, an imaging system, an Alpha Proton X-ray Spectrometer, and the Wheel Abrasion Experiment. The total rover mass was just over 10.5 kg, and the six-wheeled chassis had a total length of about 65 cm. An image of Sojourner actually taken on the surface of Mars is shown in Fig. 1.2.

Sojourner was powered by a single 0.2 m^2 Gallium Arsenide on Germanium (GaAs/Ge) solar panel installed on top of the central rover body, providing a peak electrical power output of 16 W at the Ares Vallis landing site conditions. The solar panel provided sufficient electrical energy to operate the rover for several hours per day, even in dust storm conditions.

In addition, three Lithium Thionyl Chloride (LiSOCl_2) D-cell-sized batteries were installed. Lithium Thionyl Chloride chemistry was chosen because these primary (i.e. non-rechargeable) batteries offer a very high energy density, an extremely long shelf life, a wide operating temperature range, and excellent durability. The decision to use a primary instead of a secondary battery module prevents



Fig. 1.2 Sojourner on Mars on sol 22 (Image courtesy of NASA/JPL)

recharging procedures, but keeps the overall power system less complex and lighter, as the power electronics for the recharging procedure do not need to be included. In addition, primary batteries can provide a higher energy density than secondary (i.e. rechargeable) batteries.

The batteries were enclosed in the rover's thermally protected Warm Electronics Box (WEB). The WEB was thermally insulated by silica aerogel and designed to protect the interior from the cold ambient temperatures present during nighttime. Three radioisotope heater units (RHUs), each containing about 2.5 g of plutonium-238, were installed to provide about one Watt of thermal power each, and to keep the WEB warm and thus to ensure a survival of the batteries in the harsh environmental conditions (Stone 1996).

Under normal operating conditions, the solar panel was the primary source of energy. The batteries were included as backup in case the solar panels were damaged during transport or failed upon arrival on Mars. In case the solar panels would have been damaged during the landing, the batteries would have enabled a full seven days of operation. As this was not the case, the batteries could be used to power experiments at night and to support in peak power applications. As with a complete failure of the solar panels, also a complete failure of the battery module would not have stopped the mission as solar-based operation would still have been possible (Stone 1996).

1.3.3 Mars Polar Lander Mission Power Systems

The Mars Polar Lander mission was an element of the NASA Mars Surveyor '98 program consisting of the Mars Climate Orbiter and the Mars Polar Lander. The Mars Polar Lander also carried the Deep Space 2 surface-penetrator mission to Mars. The Mars Polar Lander mission was to study the Martian weather, climate,

water and carbon dioxide levels in the southern polar *Planum Australis* region, located less than 1,000 km from the Martian south pole.

Unfortunately, the mission failed as the communication could not be regained after the atmospheric entry of the probe. Both, Mars Polar Lander and the Deep Space 2 probe, were lost.

1.3.3.1 Mars Polar Lander Power System

The Mars Polar Lander had a dry mass of 512 kg and was, similar to the Mars Pathfinder Lander, also powered by a combination of solar panels and rechargeable batteries.

The Lander could draw power from a total of six solar panels. The four larger panels were arranged in a pair of gull-shaped wings, extending about 3.6 m on either side of the lander. Two smaller panels were attached to the side of the lander to increase the total power output.

A rechargeable 16 Ah nickel-hydrogen (Ni-H₂) battery was installed into the central electronics enclosure and powered electric heaters to prevent the inside of the enclosure from getting too cold during nighttime. The lifetime of the Lander would therefore have depended on the lifetime of the battery, as the battery would finally not have enough power to prevent the lander from freezing (NASA 1998 Mars Missions 1998).

1.3.3.2 Deep Space 2 Power System

The Deep Space 2 probe was intended to separate from the Polar Lander shortly before entering the atmosphere. The probe was to dive towards the surface and bury up to about 1 meter in the Martian underground upon impact. This procedure was not only intended as proof-of-concept for high-G crashing technology (the surface impact speed was estimated in the order of 200 m/s), but also to search for water ice underneath the Martin surface. Mission and design of the Deep Space 2 probe was therefore similar to the Penetrators developed for the Russian Mars 96 mission.

The batteries installed within the Deep Space 2 probe were designed to survive the high-G impact, and to provide 600 mAh of electrical energy at temperatures as low as -80°C. As with Sojourner, a primary Lithium Thionyl Chloride (LiSOCl₂) cell chemistry was chosen. The cells used a lithium tetrachlorogallate salt instead of the more conventional lithium aluminum chloride salt to improve low-temperature performance and to reduce voltage delays. Each probe had two batteries composed of four D-sized cells weighing less than 40 grams each (NASA 1998 Mars Missions 1998).

1.4 The Present Mars Exploration Phase: 2000-2010

The present Mars exploration phase is a continuation of the technological and scientific approaches developed and applied with the 1990s missions. Orbiting probes such as Mars Odyssey, Mars Express, Rosetta, and the Reconnaissance Orbiter have increased - or still are increasing - the level of knowledge about Mars

by applying state-of-the-art remote sensing equipment. In addition, three successful landing missions were conducted in the new millennium up to date: the two Mars Exploration Rover missions and the stationary Phoenix Lander. Beagle 2, the fourth lander sent to Mars in the new millennium up to date was unfortunately lost; ESA's Mars Express Orbiter, previously carrying Beagle 2 to Mars, however, was a success.

A timeline of the present Mars exploration phase is provided in Table 1.3.

Table 1.3 Timeline of the present Mars exploration phase

Launch Date	Mission Name*	Outcome
April 7, 2001	Mars Odyssey	Success
June 2, 2003	Mars Express	Orbiter: success Lander Beagle 2: failure
June 10, 2003	MER A: Spirit	Success
July 7, 2003	MER B: Opportunity	Success
Mars 2, 2004	Rosetta	Success
August 12, 2005	Mars Reconnaissance Orbiter	Success
August 4, 2007	Phoenix	Success

*Grey background indicates surface landing elements

1.4.1 Mars Express/Beagle 2 Power System

Beagle 2 was a British landing probe included with ESA's 2003 Mars Express mission (Schmidt 2003). Beagle 2 had the shape of a bowl measuring approximately 1 x 0.25 m (diameter x depth). The cover of the Lander was hinged to the bowl-like body and was to unfold once successfully soft-landed on the surface of Mars to uncover an antenna, a robotic arm, and the scientific equipment. Unfortunately, contact to the probe was lost upon its separation from the Mars Express Orbiter and the probe was lost.

Beagle 2 was powered by a 42-cell Lithium-ion (Li-ion) battery module and solar cells. The solar cells were mounted on four deployable panels that were to fold out from within the Lander as part of the automatic landing sequence. The total solar array area was about one square metre.

As with other Mars surface probes, power system and batteries were kept in a warm compartment. This compartment insulated the interior from the harsh ambient climate. In addition, thermally absorbing materials were applied to transfer the heat of the sun into the battery compartment during daytime to ensure that the power system did not cool down too low during nighttime. The solar cells were Gallium Arsenide using Germanium substrate (GaAs/Ge) (Beagle 2 homepage 2004).

1.4.2 Mars Exploration Rover Power System

The Mars Exploration Rover (MER) missions continue NASA's successful surface exploration missions - the mid-1970 Viking missions and the mid-1990 Mars

Pathfinder missions - with two surface rovers. The objective of the MER missions was to explore the Martian surface and geology with two mobile systems, searching and characterizing a wide range of rocks and soils that may hold clues to water activity on Mars in the past. Two identical systems were deployed in 2003: the rovers Spirit (MER-A) and Opportunity (MER-B). An image of a Mars Exploration Rover is shown in Fig. 1.3.



Fig. 1.3 Artist's image of a Mars Exploration Rover (Image courtesy of NASA)

Each of the twin rovers is approximately 1.6 m long and weighs 174 kg. The central element of the rovers – again in analogy to Sojourner – is a warm electronics box (WEB). On top of the WEB there is a triangular surface called the rover equipment deck. Antennas, camera mast, and solar panels are installed onto the deck. Additional solar panels are attached to the edges of the deck by hinges, allowing the panels to be folded up during launch and transit.

The solar panels have a total area of 1.3 m². Three-layer solar cells are used, with the layers consisting of Gallium Indium Phosphorus, Gallium Arsenide and Germanium (GaInP/GaAs/Ge). The solar array can produce almost 900 Wh of energy per sol at beginning-of-life, clear sky conditions, and without dust deposition. The actual power output proved to be considerably smaller, and strongly depends on dust deposition on the panels.

Two Lithium-ion (Li-ion) battery modules are installed inside the WEB. These battery modules were included to aid in the launch, correct anomalies during cruise, and support surface operations. In later phases of the mission, the rechargeable batteries are designed to support the primary power source for the MER Rovers, the solar arrays, and to support nighttime experiments (Ratnakumar et al. 2003, Ratnakumar et al. 2006).

The WEB is heated by electric heaters, waste heat generated by electronic components as well as eight RHUs. Each RHU contains about 2.7 g of plutonium dioxide and produces about one Watt of thermal power. Sojourner, in comparison,

had three RHUs with comparable thermal output power. The RHUs considerably reduce the electric heating load and save battery power particularly in cold winter nights (NASA MER 2003). In addition, the combination of a WEB and RHUs prevents the batteries from being operated at extreme temperatures. Even the -20 to -30°C operation required by lander batteries are very challenging for state-of-the-art battery technology. Advances in battery technology may even enable operation at -40°C in the near future without a significant penalty to the room temperature specific energy, safety, stability or cycle life (Ratnakumar et al. 2000). A heated thermal enclosure nevertheless improves performance and lifetime of batteries, electronics, and scientific equipment to such an extent, that the mass penalty of the enclosure is acceptable.

1.4.3 Phoenix Power System

The Phoenix mission deployed a small robotic Lander in the Green Valley of Vastitas Borealis near the Martin north pole on May 25, 2008. The mission essentially had two goals: to investigate the geologic history of water and to evaluate a planetary habitability in the ice-soil boundary. The mission was chosen to be a stationary Lander rather than a rover in order to keep the mission costs low and to be able to increase the scientific payload by avoiding a heavy locomotion system. An image of the Phoenix Lander is shown in Fig. 1.4.

The Phoenix Lander power system consisted of a two-wing solar array. The two wings had the shape of a decagon and a functional surface of 4.2 m^2 installed onto a flexible, lightweight substrate.

Two rechargeable Lithium-ion (Li-ion) batteries with a capacity of 25 Ah were installed to provide a limited energy storage capacity (NASA Phoenix 2007). The batteries were provided by Yardney Technical Products, who had also previously supplied the Lithium-ion batteries for the twin Mars Exploration Rovers.

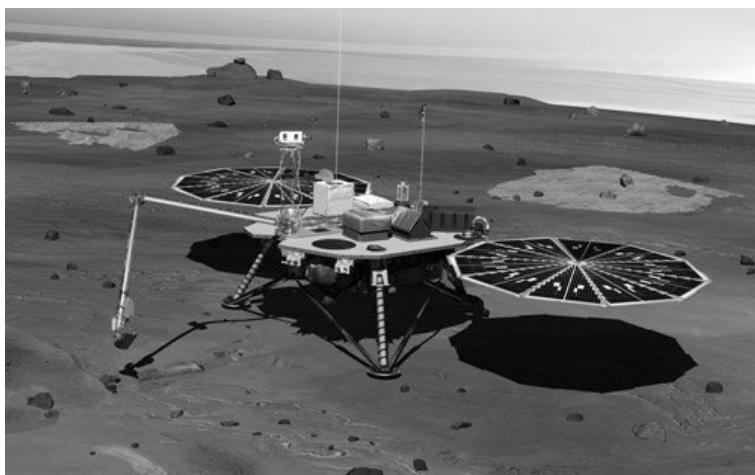


Fig. 1.4 Artist's image of the Phoenix Lander on Mars (Image courtesy of NASA)

1.5 Future Mars Exploration: 2010+

The foreseeable future will see a continuation of the robotic surface exploration of Mars. Larger, heavier, and more sophisticated systems such as the Mars Science Laboratory and the ExoMars Rover will continue to broaden our knowledge about the surface of Mars. These advanced systems will require high-performance power systems to fulfil their mission objectives.

A timeline of the future Mars exploration phase is provided in Table 1.4.

Table 1.4 Timeline of the future Mars exploration phase

Launch Date	Mission Name ¹	Outcome
2009-2011	Phobos/Grunt	Sample return mission to Phobos ²
2009-2011	Yinghuo 1	Chinese orbiter to be launched with Phobos/Grunt
end of 2011	Mars Science Laboratory	Mars surface rover
2011+	ExoMars	Orbiter, lander and surface rover
2013+	Mars Atmosphere and Volatile Evolution	Mars Orbiter
2013+	2013 Mars Science Orbiter	Successor to the Mars Reconnaissance Orbiter
2015+	Mars Sample Return	Sample return mission to the surface of Mars
...
2030+	n/a	First manned flight to Mars

¹Grey background indicates surface landing elements

²Phobos/Grunt not further discussed, as no landing operation on Mars is planned

1.5.1 Mars Science Laboratory Power System

The Mars Science Laboratory (MSL) is currently being developed as follow-up to the hugely successfully twin Mars Exploration Rovers. Being more than five times as heavy as the MERs (approx. 900 vs. 174 kg), the MSL will carry an extensive array of scientific instruments to investigate if life ever existed on Mars, to characterize climate and geology of Mars, and to conduct further investigations in preparation of human exploration.

The Mars Science Laboratory will use a radioisotope power system. As with the RTGs used with the Viking missions of the mid-1970s, electricity is produced from the heat released by the natural decay of plutonium 238 utilizing thermoelectric converters. The so-called Multi-Mission Radioisotope Thermoelectric Generator (MMRTG) is going to be used with the MSL. The MMRTG contains a total of 4.8 kg of plutonium dioxide that initially provides approximately 2,000 W of

thermal power, of which approximately 120 watts are available as electrical power output of the thermoelectric converters. The MMRTG generator is about 64 cm in diameter (i.e. measured from fin tip to fin tip) by 66 cm long and weighs about 43 kg. The MMRTG therefore offers a very high energy content with a continuous power output in a compact, package (DOE 2006).

An image of the MSL is shown in Fig. 1.5. The dark inclined cylinder-like shaped objects located in the rear of the vehicle are the heat exchangers of the radioisotope power source.



Fig. 1.5 Artist's image of the Mars Science Laboratory (Image courtesy of NASA)

Fig. 1.6. shows a cutdrawing of the GPHS-RTG that are used for Galileo, Ulysses, Cassini-Huygens and New Horizons space probes.

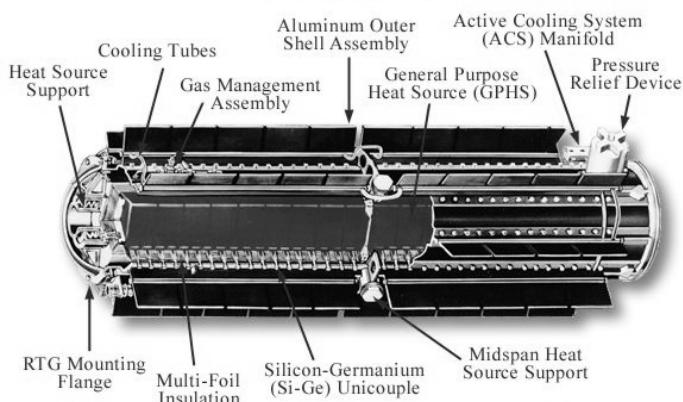


Fig. 1.6 A cutdrawing of an GPHS-RTG that is used for Galileo, Ulysses, Cassini-Huygens and the New Horizons space probes (Image courtesy of NASA)

The RTG for the Cassini-Huygens probe was significantly larger than the MMRTG currently being developed for the MSL and has at output power of 885 W at beginning of life, and some 633 W at end of mission (NASA Cassini 1997). In comparison, the MMRTG of the MSL will only have a beginning-of-life output power of 120 W.

1.5.2 ExoMars Rover Power System

ExoMars is ESA's ambitious Mars exploration mission consisting of an orbiter as well as a carrier and descent module carrying a robotic surface rover. The main scientific objective of the ExoMars mission will be to study the biological environment of the Martian surface, to search for traces of life, to investigate geochemistry and geophysics, and to explore possible risks and threats for future manned Mars surface missions.



Fig. 1.7 Artist's image of the ExoMars rover rolling down the descent module (Courtesy of ESA - image by AOES Medialab)

As the ExoMars rover is currently still under development, different versions of the rover specifications are available. The data presented in the following is taken from recent ESA technical presentations and should therefore provide an up-to-date information about the current development goals.

1.5.2.1 ExoMars Lander Power System

The ExoMars science instruments are divided into the Pasteur payload (installed on the Rover) and the Humboldt payload (installed on the lander). The lander will

therefore not only carry the rover during descent and landing, but will also be used as platform for scientific exploration once the rover is deployed.

Primary power will be provided by six solar panels hinged to the lander base carrying the rover during transfer and landing. The arrays will have a total active area of 2.9 m². In addition, a 500 Wh Lithium-ion (Li-ion) technology battery module will be installed in the lander (McCoy 2008).

1.5.2.2 ExoMars Rover Power System

The rover will weigh approximately 235 kg and thus be slightly heavier than the American Mars Exploration Rovers Spirit and Opportunity. The scientific payload will consist of the roughly 16.5 kg Pasteur package and the Geophysical Science Package, containing, among other instruments, a large drill (McCoy 2008).

The ExoMars rover power system will consist of a combination of solar panels and a Lithium-ion (Li-ion) technology secondary battery module. Primary power will be provided by five solar panels (one fixed, four deployable) having a total of 2.6 m² of active area. This is approximately twice the solar array area available with the MERs (2.6 m² with the ExoMars rover versus 1.3 m² with the Mars Exploration Rovers).

In addition, a Lithium-ion technology battery module in 6s9p configuration is going to be installed into the rover (Roe 2008). The sp-notation gives the number of cells in series ‘s’ (i.e. in a so-called string), and the number of strings installed in parallel ‘p’. A 6s9p configuration therefore specifies that the battery module consists of nine parallel strings of six cells each; this corresponds to a total of 54 single cells.

The power system of the ExoMars rovers will therefore essentially apply the proven hybrid photovoltaic/secondary battery power system approach. Applying the proven design approach, the batteries will also be installed into a central insulated compartment installed into the rover base, with the solar panels attached, respectively hinged, to the top of the central rover compartment.

1.5.3 Future Mars Surface Missions

In the coming years and decades, Mars will be further investigated with a wide range of different robotic systems. This will include stationary landers, surface rovers, and even exotic airborne robotic systems such as airships, aeroplanes, or rocket-powered hoppers (Visentin and van Winnendael 2006).

The sample return mission will be one of the most important milestones in future Mars exploration. This mission will return Martian soil and rock samples back to Earth (O’Neil and Cazaux 2000, Clark 2007). The Soviet Union’s Luna 16, 20 and 24 already succeeded in a similar mission profile in 1970s, returning soil samples from the Moon back to Earth. The upcoming Mars sample return mission has not only been a premier objective in robotic planetary exploration for decades, but could also be used to test many of the technologies also required with future manned surface exploration.

This will probably also include the proof-of-concept for power generation, propellant production and energy storage applications. In-situ resource utilisation

(IRSU) is often considered as enabler for Earth-return missions. ISRU refers to the utilisation of in-situ resources such as atmosphere or surface soils and rocks for propellant production. ISRU promises a significant reduction in Earth launch mass, as the return spacecraft can be launched from Earth with empty tanks and an ISRU plant designed to produce the propellants directly on the surface of Mars. ISRU is an interesting option for robotic return missions, and it is currently considered as almost inevitable in state-of-the-art manned return missions (Zubrin and Wagner 1996; Hoffman and Kaplan 1997).

The Automation and Robotics for Human Mars Exploration (AROMA) study, made under ESA contract between 2000 and 2002, investigated Automation & Robotics (A&R) systems required in the robotic outpost phase and the human exploration phase (von Richter 2002; Hofmann and von Richer 2003). Systems considered in this study include mini/micro rovers, automated instrument packages and automated drilling systems as well as mobile pressurized laboratories and utility trucks required with manned exploration of Mars. The output power requirements of the investigated systems range from a few watts for mini- and micro-rovers up tens of kilowatts with manned rovers (Fraser et al. 2004).

The versatile and challenging mission profiles of future Mars exploration programs require a careful evaluation of both existing and not yet technically mature power system concepts in order to pursue the development of innovative power system elements in a timely manner (Fraser 2001).

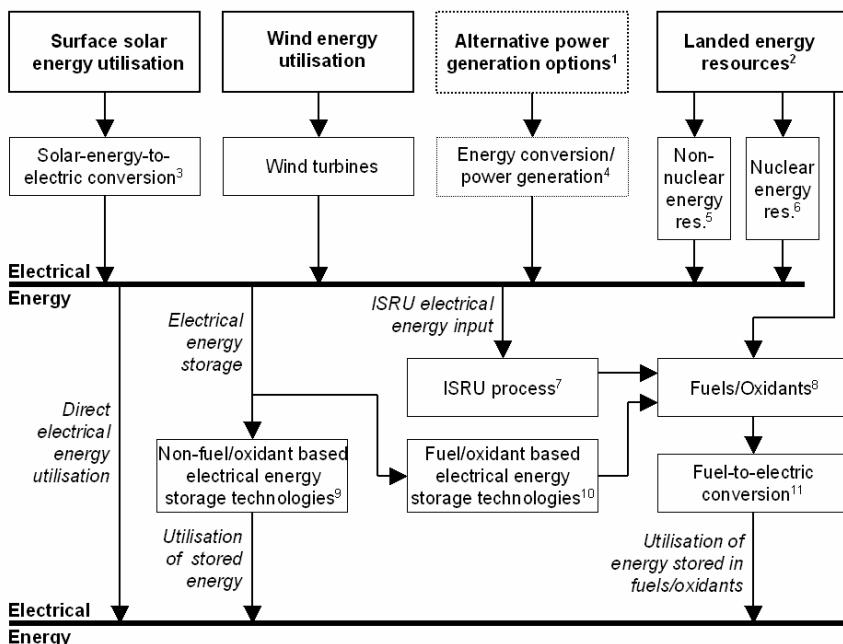


Fig. 1.8 Overview of power generation options and energy pathways for Mars surface applications

1.6 Mars Surface Power Systems Options

Orbital parameters, atmosphere and surface geology of Mars enable a wide range of different power generation options. On Mars, solar energy is weaker than on Earth and Moon, but – other than with missions to the outer planets - utilization is feasible with a reasonably large solar panel area. The presence of a (thin) atmosphere enables wind energy utilisation. Atmospheric composition and surface geology enable in-situ resource utilisation for propellant production. Landing resources on the

Table 1.5 Description of synonyms used in Fig. 1.8.

Index	Synonym used in Fig. 1.8.	Non-exhaustive list of possible power system elements
1	Alternative power generation options	e.g. geothermal energy, solar power satellites/beamed power from space
2	Landed energy resources	e.g. primary batteries, fuels (e.g. hydrogen) or feedstock species, nuclear energy resources
3	Solar-energy-to-electric conversion	e.g. photovoltaics, solar dynamic power systems
4	Energy conversion/power generation	Technologies required for power generation or energy conversion from an alternative energy source (e.g. antenna for beamed power from space)
5	Landed non-nuclear energy resource	e.g. primary batteries, readily-fuelled fuel cells or heat engines
6	Landed nuclear energy resource	e.g. radioisotope generators, fission reactors with static or dynamic conversion systems
7	ISRU process	In this context: Utilisation of in-situ resources (e.g. atmosphere, soil, rocks) for propellant and/or oxidant production
8	Fuels/oxidants	e.g. hydrogen/oxygen, methane/oxygen, carbon monoxide/oxygen, methanol/oxygen
9	Non-fuel/oxidant based electrical energy storage technologies	e.g. secondary batteries, supercapacitors, flywheels, compressed air energy storage, superconducting magnetic energy storage
10	Fuel/oxidant based electrical energy storage technologies	Fuel/oxidant production for electrical energy storage purposed (can be re-converted into electrical energy by a fuel-to-electric conversion technology)
11	Fuel-to-electric conversion	e.g. primary fuel cells, heat engines, static conversion systems (thermoelectric, alkali metal thermoelectric converter), dynamic conversion systems (e.g. Brayton cycle)

surface of Mars is also an option, particularly with smaller missions, as demonstrated in the past. In addition, there are alternative power generation options, such as beaming power from space that might be an option in the decades and centuries to come.

An overview of the principal power generation and energy storage options is provided in Fig. 1.8. This schematic provides a non-exhaustive list of pathways from energy resources to electrical energy with a focus on near-future applications. Some comments regarding non-intuitive synonyms used in Fig. 1.8. are compiled in Table 1.5.

1.6.1 Energy Pathways Used in Past and Present Mars Surface Mission

As described above, previous and present missions were essentially based on two energy pathways: the missions either relied on landed energy resources or surface solar energy utilisation.

Landed resources were primarily RTGs, but also primary batteries (e.g. with Sojourner). The RTGs were operated in combination with rechargeable batteries for intermediate energy storage purposes (e.g. Viking Landers, Mars 96 Small Stations).

Surface solar energy utilisation was always combined with either secondary batteries (e.g. Pathfinder Lander) or primary batteries (e.g. Sojourner) in order to handle the diurnal variations in solar power generation and to have energy for peak power and nighttime operations.

1.6.2 Energy Pathways Potentially Used in Near-Future Mars Surface Missions

In near-future Mars surface missions, the energy pathways will most likely still be focused on surface solar energy utilisation and landed energy resources. Both approaches will co-exist and are going to be gradually optimised with technological advances on the one hand, and as a function of mission profile, power and energy requirements, economic considerations and – always an issue with nuclear energy – safety issues on the other hand.

Robotic rovers of the size and the performance of the Mars Exploration Rovers can be powered by solar energy if the limitations in magnitude and availability of output power can be accepted. These limitations are an issue with such as a rather small and low-power application, but when moving to high-performance applications such as manned rovers, where tens of kilowatts of output power have to be continuously available over the course of up to 20 days, solar energy and secondary batteries alone certainly are not the power system options of choice. This becomes obvious when comparing the very strongly simplified specific energy estimates compiled in Table 1.6.

A factor of roughly 200 is derived between the state-of-the-art specific energy of RTGs and the long term performance goals for secondary batteries in this strongly simplified and application-dependent estimation. Even fuelled systems – a hydrogen/oxygen fuel cell is included for comparison – cannot match the performance of nuclear technology. Only photovoltaics, and similar figures may also be derived for other surface solar energy utilisation technologies (and with certain

limitations this is also true for wind energy utilisation) can provide similar performance figures if the system is operated over long periods of time. With Mars, this implies that counter-measures have to be taken to prevent a gradual reduction in output power due to dust deposition on the panels. A 0.28% performance reduction per sol due to dust deposition on solar panels was e.g. measured with Pathfinder. Such a level of performance degradation cannot be accepted if a solar power system is designed for a long operational lifetime.

Table 1.6 Specific energy estimates for selected technologies

Technology	Specific energy	Derivation
RTG	46 kWh kg ⁻¹	MMRTG mass: 43 kg MMRTG power: 120 W power output MMRTG operation: 1 year (687 sols) PV mass: 5 kg m ⁻²
Photovoltaics	41 kWh kg ⁻¹	PV efficiency: 10% average Solar radiation: 3 kWh m ⁻² sol ⁻¹ PV operation: 1 year (687 sols)
H ₂ /O ₂ fuel cell	>1 kWh kg ⁻¹	Including mass of oxygen and storage vessels, limited fuel-to-electric conversion efficiency
Secondary battery	0.2 kWh kg ⁻¹	USABC long term goal for EV batteries excluding limited depth-of-discharge and energetic discharge efficiency ¹

¹ Secondary battery figures are long-term performance targets defined by the United States Advanced Battery Consortium - USABC (USABC 2008; Chalk and Miller 2008) and do not reflect the current status of battery technology

In near-future missions, primary power generation will therefore most likely be achieved with a nuclear and/or a photovoltaic power system.

Energy storage technologies such as regenerative fuel cells or secondary batteries will be used for intermediate energy storage, but not for primary power generation. Even if fuelled power system elements are included in a Mars surface mission, fuel and oxidant are not likely to be readily landed on Mars, but will rather be produced by ISRU technology, optionally using a feedstock of species brought from Earth (see Chap. 17).

Wind energy utilisation is an option, but the low atmospheric pressure (<10 mbar) requires high wind speeds and/or very large turbine diameters. In addition, the diurnal and seasonal availability of wind energy can hardly be predicted accurately, and large energy storage capacities would therefore have to be included if large diurnal and seasonal variations in output power availability are not acceptable.

Output power of a surface solar energy system is more predictable, considering the diurnal solar cycle and an additional uncertainty due to seasonal variations and

seasonal phenomena such as dust storms. Alternative surface solar energy utilisation technologies may be applied, e.g. solar dynamic systems, if they can provide considerable advantages compared to photovoltaic power generation (for a discussion see Chap. 5).

Table 1.7 Surface solar energy utilization versus nuclear power generation

Issue	Surface solar energy utilisation (PV, dynamic system)	Nuclear scenario (RTG, DIPS, fission reactors)
Specific energy with respect to the Earth launch mass	Very high; depends on location and operational lifetime ¹	Very high; depends on operational lifetime or mission duration ²
System lifetime	Very long, but depends on maintenance ¹	Very long ²
Continuous power output	Not available; intermittent storage system necessary	Available
Reliability	Very high, but depends on installation, location and operation ¹	Very high
Onsite-installation	Regularly required with large solar panels	Not required with RTGs; tbd with fission reactors
Maintenance	Maintenance may be required ¹	No maintenance with RTGs; tbd with fission reactors
Operation in harsh environmental conditions	Dust suspended in atmosphere is a problem	No problem
Output power predictability	Limited; depends on weather, season and dust deposition	Very high; output power degradation rate of RTGs is known a priori
Dependence on orientation	Limited; problem with mobile systems	No problem
Check prior to launch	Limited; no full check if panels have to be unfolded or installed	Possible
Power during launch and transfer to Mars	Normally not possible	Possible
Safety during launch	No problem	Launch failure may have catastrophic effect
Public opinion	Very positive	Problem, particularly in Europe

¹ Dust deposition is a major problem affecting lifetime, maintenance and useable energy output of surface solar power systems ($0.28\% \text{ sol}^{-1}$ performance reduction due to dust deposition with Pathfinder)

² Useable energy output often limited by mission duration rather than RTG lifetime

In general, the possibilities of surface solar energy utilisation systems have to be critically compared to a nuclear scenario in order to identify if a primary-solar solution is also an option with larger human precursor and manned missions. A comprehensive comparison of these two most promising power generation approaches is given in Table 1.7.

The comparison of surface solar energy utilisation and nuclear power generation shows that both technologies have advantages and disadvantages. A generally valid conclusion as to when either of the two approaches can or should be applied in Mars surface missions cannot be made. Both options have to be carefully evaluated for a specific mission profile, the technological availability and the broader Mars exploration plans.

There will certainly be missions where surface solar energy utilisation is the technology of choice, and there will also be missions where a nuclear power system will perform very favourably. The compact, robust and reliable nuclear power sources generally are a hard match for surface solar energy utilisation in mobile applications and in high-power applications where a continuous supply with electrical energy is required day and night.

As this will most likely be the case in human surface exploration, nuclear power generation based on advanced RTGs, highly-efficient dynamic isotope power systems (DIPS) and eventually even small fission reactors will most likely see a renaissance (see Chaps. 7-10).

1.7 Summary

A safe, reliable and sufficient supply with electrical energy is critical in Mars surface mission design. A wide range of different power system technologies are available and either technically feasible or already tested under laboratory conditions. Some technologies are technically mature and space-proven, some are new and innovative and promise performance characteristics unmatched by state-of-the-art power system options. A comprehensive evaluation is therefore necessary to find the power system elements best suited for a specific application.

Surface solar energy utilisation is an option on the surface of Mars. The irradiance is considerably lower than on Earth, and is even further reduced by strong seasonal dust storms (see Chap. 2). Dust deposition on solar panels is another issue that gradually reduces the output power as long as no countermeasures are taken. In addition, the operating temperatures can be considerably lower than in most other photovoltaic applications. The development of low intensity low temperature (LILT) solar cells is therefore of significant importance. Solar dynamic power systems remain an option.

Wind energy utilisation is also an option on Mars. The thin Martian atmosphere requires very large turbine blades and/or very high wind speeds to produce a useable power output. Diurnal and seasonal variations in wind energy have to be considered. Alternative means of power generation and/or a sufficient energy storage capacity therefore have to be included if wind is considered as primary energy source.

Nuclear power is also an option on Mars. Radioisotope thermoelectric generation is a proven technology and has been successfully applied in a number of space missions within the last decades. This technology is light, compact, reliable and provides a continuous and very predictable output power. In addition, dynamic isotope power systems can be considered as a replacement for thermoelectric generators to increase the electrical output power without increasing the plutonium mass. Small fission reactors have already been developed for space applications in the past and are attractive options for large, manned surface missions where continuous electrical output power levels of 100 or even 250 kW are required.

Two energy pathways have been extensively applied in previous and present missions to the surface of Mars: surface solar energy utilisation and radioisotope power generation, both with battery storage for backup and peak output power applications. Future Mars surface missions will most likely see a continuation and further development of these two options. Large robotic and ultimately manned missions will require a safe, reliable and predictable stationary power supply not only for daily operations, but also for propellant production utilising in-situ resource utilisation, as well as for large mobile exploration systems.

Non-nuclear technologies such as solar and wind energy utilisation are limited on a larger scale due to issues with on-site installation, maintenance, and diurnal as well as seasonal variations in output power. Nuclear technologies, on the other hand, provide an almost constant output power in a very compact and robust package. There are no significant diurnal or seasonal variations in output power, and installation and maintenance is easier than setting up large arrays of wind turbines or hundreds of square meters of solar panels. Nuclear energy is therefore considered as *enabling technology* and, from today's perspective, will be increasingly applied when moving towards larger human precursor and manned missions in the near future.

References

- Beagle 2 homepage, Power Section (2004),
<http://www.beagle2.com/technology/power.htm>
- Chalk, S.G., Miller, J.F.: Key challenges and recent progress in batteries, fuel cells, and hydrogen storage for clean energy systems. *J. Power Sources* 159, 73–80 (2006)
- Clark, B.C.: Mars sample return: the critical next step. *Acta Astronaut* 61, 95–100 (2007)
- DOE, Space Radioisotope Power Systems - Multi-Mission Radioisotope Thermoelectric Generator (2006), <http://nuclear.gov/pdfFiles/MMRTG.pdf>
- Fraser, S.D.: Non-nuclear power system options for a mission to Mars and derived terrestrial applications. Diploma thesis, Graz University of Technology (2001)
- Fraser, S.D., Monsberger, M., Hacker, V.: Fuel cell power system options for Mars rovers. In: Proc. 2nd International Conference on Green Propellants for Space Propulsion, Sardinia, Italy, June 7–8 (2004)
- Hoffman, S.J., Kaplan, D.I. (eds.): The reference mission of the NASA Mars exploration study team. NASA Special Publication 6107 (1997)
- Hofmann, P., von Richer, A.: Automation and robotics for human Mars exploration (AROMA). *Acta Astronaut* 53, 399–404 (2003)

- Linkin, V., Harri, A.-M., Lipatov, A., Belostotskaja, K., Derbunovich, B., Ekonomov, A., et al.: A sophisticated lander for scientific exploration of Mars: scientific objectives and implementation of the Mars-96 Small Station. *Planet Space Sci.* 46, 717–737 (1998)
- NASA Cassini. Cassini Launch. Press kit (October 1997)
- NASA 1998 Mars Missions. 1998 Mars missions. Press kit (December 1998)
- NASA MER. Mars Exploration Rover Launches. Press kit (July 2003)
- NASA Pathfinder. Mars Pathfinder Landing. Press kit (July 1997)
- NASA Phoenix. Phoenix launch – mission to the Martian polar north. Press kit (August 2007)
- NASA Viking. Viking 1 Lander. NSSDC ID: 1975-075C (2009)
- McCoy, D.: ExoMars Technical Presentation to PB-HME. ESTEC, September 26 (2008)
- O’Neil, W.J., Cazaux, C.: The Mars sample return project. *Acta Astronaut* 47, 453–465 (2000)
- Ratnakumar, B.V., Smart, M.C., Huang, C.K., Perrone, D., Surampudi, S., Greenbaum, S.G.: Lithium ion batteries for Mars exploration missions. *Electrochim Acta* 45, 1513–1517 (2000)
- Ratnakumar, B.V., Smart, M.C., Kinderl, A., Frank, H., Ewell, R., Surampudi, S.: Lithium batteries for aerospace applications: Mars Exploration Rover. *J. Power Sources* 119-121, 906–910 (2003)
- Ratnakumar, B.V., Smart, M.C., Whitcanack, L.D., Ewell, R.C.: The impedance characteristics of Mars Exploration Rover Li-ion batteries. *J. Power Sources* 159, 1428–1439 (2006)
- Roe, M.: ExoMars Rover Vehicle. European Industry Day, TAS-I, Turin, Italy (2008)
- Schmidt, R.: Mars Express – ESA’s first mission to planet Mars. *Acta Astronaut* 52, 197–202 (2003)
- Shirley, D.L.: Touching Mars: status of the Mars robotic exploration program. *Acta Astronaut* 45, 249–265 (1999)
- Surkov, Y.A., Kremnev, R.S.: Mars-96 mission: Mars exploration with the use of penetrators. *Planet Space Sci.* 46, 1689–1696 (1998)
- USABC, Goals for advanced batteries for EVs. Downloaded from the USABC energy storage systems goals (2008), <http://www.uscar.org>
- Visentin, G., van Winnendael, M.: Robotics options for low-cost planetary missions. *Acta Astronaut* 59, 750–756 (2006)
- von Richter, A.: Automation and robotics for human Mars exploration (AROMA): executive summary. ESTEC/Contract 14563/oo/NL/WK (2002)
- Zubrin, R., Wagner, R.: The case for Mars: the plan to settle the red planet and why we must. Touchstone, New York (1996)

Chapter 2

Available Solar Energy and Weather Forecasting on Mars Surface

Viorel Badescu

Polytechnic University of Bucharest, Romania

2.1 Introduction

The design of both photovoltaic and photothermal systems operating on Mars requires detailed information on the solar radiation flux incident on the surface of the planet as a function of latitude, longitude, time of day and year. The atmospheric (vertical) optical depth τ is an indicator of solar radiation attenuation in the atmosphere. One can concisely define the optical depth as follows. Let us assume that the sun is at zenith and denote by I_0 and I_{ground} the direct solar irradiance at the top of the atmosphere and at ground level, respectively. Then, usage of τ allows to write the Beer's law as $I_{ground}/I_0 = e^{-\tau}$. Suspended dust absorbs solar radiation and emits (and absorbs) longer-wavelength radiation. Mars may be considered "clear" when the dust content in the atmosphere is low, but when local or global storms occur the optical depth increases and the direct beam solar radiation decreases drastically.

There is quite a literature on the interaction of radiation and Martian atmosphere (see e.g. Crisp (1990), Lindner (1990), Hourdin (1992), Santee and Crisp (1993), Kahn (1995)). As is known, the Martian atmosphere consists of gases (95 % CO₂, 2.7 % N₂ and 1.6 % Ar, traces of water vapor and ozone) and dust. Carbon dioxide is by far the most important gaseous absorber and emitter, mostly through its 15 μm band (Hourdin 1992). Dust is always present, with abundances that vary significantly with season and from year to year (Beish and Parker 1990; Clancy and Lee 1991; Martin and Zurek 1993). Occasionally, winds on Mars exceed the threshold value for setting particles into motion, and short-lived (~day) local dust storm result. Approximately once or several times each Martian year, especially near the perihelion orbital position, a local dust storm expands to global proportions (Pollack et al. 1979; Zurek and Martin 1993).

Earlier studies are available concerning the direct beam component of solar radiation on Mars surface (Van Hemelrijk 1985, 1987). A more complete set of data is provided in Appelbaum and Flood (1990), Appelbaum et al. (1993) where information on direct, diffuse and global solar radiation is given. These data are based on the measured optical depth of the Martian atmosphere derived from images taken of the Sun and Phobos with a special diode on the Viking Landers cameras. It is useful to recall that the Viking Lander 1 (VL1) separated on August 20, 1976

from the Orbiter and touched down at Chryse Planitia. Viking Lander 2 (VL2) landed at Utopia Planitia on September 3, 1976 (VL2 ended communications on April 11, 1980 and VL1 on November 13, 1982). The solar radiation data of Appelbaum and Flood (1990), Appelbaum et al. (1993) could be considered as representative for a larger range of latitudes because of the similarity in the properties of the dust suspended above the two landing sites.

The data provided by Appelbaum and co-workers refer to solar radiation on a horizontal surface. It is well known, however, that generally the amount of collected solar energy depends on the orientation and tilt of the receiving surface. A large number of papers has studied the optimum tilt and azimuth angles for different sites on Earth (see e.g. Nijegorodov et al. (1994), Badescu (2008)). It is the purpose of this chapter to deal with the optimum slope and orientation of a surface receiving solar radiation on the Mars surface.

The solar energy systems are usually classified according to the kind of radiation they are designed for. Some solar energy devices use solar global radiation. They are quite attractive in the case of Mars, because, first, they imply a simpler technology and second, they use *both* direct and diffuse components. Indeed, it was proven that there is a large diffuse component of insolation even at high optical depth, when the direct component practically vanishes (Appelbaum and Flood 1990; Appelbaum et al. 1993). Other solar energy devices are designed to use the direct component only. This component can be concentrated. As a result of the increase in concentration, the power and the efficiency of space or ground-based solar power plants increase (Badescu and Dinu 1993, 1995, 1996; Badescu 1994, 1995, 1996).

In this chapter we focus on both direct and global solar irradiances. A procedure to evaluate direct, diffuse and global radiation on arbitrarily tilted surfaces is presented in section 2.2. The core of this procedure is the model proposed in Appelbaum and Flood (1990) and developed in Appelbaum et al. (1993) to compute the solar radiation components on a horizontal surface on Mars. A simple model is used to derive the values of solar radiation on an arbitrarily tilted and oriented surface from the values computed on a horizontal surface. A number of usual strategies concerning the adjustment of the orientation of a solar receiver are studied here. Mainly they differ in the solar device design to collect either solar direct or global radiation. For each strategy the optimum tilt angle is evaluated. Most part of the primary data we use here is based on Viking Landers measurements. Of course, the results we obtain characterize the VL1 and VL2 sites but they can be used as a first approximation for other locations at similar latitudes.

Section 2.3 reviews a series of procedures that can be used to forecast meteorological parameters on the surface of Mars. The techniques envisaged could be valuable for a sound preparation of the future systems operating on Mars. Indeed, simulations are necessary to evaluate the usefulness of the equipment under various climatic situations. This sort of simulations requires a detailed knowledge of the time variation of meteorological parameters. For instance, during testing of solar cells operation (see Badescu (1998) and Chap. 3) or during simulations of a solar collector- thermal engine system (see Badescu et al. (2000) and Chap. 5) the atmospheric pressure and temperature were factors controlling the thermal losses.

But time-dependent simulations need *continuous* series of meteorological data. Providing such data is not a simple task even on Earth. More than that, the available data is limited on Mars. This makes it useful to develop models of synthetic data generation. Data synthesis was already used on Earth in case of places lacking meteorological data or to complete the gaps in existing measured series. Models were developed for a number of terrestrial meteorological and actinometric parameters such as daily and hourly global irradiation (Festa and Ratto 1993) and point cloudiness (Badescu 1997, 2008). Section 2.3.1 deals with statistical inference relations of atmospheric optical depth developed for clear sky and dust storm conditions. Section 2.3.2 refers to forecasting of daily averaged atmospheric pressure while Sect. 2.3.3 to forecasting of hourly averages of various meteorological parameters.

2.2 Available Solar Energy on Mars Surface

One denotes by ϕ and Γ the latitude and longitude of a site on Mars surface, respectively. Here the convention is: $-90^\circ \leq \phi \leq 90^\circ$ (North positive) and $-180^\circ \leq \Gamma \leq 180^\circ$ (East positive). As an example VL1 and VL2 sites are at $\phi_1 = 22.27^\circ$; $\Gamma_1 = 47.97^\circ$ and $\phi_2 = 47.57^\circ$; $\Gamma_2 = 134.26^\circ$, respectively. One assumes the Martian year begins at the vernal equinox, when the areocentric longitude L_s of the Sun, measured in the orbital plane of the planet, is $L_s=0^\circ$. A Martian year has 669 Martian days. A Martian day (= sol) has 24.65 terrestrial hours. The seasons on Mars are unequal in duration as a result of the appreciable eccentricity of the Mars orbit. Table 2.1 gives some details about the Martian seasons. Note that on the Northern Hemisphere the areocentric longitude intervals for spring, summer, autumn and winter are, respectively: $[0^\circ \text{ to } 90^\circ]$, $[90^\circ \text{ to } 180^\circ]$, $[180^\circ \text{ to } 270^\circ]$ and $[270^\circ \text{ to } 360^\circ]$, respectively.

Table 2.1 Details about the Martian seasons

Northern hemisphere	Southern hemisphere	Duration (sols)	Range (sols)	Areocentric Longitude, L_s (degrees)
spring	autumn	194	0-194	0-90
summer	winter	178	195-372	90-180
autumn	spring	143	373-515	180-270
winter	summer	154	516-669	270-360

It is convenient to divide one sol into 24 Martian hours (1 sol = 24 h_{Mars}) and one Martian hour into 3600 Martian seconds (1 $h_{\text{Mars}} = 3600 s_{\text{Mars}}$). Consequently, 1 sol = $N_{\text{sec/sol}}$ and 1 Martian year = T where $N_{\text{sec/sol}}=86400 s_{\text{Mars}}$ and $T=57801600 s_{\text{Mars}}$. One denotes by t the time during a Martian year, measured in Martian seconds ($t=0$ when the year begins, i.e. when $L_s=0^\circ$). Then, the noon of the n th sol occurs

at time $t_{\text{noon},n} = N_{\text{sec/sol}}(n-1/2)$. It is obvious that the sols are equal in duration but the variation of the areocentric longitude of the Sun during one sol depends on the sol's number. One denotes by $L_{s,\text{noon},n}$ the areocentric longitude of the Sun at the noon of the n th sol. $L_{s,\text{noon},n}$ can be computed numerically as a function of n as shown in next section.

2.2.1 Model

The relationship between the true anomaly angle v and the areocentric longitude L_s is given by Eq. (3) of Pal and Ureche (1983):

$$v = L_s - 248^\circ \quad , \quad (2.1)$$

where 248° is the areocentric longitude of Mars perihelion. The true anomaly is a function of the eccentric anomaly E (see Eq. (A.5.8) of Pal and Ureche (1983))

$$\frac{v}{2} = \arctg \left(\sqrt{\frac{1+e}{1-e}} \tg \frac{E}{2} \right) \quad , \quad (2.2)$$

where $e (=0.093377)$ is the eccentricity of Mars orbit. The relation between the time t and the eccentric anomaly E ($t=0$ if $L_s=0^\circ$) is given by Eq. (A.5.14) of Pal and Ureche (1983)

$$\frac{2\pi(t-t_{\text{perihelion}})}{T} = E - e \sin E \quad , \quad (2.3)$$

where T is the duration of a Martian year and $t_{\text{perihelion}}$ is the time of Mars perihelion (when $v=0^\circ$ and $L_s=248^\circ$). By using Eqs. (2.1)-(2.3) one obtains $t_{\text{perihelion}}$ as

$$t_{\text{perihelion}} = -\frac{T}{2\pi} \cdot \quad (2.4)$$

The noon of the n th sol occurs at the time

$$t_{\text{noon},n} = N_{\text{sec/sol}}(n-1/2) \quad (2.5)$$

By using Eqs. (2.1)-(2.3), the following relationship between $t_{\text{noon},n}$ and $L_{s,\text{noon},n}$ can be derived:

$$\begin{aligned} \frac{2\pi}{T}(t_{\text{noon},n} - t_{\text{perihelion}}) &= 2\arctg \left[\sqrt{\frac{1-e}{1+e}} \tg \frac{L_{s,\text{noon},n} - 248^\circ}{2} \right] \\ &\quad - \sin \left[2\arctg \left(\sqrt{\frac{1-e}{1+e}} \tg \frac{L_{s,\text{noon},n} - 248^\circ}{2} \right) \right]. \end{aligned} \quad (2.6)$$

Once n is known, $t_{\text{noon},n}$ can be computed using Eq. (2.5). Then $L_{s,\text{noon},n}$ is obtained by solving numerically Eq. (2.6).

2.2.1.1 Various Angles Related to the Sun's Apparent Movement

The zenith angle of the Sun θ_z can be computed by Nijegorodov et al. (1994):

$$\cos \theta_z = \cos \phi \cos \delta \cos \omega + \sin \phi \sin \delta , \quad (2.7)$$

where ω is the hour angle while δ is the solar declination angle. In case the Mars solar time t_{solar} is measured in Martian hours the relationship between solar time and hour angle ω (degrees) is (Appelbaum and Flood 1990):

$$\omega = 15 \cdot t_{solar} - 180^\circ . \quad (2.8)$$

ω is negative in the morning and vanishes at the solar noon. The solar declination angle δ is given by (Appelbaum and Flood 1990):

$$\sin \delta = \sin \delta_0 \sin L_s , \quad (2.9)$$

where $\delta_0 (=24.936^\circ)$ is the Mars obliquity of the rotation axis to the orbital plane. Let us consider a plane of arbitrary tilt and azimuth angles β and γ , respectively. The conventions assumed here are: $\beta \in [-90^\circ, 90^\circ]$ is positive facing South and negative facing North while, $\gamma \in [-180^\circ, 180^\circ]$ is East negative and West positive.

The angle of incidence θ of the Sun rays on the tilted surface is (Nijegorodov et al. 1994)

$$\begin{aligned} \cos \theta = & \sin \delta \sin \phi \cos \beta - \sin \delta \cos \phi \sin \beta \cos \gamma + \\ & + \cos \delta \cos \phi \cos \beta \cos \omega + \cos \delta \sin \phi \sin \beta \cos \gamma \cos \omega \\ & + \cos \delta \sin \beta \sin \gamma \sin \omega . \end{aligned} \quad (2.10)$$

The sunset hour angle $\omega_{ss,H}$ (degrees) for a horizontal surface is given by (Nijegorodov et al. 1994)

$$\omega_{ss,H} = \frac{180^\circ}{\pi} \arccos(-\tan \phi \tan \delta) . \quad (2.11)$$

The evaluation of the sunrise and sunset hour angles, $\omega_{sr,T}$ and $\omega_{ss,T}$, respectively, for an arbitrary tilted and oriented surface located at an arbitrary latitude is more complicated. Depending on the surface orientation and time of the year, the Sun may rise and set on an inclined surface twice, once or not at all during a day. Different analytical procedures are available to compute $\omega_{sr,T}$ and $\omega_{ss,T}$ (see, e.g. Jones (1980), Andersen (1980), Iqbal (1981), Chowdury and Rahman (1987), Nijegorodov et al. (1994), Kambezidis (1997)). During various tests none of them proved to be reliable because of graphical and/or intrinsic errors. So, we prepared a simple search algorithm to evaluate numerically the sunrise and sunset hour angles for a tilted surface.

2.2.1.2 Solar Irradiance on Tilted Surfaces

The beam irradiance on a surface perpendicular to the Sun's rays at the top of the Martian atmosphere, I_0 , is a function of the areocentric longitude L_s (Appelbaum et al. 1993):

$$I_0 = 590 \frac{[1 + e \cos(L_s - 248^\circ)]^2}{(1 - e^2)^2} , \quad (2.12)$$

where 590 W/m^2 is the *mean* beam irradiance. For a given sol number n ($1 \leq n \leq 669$) we evaluated the areocentric longitude at the solar noon $L_{s,noon,n}$. This value was used in Eq. (2.12) to evaluate I_0 which was assumed to be constant during the given sol.

The direct irradiance on a surface perpendicular to the Sun's rays, placed on the Martian ground, I_\perp , is given by (Appelbaum et al. 1993):

$$I_\perp = I_0 \exp\left(-\frac{\tau}{\cos \theta_z}\right) . \quad (2.13)$$

In this chapter we used the values of the optical depth measured by VL1 and VL2 and reported in Figs. 1 and 2 of Appelbaum and Flood (1990) as a function of the areocentric longitude. The beam irradiance on a surface of arbitrary tilt and orientation I_T is

$$I_T = I_\perp \cos \theta . \quad (2.14)$$

When the surface is horizontal ($\beta=0^\circ$) the direct irradiance I_H is given by

$$I_H = I_\perp \cos \theta_z . \quad (2.15)$$

A multispectral and multiple-scattering calculation was performed in Appelbaum et al. (1993) to determine the wavelength - integrated, net solar flux at a full range of pressure levels in the atmosphere and on the Martian surface. The cited authors obtained a four - dimensional table of the net fluxes parameterized in terms of optical depth, solar zenith angle and surface albedo a . The global solar irradiance G_H on a horizontal surface is then given by the normalized net solar flux function $f(\theta_z, \tau, a)$ according to Eq. (15) of Appelbaum et al. (1993):

$$G_H = I_0 \cos \theta_z \frac{f(\theta_z, \tau, a)}{1-a} . \quad (2.16)$$

The values of the function f are tabulated in Appelbaum et al. (1993) as a function of θ_z and τ for $a=0.1$ (their Figs. 2a and 2b) and for $a=0.4$ (their Figs. 3a and 3b). For other albedo values, the function f can be obtained by interpolating between the values for $a=0.1$ and 0.4 .

The albedo of a surface is the ratio between the reflected component by the surface and the incident solar radiation. The major variations in the Martian albedo during the Viking Landers activities were associated with the two global

dust storms of 1977, which played an important role in the deposition and removal of fine dust materials. Table 1 of (Appelbaum et al. 1993) gives discrete values of the albedo for a range of latitudes and longitudes covering the whole surface of Mars. That table should be used for "clear" days, when the mean and maximum albedo values are 0.27 and 0.40, respectively. For periods with dust storms an albedo function a was introduced in Appelbaum et al. (1993) as follows:

$$a = \max [a_{Table}, \min(0.18\tau, 0.4)] , \quad (2.17)$$

where a_{Table} is the albedo value from their Table 1. This function is based on the assumption that during dust storms the albedo increases above the tabulated value, presumably due to dust deposition on the surface.

The diffuse solar irradiance D_H on a horizontal surface is given by:

$$D_H = G_H - I_H . \quad (2.18)$$

In some particular circumstances the anisotropy of diffuse radiation might be more significant on Mars than on Earth. However, a simple (Liu-Jordan) isotropic model is used here to evaluate the diffuse solar irradiance D_T incident on a tilted surface (Nijegorodov et al. 1994):

$$D_T = \frac{1 + \cos \beta}{2} G_H . \quad (2.19)$$

The reflected solar irradiance R_T on the tilted surface is given by:

$$R_T = a \frac{1 - \cos \beta}{2} G_H . \quad (2.20)$$

The global solar irradiance G_T on the tilted surface is obtained as the sum of the direct, diffuse and reflected irradiances:

$$G_T = I_T + D_T + R_T . \quad (2.21)$$

Note that the diffuse and reflected components of G_T (i.e. D_T and R_T , respectively) depend on the tilt angle β but not on the azimuth angle γ . On the other hand, the direct component I_T is a function of θ (see Eq. (2.14)) which depends on β and γ . Thus, generally G_T depends on ω , β and γ . For a given hour angle ω there exists an optimum tilt and orientation which maximizes the global irradiance G_T .

2.2.1.3 Solar Irradiation on Tilted Surfaces

The daily direct irradiation \tilde{I}_T ($\text{Jm}^{-2}\text{sol}^{-1}$) on a tilted surface is obtained by integrating the direct irradiance between sunrise and sunset

$$\tilde{I}_T = \frac{12.325 \cdot 3600}{12 \cdot 15} \int_{\omega_{sr,T}}^{\omega_{ss,T}} I_T d\omega . \quad (2.22)$$

Here the hour angle is measured in degrees. Similar relationships apply in the case of daily diffuse and reflected irradiations, \tilde{D}_T and \tilde{R}_T ($\text{Jm}^{-2}\text{sol}^{-1}$), respectively

$$\tilde{D}_T = \frac{12.325 \cdot 3600}{12 \cdot 15} \int_{\omega_{sr,T}}^{\omega_{ss,T}} D_T d\omega, \quad (2.23)$$

$$\tilde{R}_T = \frac{12.325 \cdot 3600}{12 \cdot 15} \int_{\omega_{sr,T}}^{\omega_{ss,T}} R_T d\omega. \quad (2.24)$$

The daily global irradiation \tilde{G}_T ($\text{Jm}^{-2}\text{sol}^{-1}$) is given by

$$\tilde{G}_T = \tilde{I}_T + \tilde{D}_T + \tilde{R}_T. \quad (2.25)$$

\tilde{D}_T and \tilde{R}_T depend on the tilt angle while \tilde{I}_T and \tilde{G}_T are functions of both tilt and azimuth angles. Sometimes irradiations on a longer time interval than a single sol are considered. They can be obtained by summing up the daily irradiations for all the sols within that interval. As an example, the yearly solar global irradiation is given by $\sum_{i=1}^{669} \tilde{G}_{T,i}$.

2.2.1.4 Strategies to Collect Solar Radiation

Different strategies can be adopted to collect solar radiation. They can be classified according to the kind of solar energy equipment.

In the case of solar concentrators one wants to have a solar direct irradiance I_T as high as possible. Then, the best strategy (say P) is to keep the receiving surface always perpendicular to the Sun's rays. This requires usually an expensive tracking mechanism with two rotation axes.

In the case of flat plate solar collectors, it is the global solar irradiance G_T which has to be maximized. A preliminary discussion is necessary. Measurements at VL1 and VL2 sites showed that the optical depth during the morning is systematically larger than during the afternoon, a result indicative of the presence of a ground fog in the morning (Pollack et al. 1979). However, in previous studies (Appelbaum and Flood 1990; Appelbaum et al. 1993) the optical depth was considered constant during the day. This approximation will be used here, too. As a consequence, the daily variation of solar global irradiance (and of its components) on a horizontal surface is symmetric around the noon value. In this case it is well known that the optimum azimuth angle is $\gamma=0^\circ$ (i.e. the receiving surface should be South oriented in the Northern Hemisphere and North oriented in the Southern Hemisphere). Thus, only the tilt angle has to be optimized. Two optimization strategies can be envisaged:

- (1) the tilt angle is continuously adjusted (this strategy will be denoted as InOp - instantaneous optimum - in the following) and
- (2) the tilt angle is adjusted from time to time (when the adjustment is made once a day one speaks about a DOp - daily optimum - strategy).

In the first case one needs a tracking mechanism with a single rotation axis, which increases the cost and could decrease the strength of the device. In the second case the system is less expensive and more robust. Of course, the collected solar energy is expected to be larger in the case of the InOp strategy. Choosing between the two strategies should take into account all the above considerations and many other practical aspects related to the particularities of the solar device.

2.2.2 Meteorological Data

There are measurements available on Mars for atmospheric temperature and pressure (both at surface level and at high altitudes), the wind speed on various directions and the solar irradiation. All these data refer to a number of places (mainly from the Northern Hemisphere) and cover all the seasons of a Martian year.

A few basic facts about Martian meteorology are now summarized for the readers convenience. The near-surface atmospheric temperature at Viking Landers latitude ranges from about -17°C during summer to -107°C in winter. Also, the daily average atmospheric pressure ranges between about 7 and 10 hPa, with a minimum near areocentric longitude $L_s=150^\circ$ and a maximum around $L_s=300^\circ$ (Badescu 2001c). The solar constant at the level of Mars orbit is 590 W/m². The solar global irradiance on a horizontal surface at ground level could be as high as 400 W/m² at the noon of a clear summer day and as low as 80 W/m² at midday during a winter dust storm (Badescu 1998).

Some of the available meteorological information for Mars refers to the Viking Landers (VL) sites. The data can be found in the Viking Lander Meteorology and Atmospheric Opacity Data Set Archive from the Planetary Data System database (Lee 1995).Four electronic files from this data set are used here, namely (i) VL_MBIN.DAT and VL_TBIN.DAT (which contain values of surface pressure, ambient temperature and zonal and meridional wind speed), (ii) VL_OPAC.DAT (which contains atmospheric optical depth data) and (iii) VL_AVEP.DAT (which contains daily average pressure values and relevant statistics on a day by day basis).

The file VL_MBIN.DAT (8325 records) contains binned and splined data obtained from the Viking Meteorology Instrument System (VMIS) through portions of the Viking Lander 1 mission. The file VL_TBIN.DAT (25750 records) contains binned and splined data obtained from the VMIS through most of the Viking Lander 2 mission and the early days of the Lander 1 mission. In both cases the numbering of the Martian solar days (sols) started at day 0 when each Lander touched down. Both files VL_MBIN.DAT and VL_TBIN.DAT contain mean values of meteorological parameters in each bin. The bin number starts with 0 at local midnight and increments by 1 each 1/25 of a sol.

The file VL_OPAC.DAT (1044 records) stores measurements of atmospheric optical depth and associated error estimates. Each record contains the optical depth value and its associated local solar time (in Earth hours) beginning at midnight (Lee 1995). The measurements extended over one and a quarter Mars years (1 Mars year = 1.88 Earth years). Note that during great dust storms the optical depth was sometimes so large that the Sun's disk could not be detected.

Consequently, only lower bounds could be estimated for τ in these cases (Pollack et al. 1979; Zurek and Martin 1993).

To produce a consistent set of data the following procedure was adopted. Each record from the files VL_MBIN.DAT and VL_TBIN.DAT was associated with a local solar time t_{solar} (in Earth hours) by using the relation $t_{solar} = (n_{bin} - 1/2)(24/25)$ where n_{bin} is the bin number. As the files VL_MBIN.DAT and VL_TBIN.DAT are more "dense" than the VL_OPAC.DAT file, a time interpolation procedure was used to obtain the values of the atmospheric pressure and temperature and wind velocity associated with a given record of atmospheric optical depth. As a time interpolation parameter we used here the local solar time rather than the areocentric longitude of the Sun. Details about the interpolation procedure and its validation are given in Badescu (1998b).

The atmospheric pressure values are missing for eleven records. In these cases the following procedure was adopted:

- (i) in case of the record VL2, year 1, areocentric longitude = 229.808° (autumn), sol = 200, hour = 11.41, we used the arithmetic mean of the pressure values corresponding to sol = 199, hour = 13.87 and sol = 201, hour = 11.42, respectively.
- (ii) in the case of the record VL2, year 2, areocentric longitude = 19.228° (spring), sol = 454, hour = 9.62, we used the arithmetic mean of the pressure values corresponding to sol = 453, hour = 17.33 and sol = 454, hour = 17.25, respectively.
- (iii) in the case of VL2, year 2, the first nine records during summer (areocentric longitudes between 93.363° and 100.510°) we used the pressure value corresponding to the tenth record (areocentric longitude = 101.697° , sol = 634, hour = 6.98).

As a result of the interpolation procedure we obtained a set of *complete* records. Each complete record associates to a given sol number and a local solar time value a set of four meteorological values, namely optical depth τ and atmospheric temperature T_a , atmospheric pressure p_a and wind speed w . The number of *complete* records is rather small (Table 2.2).

Table 2.2 The number of *complete* meteorological data records available at VL1 and VL2 sites

		Viking Lander 1		Viking Lander 2	
		Sol numbers	Number of complete records	Sol numbers	Number of complete records
Year 1	spring	-	-	-	-
	summer	6 to 155	22	6 to 111	37
	autumn	168 to 306	108	120 to 257	77
	winter	307 to 350	56	269 to 415	55
Year 2	spring	-	-	416 to 608	147
	summer	-	-	610 to 784	58
	autumn	-	-	792 to 872	25
	winter	-	-	-	-

The file VL_AVEP.DAT contains summary pressure data presented on a sol-by-sol basis (3297 recordings). The data set consists of values of the daily average pressure, minimum daily pressure, maximum variation in pressure during each sol and the standard deviation in the daily average pressure. The file contains data for VL1 (sols 0 to 2245) and VL2 (sols 0 to 1050). The data are not continuous.

Many procedures detailed in Sect. 2.3 use as input various parameters including the value of the atmospheric optical depth. However, the number of available optical depth data is relatively small as compared to other meteorological parameters. To extend the Viking record of dust opacity beyond the period of lander Sun diode measurements, less direct indicators of substantial overhead dust opacity must be used as proposed in Zurek and Martin (1993). These include the observed visual darkening of the scene, the restricted diurnal range in the near-surface temperatures and the daily variations of surface pressure and near-surface winds. In Sect. 2.3 we propose a number of statistical correlations between the atmospheric optical depth and some indicators related to the daily variation of atmospheric pressure on Mars surface.

2.2.3 Results and Discussions

In order to compare the different strategies defined in Sect. 2.2.1 we made computations for both VL1 and VL2 locations. We selected a number of particular sols which are representative for a given season and optical depth (Table 2.3). Various actinometric parameters (direct and global irradiances and irradiations) have been maximized numerically. The optimization variable was usually the tilt angle of the surface receiving solar radiation.

Table 2.3 Particular sols selected for analysis

Location	Sol number	Season	Weather	Optical depth
VL1	88	spring	"Clear" day	0.5
	421	autumn	Dust storm	2.8
	507	autumn	Between two dust storms	1.15
VL2	31	spring	"Clear" day	0.45
	355	summer	"Clear" day	0.3
	562	winter	Dust storm	2.3

First we refer to the solar devices designed to collect global solar radiation. It is important to know if the continuous adjustment of the tilt angle (the InOp strategy) leads to a significant increase in the collected solar energy. We compared the results of the InOp strategy with the solar global irradiance incident on a horizontal surface (H).

During the "clear" springtime days the direct irradiance at VL1 site exceeds the diffuse irradiance during most part of the day (Fig. 2.1a). The continuous adjustment of the tilt angle doesn't increase significantly the amount of collected solar global radiation. This is mainly due to the fact that the optimum tilt angle is

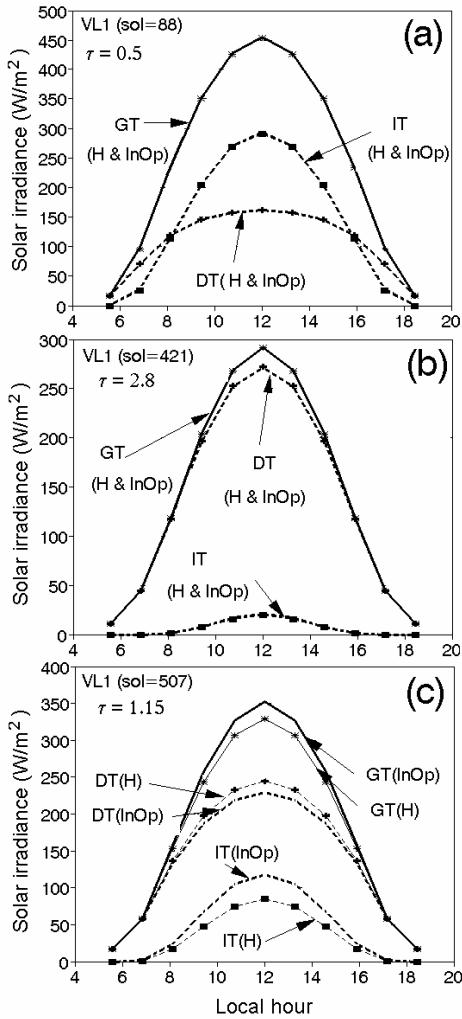


Fig. 2.1 Direct, diffuse and global solar irradiance (IT, DT and GT, respectively) for VL1 site. A horizontal surface (H) and a surface whose tilt angle is continuously optimized (InOp strategy) are considered. (a) sol number = 88, optical depth = 0.5; (b) sol number = 421, optical depth = 2.8; (c) sol number = 507, optical depth = 1.15

relatively close to $\beta=0^\circ$, i.e. the optimum tilted receiving surface is nearly horizontal (see Fig. 2.3a). In the middle of the autumn dust storm the direct solar radiation is very low but there is a relatively large diffuse component (Fig. 2.1b). Again, there is no real need for the continuous adjustment of the tilt angle because its optimum value is close to $\beta=0^\circ$ (see Fig. 2.3a). However, when the dust storm decreases in intensity, the optical depth decreases, too. The diffuse component exceeds the direct component and the continuous adjustment of the tilt angle

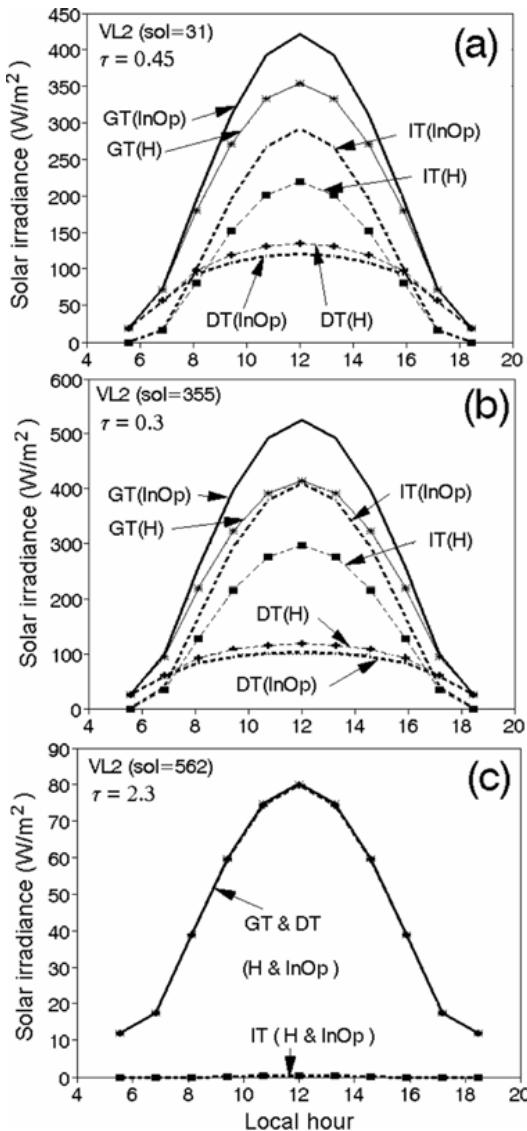


Fig. 2.2 Direct, diffuse and global solar irradiance (IT, DT and GT, respectively) for VL2 site). A horizontal surface (H) and a surface whose tilt angle is continuously optimized (InOp strategy) are considered. **(a)** sol number = 31, optical depth = 0.45; **(b)** sol number = 355, optical depth = 0.3; **(c)** sol number = 562, optical depth = 2.3

increases the collected global solar irradiance G_T (Fig. 2.1c). Note that the tilt angle adjustment increases the direct component but decreases the diffuse component. The optimum tilt angle is significantly different from $\beta=0^\circ$ (see Fig. 2.3a).

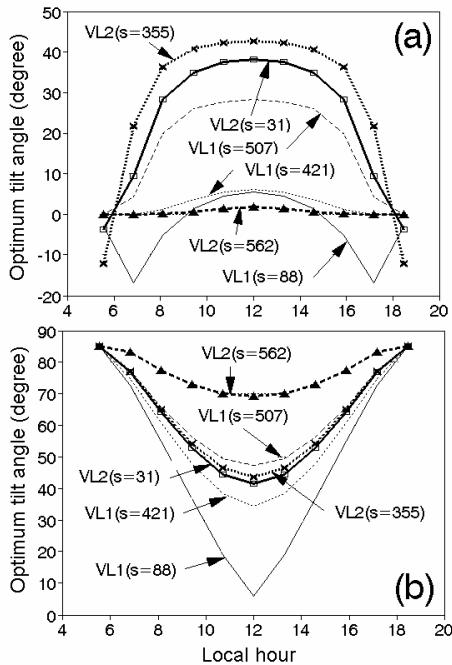


Fig. 2.3 (a) Daily variation of the optimum tilt angle allowing maximum incident global solar irradiance, G_T (InOp strategy). The cases of Figs. 2.1 and 2.2 are considered. **(b)** Daily variation of the optimum tilt angle allowing maximum incident direct solar irradiance, I_T (P strategy). The cases of Figs. 2.1 and 2.2 are considered again, but the surface azimuth angle is continuously adjusted to follow the Sun's azimuth angle

During the spring "clear" days at VL2 site the continuous adjustment of the tilt angle significantly increases the global solar irradiance as compared to the value on a horizontal surface (Fig. 2.2a). It is interesting to note that this increase is associated with an increase of the direct component and a decrease of the diffuse component. The same remarks apply to the summer "clear" days (Fig. 2.2b). In both cases the optimum tilt angle of the InOp strategy is significantly different from zero (Fig. 2.3a). During the winter dust storm days the direct component vanishes and the diffuse component is very small (Fig. 2.2c). The optimum tilt angle is very close to $\beta=0^\circ$ (Fig. 2.3a).

When the direct solar irradiance I_T has to be maximized (strategy P), the optimum tilt angle has a daily variation which is obviously different from that obtained in the case of maximizing G_T (compare Fig. 2.3b and Fig. 2.3a). Because the receiving surface is perpendicular to the Sun's rays the optimum tilt angle β_{opt} equals the solar zenith angle. Note that the azimuth angle γ of that surface has to be continuously adjusted, too. The optimum adjustment of β and γ doesn't depend on the optical depth τ but is a function of astronomical and site-specific parameters only.

The results above show that the continuous adjustment of the tilt angle (InOp strategy) is not always justified when the maximum solar global irradiance has to be collected. For some particular cases, characterized by a given location on the Mars surface, time of the year and atmospheric optical depth, the DOp strategy may lead to a performance comparable to that of InOp. Thus, using a receiving surface of a constant slope during the day could be an attractive technical solution due to its simplicity. Additional information is needed, however, in order to understand the influence of the surface orientation.

This could be of interest if solar radiation is used as an energy source for transportation. We computed the direct, diffuse and global daily irradiations for surfaces of different tilts and orientations kept constant for the same day. The sols of Table 2.3 were tested again. Some results are given in Figs. 2.4-2.7. In the case of VL1 site the following remarks were made. In the spring "clear" day (sol=88) the daily direct irradiation obviously depends on the tilt angle of the receiving surface (Fig. 2.4a). The influence of the azimuth angle is small for tilt angles close to the optimum but increases at higher slope angles. The daily global irradiation shows a dependence on β and γ similar to that of the direct irradiance (compare Figs. 2.5a and 2.4a). This is mainly due to the fact that the diffuse component depends on the tilt angle β only. During the autumn dust storm day (sol = 421) the direct daily irradiation is very small (Fig. 2.4b). However, an optimum orientation (azimuth angle 0°, tilt angle around 30°) is clearly pointed out. When the global irradiation is considered, the optimum orientation exists, too, but it is less obvious than in the case of the direct irradiation (compare Figs. 2.5b and 2.4b).

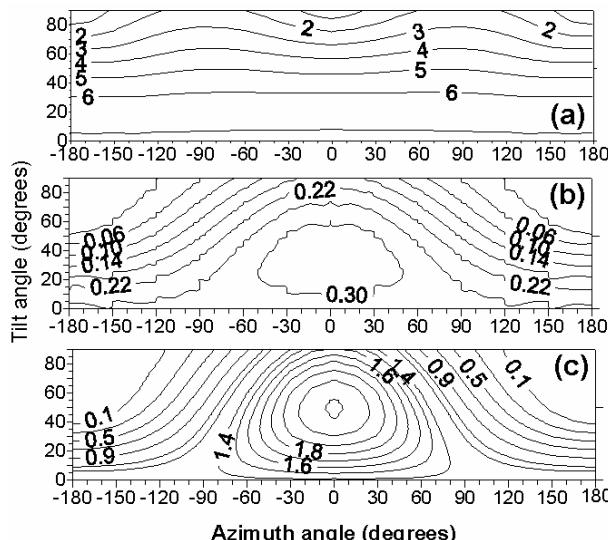


Fig. 2.4 Daily direct solar irradiation ($\text{MJ}/\text{m}^2/\text{sol}$) incident on surfaces of different orientations at the VL1 site. (a) sol number = 88, optical depth = 0.5; (b) sol number = 421, optical depth = 2.8; (c) sol number = 507, optical depth = 1.15

The rather large value of the global irradiation is mainly due to its diffuse component. In the case of the other autumn day (sol=507), whose optical depth is 1.15, an optimum orientation exists, too (azimuth angle 0°, tilt angle around 50°) (Fig. 2.4c). In this case the direct solar irradiation is relatively important and its magnitude is very sensitive to the azimuth angle. There is an optimum orientation for the daily global irradiation, too (Fig. 2.5c). However, the optimum tilt angle (around) 25° is obviously different from that we obtained for the direct irradiation (compare Figs. 2.5c and 2.4c). Also, the dependence of the global irradiation on the azimuth angle is less important than in the case of the direct irradiation. This remark could be useful in the case of a solar powered vehicle. Indeed, during movement, the vehicle changes its position as reported to the Sun. In order to avoid large fluctuations in the received energy it seems more appropriate to use global rather than direct solar radiation. This means that the systems using unconcentrated solar energy should be preferred to those based on solar concentrators.

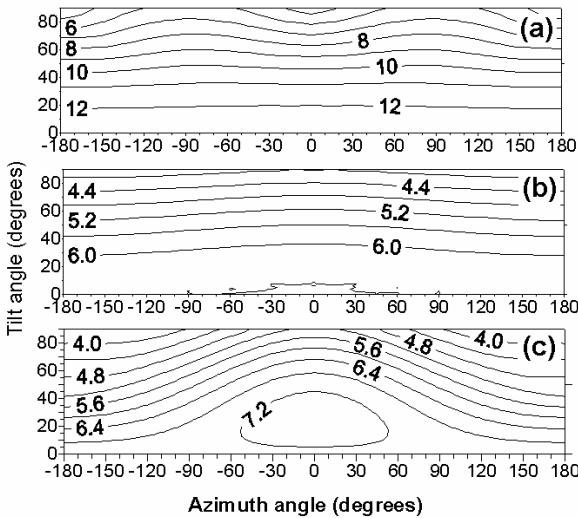


Fig. 2.5 Daily global solar irradiation ($\text{MJ}/\text{m}^2/\text{sol}$) incident on surfaces of different orientations at the VL1 site. (a) sol number = 88, optical depth = 0.5; (b) sol number = 421, optical depth = 2.8; (c) sol number = 507, optical depth = 1.15

When the VL2 site is considered one sees again an obvious optimum orientation in the case of the "clear" spring and summer days (Figs. 2.6a and 2.6b). The direct daily irradiation has a strong dependence on the azimuth angle. For the same "clear" days the global daily irradiation is a maximum for an optimum tilt angle whose values are close to those obtained for the direct daily irradiation (compare Figs. 2.7a and 2.7b, and Figs. 2.6a and 2.6b, respectively). There is a difference as compared to the results obtained in the case of the VL1 site. Indeed, in the last case the optimum tilt angles for direct and global radiation, respectively, were different

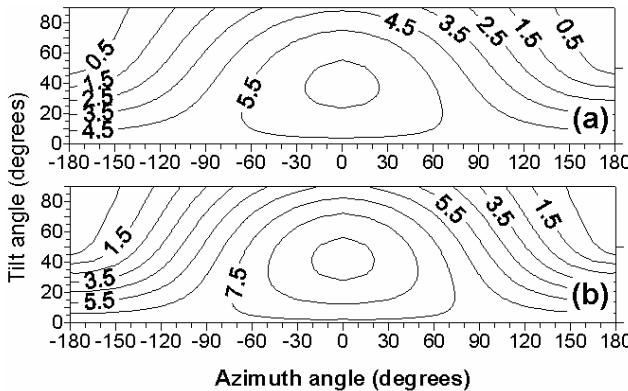


Fig. 2.6 Daily direct solar irradiation ($\text{MJ}/\text{m}^2/\text{sol}$) incident on surfaces of different orientations (VL2 site) **(a)** sol number = 31, optical depth = 0.45; **(b)** sol number = 355, optical depth = 0.3

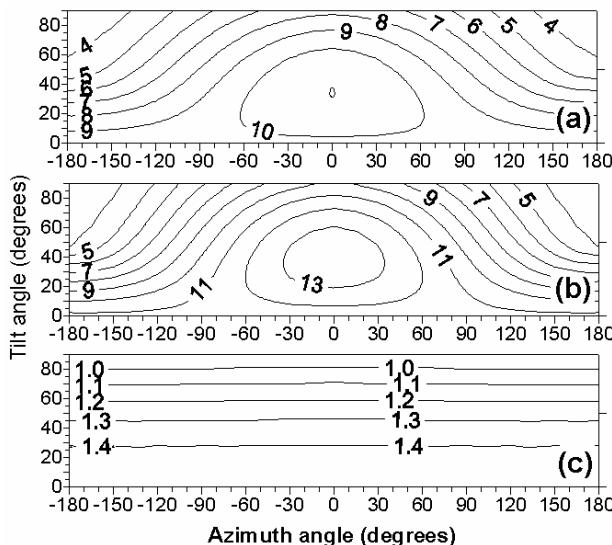


Fig. 2.7 Daily global solar irradiation ($\text{MJ}/\text{m}^2/\text{sol}$) incident on surfaces of different orientations (VL2 site). **(a)** sol number = 31, optical depth = 0.45; **(b)** sol number = 355, optical depth = 0.3; **(c)** sol number = 562, optical depth = 2.3

during the "clear" days (compare Figs. 2.4a and 2.5a). The reason is that at VL2 site the direct radiation is the main component while at VL1 location both direct and diffuse components have a comparable magnitude (see Figs. 2.4a and 2.5a for VL1 and Figs. 2.6a and 2.7a for VL2 site). During spring and summer "clear" days

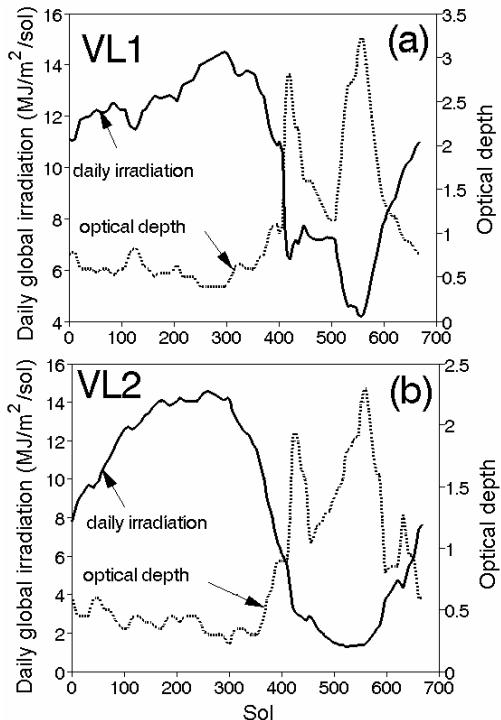


Fig. 2.8 Yearly variation of the atmospheric optical depth and daily solar global irradiation on a horizontal surface on Mars. **(a)** VL1 site and **(b)** VL2 site

the dependence of global daily irradiation \tilde{G}_T , on the azimuth angle is less important than in the case of direct irradiation (compare, for example, Figs. 2.6a and 2.7a). In the dust storm winter days ($\text{sol} = 562$) the direct irradiation practically vanishes while the global daily irradiation is very small (Fig. 2.7c). The optimum tilt angle is close to $\beta=0^\circ$. There is practically no dependence of \tilde{G}_T on the azimuth angle as the diffuse radiation is the only component of the global radiation.

A more complete picture about the opportunity of using solar radiation as an energy source is obtained by studying its availability during the whole Martian year. Figures 2.8a and 2.8b show the yearly variation of the atmospheric optical depth for both VL1 and VL2 sites. The information was derived by discretizing the data shown in Figs. 1 and 2 of Appelbaum and Flood (1990). By using the optical depth data we computed the daily global irradiation on a horizontal surface (Figs. 2.8a and 2.8b). The results we obtained are in good agreement with those reported in Fig. 9 of Appelbaum et al. (1993). There are, however, some minor differences. As an example, in the case of VL1 location the first peak of the optical depth around sol 425 determines a local minimum in the daily solar global irradiation (Fig. 2.8a). This local minimum is not represented on Fig. 9a of

Appelbaum et al. (1993). As one sees, in the Northern Hemisphere of Mars the spring and summer (sols 1-372) are generally "clear" and the level of global solar irradiation on horizontal surface is high. However, during autumn and winter (sols 373-669), when dust storms occur frequently, the daily global solar irradiation strongly decreases. During these seasons diffuse solar radiation usually exceeds the direct component.

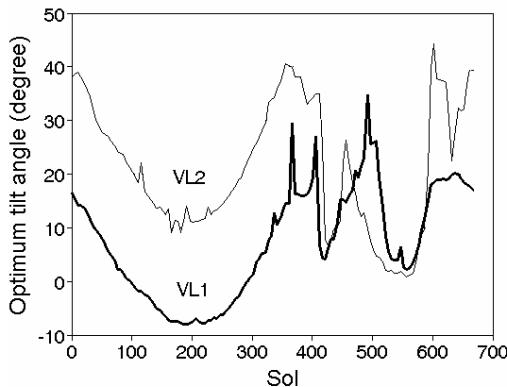


Fig. 2.9 Yearly variation of the daily optimum tilt angle $\beta_{opt,day}$ at VL1 and VL2 sites

The results above proved that the DOp strategy (once per day adjustment of the receiving surface tilt angle) is a reliable option. The yearly variation of the daily optimum value $\beta_{opt,day}$ of the tilt angle is shown in Fig. 2.9 for both VL1 and VL2. During spring and summer (sols 1-372) the time variation of $\beta_{opt,day}$ is similar at both locations because it is mainly controlled by the astronomical parameters. During the dust storm seasons (autumn and winter, sols 373-669) the time variation of $\beta_{opt,day}$ is mainly controlled by the atmospheric optical depth. Generally, when the optical depth increases, the direct component decreases and the incident solar flux mainly consists of diffuse radiation. Consequently, the optimum tilt angle decreases. This explains why the local minimum values of $\beta_{opt,day}$ correspond to those sols when the optical depth is a local maximum (compare Fig. 2.9, on one hand, and Figs. 2.8a and 2.8b, on the other hand).

Four strategies to collect solar radiation were analyzed during the whole Martian year at VL1 and VL2 sites. In order of increasing technical complexity they are: H (a horizontal receiving surface), DOp, InOp and P. Some of the results obtained are shown in Figs. 2.10 and 2.11.

For the VL1 location the following remarks are made. The best strategy to collect direct and global radiation is to keep the receiving surface perpendicular on Sun's rays (strategy P: see Figs. 2.10a and 2.10c). During the "clear" seasons (spring and summer) this strategy is obviously better than the three other strategies. However, during the dust storm seasons (autumn and winter) the P-strategy is only slightly better than the other ones. The strategies H, InOp and

DOp give comparable results during the whole year. Consequently, keeping the receiving surface horizontal (H strategy) looks quite attractive because it is the simplest from a technological point of view. The H strategy collects the maximum diffuse solar radiation, as expected, while the P strategy is the worst from this view-point (Fig. 2.10b).

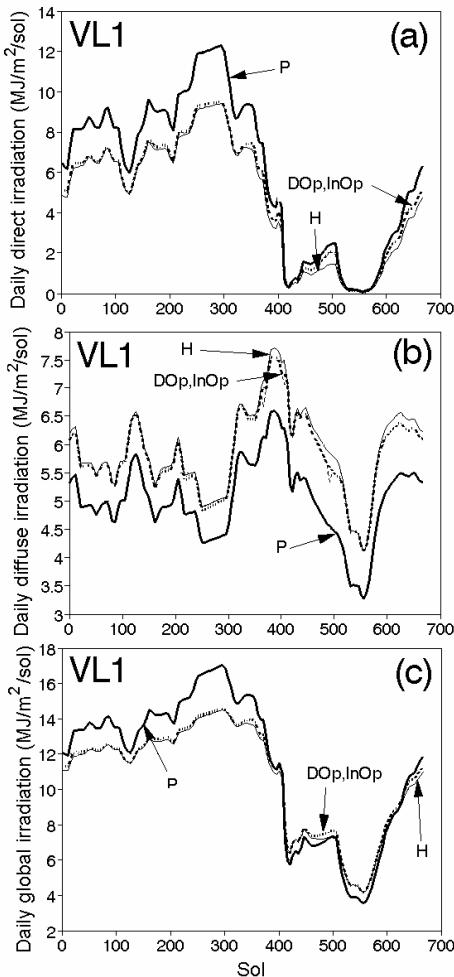


Fig. 2.10 Yearly variation of the daily direct, diffuse and global solar irradiation at VL1 site. Four strategies to collect solar energy were considered: H - horizontal receiving surface; DOp - South oriented surface with tilt angle adjusted once a day; InOp - South oriented surface with tilt angle continuously adjusted to receive the maximum global solar irradiance; P - receiving surface perpendicular to Sun's rays. (a) daily direct solar irradiation, (b) daily diffuse solar irradiation and (c) daily global solar irradiation

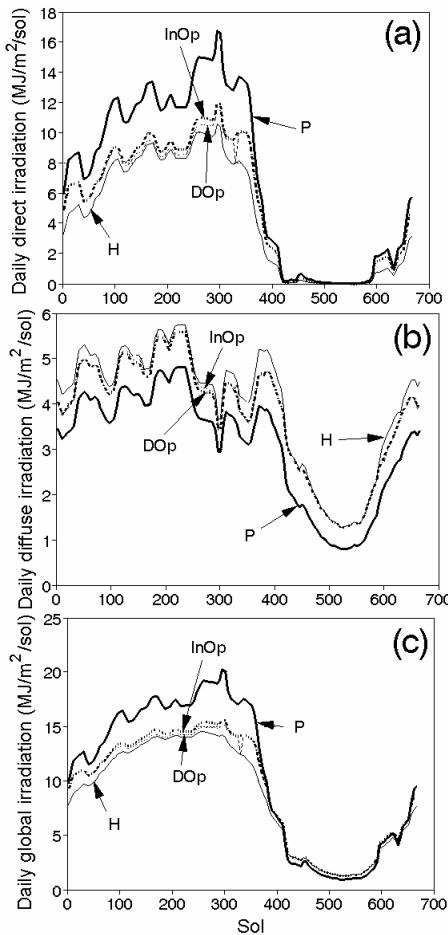


Fig. 2.11 The same as Fig. 2.10 but for VL2 site

For the VL2 location the above remarks roughly maintain (Fig. 2.11). The P strategy is the most appropriate way of collecting direct and global irradiation in the "clear" seasons (spring and summer, i.e. sols 1 to 372) (Figs. 2.11a and 2.11c). For the higher latitudes of VL2 the difference among the performance of the four strategies is more important than in the case of the places closer to the equator (compare Fig. 2.11a and Fig. 2.10a).

Let us look to the daily global solar irradiation during the dust storm seasons at VL1 and VL2 sites (Figs. 2.10c and 2.11c, respectively). There is a very small difference in the performance of all the four strategies. This means that tilting the receiving surface in order to avoid the dust deposition doesn't decrease significantly the amount of collected solar energy.

The results above emphasize rather small differences between the energy collected using the strategies H, InOp and DOp, respectively. Of course, the

simpler strategies, namely DOp and H should normally be preferred. However, our results are based on measurements performed on a relatively short time interval at VL1 and VL2 sites. Of course, one expects the time variation of the tilt angle shown in Fig. 2.9 be the same on days with small atmospheric optical depth and similar areocentric longitudes. But the dust storms could strongly modify this time-dependence. In these cases the InOp strategy becomes attractive as it is based on sensors which maximize the "instantaneous" global irradiance.

Table 2.4 Optimum tilt angles for a South oriented surface placed at VL1 and VL2 sites (Northern Hemisphere)

Location	Time interval	Optimum tilt angle (degrees)	Maximum global irradiation (MJ/m ² /interval)
VL1	spring	1.24	2368.1
	summer	0.57	2459.5
	autumn	16.34	1186.9
	winter	15.58	1144.0
	spring & summer	0.99	4824.5
	autumn & winter	15.97	2366.0
	whole year	4.97	7102.5
	spring	20.81	2408.7
	summer	22.87	2493.1
VL2	autumn	28.00	554.0
	winter	30.26	566.9
	spring & summer	21.84	4929.0
	autumn & winter	29.46	1150.9
	whole year	22.94	6026.5

Values obtained by maximizing the incident solar global irradiation during the specified time intervals

A number of strategies simpler than DOp could also be envisaged. For example, the tilt angle could be adjusted once per season, per semester or per year, respectively. Computations were made for both VL1 and VL2 sites (Table 2.4). As one can see, using a single tilt angle adjustment per year means to lose less than 1 % from the energy collected by using two adjustments per year. This applies for both VL1 and VL2 locations. However, extrapolations to the future should be made with caution and additional measurements are required for a better knowledge and understanding of phenomena.

2.3 Weather Forecasting

The areas around both Viking Landers are flat (ensuring secure landing) and not too close to the Martian poles (the solar radiation energy flux needed by the photovoltaic cells was rather large). Thus, one could expect that these regions will be used as landing sites for future Martian missions. As an example, the 1997

Pathfinder mission landed in Ares Vallis at 19.17°N and 33.21°W (about 815 km distant from VL1 site). This gives a broad perspective to the meteorological data collected by the Viking mission.

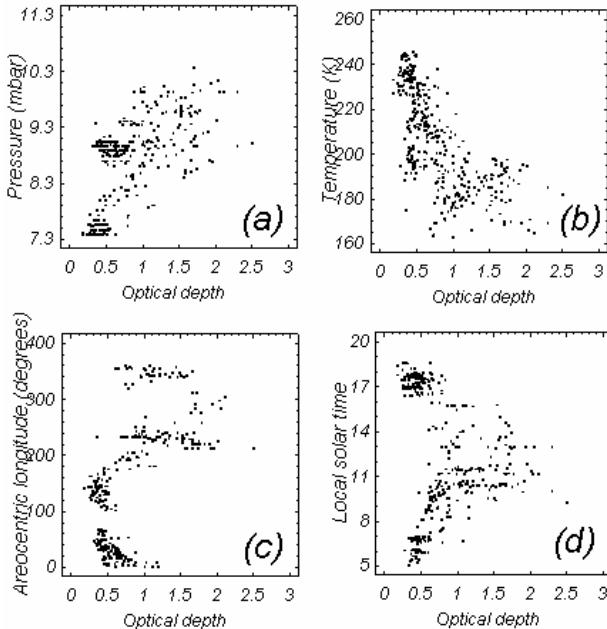


Fig. 2.12 Scatter plots obtained by using all *complete* recordings available for VL2 site. **(a)** Near-surface atmospheric pressure vs optical depth; **(b)** Near-surface temperature vs optical depth; **(c)** Areocentric longitude vs optical depth; **(d)** local solar time vs optical depth

Figure 2.12 shows a number of scatter plots obtained by using all the *complete* records available at VL2 site. Roughly, the optical depth increases when the near-surface atmospheric pressure increases (Fig. 2.12a). The large data scatter is due to the well-documented seasonal change. On average, the optical depth increases when the ambient temperature decreases (Fig. 2.12b). However, at low optical depth values there is a larger range of variation for the temperature. Consequently, one could expect that the ambient temperature is not the best regression parameter on clear skies. There is no obvious statistical proportionality between optical depth and areocentric longitude (Fig. 2.12c). The available data at VL2 site covers all the four seasons, with a higher density during spring and summer (or, in other words, during the clear sky season). The optical depth is small during spring and summer and large during autumn and winter, as known. Data is available for almost all times of the day (Fig. 2.12d). There is no obvious statistical proportionality between optical depth and local solar time. At sunrise and sunset the optical depth is lower. The morning optical depth values are somewhat higher than the near-sunset values. This is in agreement with reports of Colburn et al. (1989) for VL1 site,

who explained the morning enhancement through the presence of water ice hazes and fogs.

The above qualitative features also apply to VL1 data, which contain larger optical depth values.

2.3.1 Inference of Atmospheric Optical Depth

A number of simple relationships between atmospheric optical depth and various meteorological parameters were tested. First, the following two-parameter non-linear relations were fitted to the data (least square method):

$$\tau = ap^b + cL_s^d, \quad \tau = aT^b + cL_s^d. \quad (2.26,27)$$

Here p , T and L_s are atmospheric pressure (hPa), ambient temperature (K) and areocentric longitude (degrees), while a to c are coefficients to be computed by the regression. Use of the areocentric longitude as a regression parameter follows (Pollack et al. 1979) where the authors noted that during the decay phase of the 1977 dust storms the logarithm of the optical depth had an approximately linear dependence on time. (In fact each dust storm is characterized by two phases, which are distinguished by the slope of the curve). Equations (2.26) and (2.27) were applied to three data sets of complete recordings, namely (i) all available data at VL1 site, (ii) all available data at VL2 site and (iii) all available data at both VL1 and VL2 sites. Results are given in Table 2.5 and show that pressure is a somewhat better regression parameter than ambient temperature. The accuracy of the regression is worse in case of both VL1 data and the whole data set (i.e. VL1 & VL2).

Table 2.5 Regression coefficients in Eqs. (2.26) and (2.27)

Viking Lander site	Number of recordings	Equation	a	b	c	d	MSE
VL1	186	Eq (2.26)	1.69570	-0.8075	0.00067	1.28059	0.3724
		Eq (2.27)	0.02525	0.42188	0.00059	1.16468	0.3850
VL2	388	Eq (2.26)	0.00124	2.74616	0.00826	0.66346	0.0968
		Eq (2.27)	6793.30	-1.7650	0.00046	0.96340	0.1098
VL1&VL2	574	Eq (2.26)	0.02958	1.22180	0.02003	0.59641	0.3175
		Eq (2.27)	6.05091	-0.5367	0.00976	0.59213	0.3347

Units to be used in Eqs. (2.26) and (2.27): atmospheric pressure p (hPa), ambient temperature T (K) and areocentric longitude L_s (degrees). Mean square error (MSE) is defined as the residual sum of squares divided by the number of recordings

Figure 2.13a shows the best-fit Eq. (2.26) applied to VL2 data. The larger residuals are associated with the higher optical depth values, as expected, because only lower bounds could be estimated for τ in these cases.

Second, the following three parameter linear relation was fitted to the three data sets above:

$$\tau = a + bp + cT + dL_s . \quad (2.28)$$

The accuracy of fitting Eq. (2.28) to the data is similar to that of Eqs (2.26) and (2.27). Figure 2.13b shows the best-fit Eq. (2.28) applied to VL2 data. The smaller residuals are associated with optical depth values around 0.5.

The period $L_s = 160^\circ - 330^\circ$ and variants thereof have been referred to as the “dust storm season”. This period is nearly centred on perihelion, which is the time of maximum insolation on Mars (Martin and Zurek 1993). The two 1977 dust storms (A and B) lasted between $L_s=210^\circ$ to 230° and 280° to 300° , respectively (Hourdin et al. 1995). Pollack et al. (1979) noted that the optical depth had approximately the same value ($\tau \sim 1$) just prior to the onset of both 1977 A and B global dust storms at both VL1 and VL2 sites. These authors used $\tau = 1$ as a generic threshold between a “dust - free” atmosphere and a significantly dusty atmosphere, respectively.

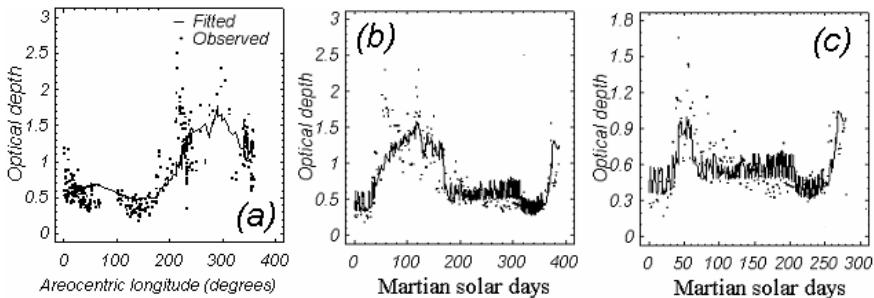


Fig. 2.13 Observed and fitted optical depth values at VL2 site. (a) Fit using Eq. (2.26). All complete recordings available were used. (b) Fit using Eq. (2.28). All complete recordings available were used. (c) Fit using Eq. (2.28). All complete recordings associated to clear sky periods were used

The three previous data sets (i) to (iii) were split into six subsets as follows: three “clear” sky data sets (data associated with optical depths $\tau < 1$) and three dust storm data sets (data associated with optical depths $\tau > 1$). The accuracy of the regression Eq. (2.28) for the three dust storm data sets was relatively poor. During the three “clear” sky data sets the accuracy was good; the values of the regression coefficients are listed in Table 2.6.

There is an obvious increase in accuracy when the regression uses clear sky data only (compare the MSE values in Tables 2.6 and 2.5, respectively). Figure 2.13c shows the best-fit Eq. (2.28) applied at VL2 site during clear sky periods. A few optical depth values larger than 1 were included in order to keep the data sequence continuous.

Table 2.6 Coefficients in Eq. (2.28) when applied to “clear sky” periods

Viking Lander site	Number of recordings	a	b	c	d	MSE
VL1	30	-3.655	0.463	0.00052	0.00606	0.0051
VL2	280	0.3408	0.137	-0.00474	0.00085	0.0378
VL1&VL2	310	0.7601	0.094	-0.00494	0.00087	0.0400

Units to be used in Eq. (2.28): atmospheric pressure p (hPa), ambient temperature T (K) and areocentric longitude L_s (degrees). Mean square error (MSE) is defined as the residual sum of squares divided by the number of recordings

The data sets associated with the dust storm season were also split, into “moderate” opacity data sets ($1 < \tau < 2$) and “high” opacity data sets ($2 < \tau < 4$). No significant increase in the regression accuracy was obtained.

In average, the optical depth increases when the atmospheric pressure increases (Fig. 2.12a). This suggested to split the three sets (i) to (iii) into nine subsets according to their atmospheric pressure. The following relationships were derived as best fits to these data subsets:

$$\tau = a + bp, \quad \tau = a + bp + cT, \quad \tau = a + bp + cT + dL_s + eh. \quad (2.29, 30, 31)$$

Here h is the local solar time (in Earth hours). Figure 2.14a shows that regression’s accuracy increases by increasing the relation’s complexity (i.e. the number of regression coefficients), as expected. Table 2.7 shows some results.

Table 2.7 Coefficients in Eqs. (2.29), (2.30) and (2.31) when applied to VL1&VL2 data

Pres- sure (hPa)	Number of records.	Equation	a	b	c	d	e	MSE
7 – 8	127	Eq (2.29)	-0.6940	0.16815	-	-	-	0.1186
		Eq (2.30)	0.62996	0.11613	-0.0042	-	-	0.1137
		Eq (2.31)	-0.9702	0.17740	-0.0042	0.00741	0.00366	0.0601
8 - 9	238	Eq (2.29)	11.6704	-1.2090	-	-	-	0.3593
		Eq (2.30)	14.4852	-1.0802	-0.0188	-	-	0.2657
9–10	192	Eq (2.31)	10.5692	-0.8625	-0.0099	-0.00165	-0.00985	0.2451
		Eq (2.29)	-0.5965	0.21826	-	-	-	0.6391
		Eq (2.30)	7.50937	-0.3028	-0.0171	-	-	0.6193
		Eq (2.31)	-6.8110	0.31131	0.0238	0.00703	-0.06515	0.3362

Units to be used in Eqs. (2.29), (2.30) and (2.31): atmospheric pressure p (hPa), ambient temperature T (K), areocentric longitude L_s (degrees) and time h (Earth hours). Mean square error (MSE) is defined as the residual sum of squares divided by the number of recordings

Equation (2.31) gives a significant increase in accuracy at higher atmospheric pressure (9–10 hPa at VL1 site and 10–11 hPa at VL2 site). Therefore, the influence of the areocentric longitude and local solar time on the regression's accuracy is important mainly during the dust storm season. Generally, the regression accuracy is very good at low atmospheric pressure (i.e. on clear sky days). This is obvious for the pressure range 7–8 hPa at both VL1 and VL2 sites. The regression accuracy is better at VL2 site than at VL1 site, especially at higher atmospheric pressure. This is a consequence of the higher optical depth at VL1 site during the dust storm season (compare Fig. 2.12a and Fig. 2.14d, for instance).

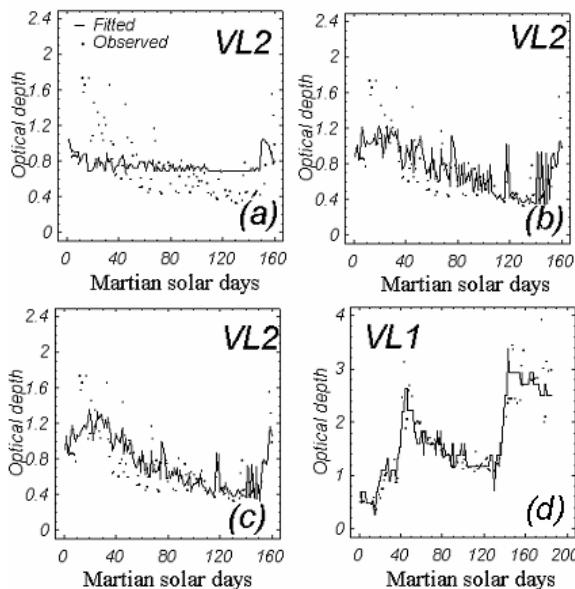


Fig. 2.14 Observed and fitted optical depth values at VL1 and VL2 sites. (a) Values computed by Eq. (2.29); (b) Values computed by Eq. (2.30); (c) Values computed by Eq. (2.31). (d) Values computed by Eq. (2.32). All available complete recordings were used

Another correlation we analyzed is that between optical depth and the daily variation of surface pressure Δp . This was inspired by Zurek and Martin (1993) who noted that among the indirect indicators of overhead dust opacity the daily variation of surface pressure is the least sensitive to local or regionally confined dust hazes. They concluded that Δp is strongly correlated to the optical depth during planet-encircling dust storms. Another indicator of the atmospheric pressure diurnal excursion is the standard deviation of daily surface pressure $\sigma_{p,day}$. This spreading indicator was studied in Leovy et al. (1985), Tillman (1985, 1988). Discussions concerning the signature of local and non-local dust storms on the diurnal (both once daily and twice daily) variation of surface pressure can be found in Zurek and Leovy (1981), Tillman (1988).

A number of relationships based on Δp and $\sigma_{p,day}$ were tested. Here we report the following simple linear relations

$$\tau = a + b\Delta p, \quad \tau = a + b\sigma_{p,day}. \quad (2.32, 2.33)$$

Both Δp and $\sigma_{p,day}$ enter in hPa. Equations (2.32) and (2.33) were applied to all available data at VL1 and VL2 sites. Results are shown in Table 2.8.

Table 2.8 Coefficients to be used in Eqs. (2.32) and (2.33)

Viking Lander site	Number of recordings	Equation	a	b	MSE
VL1	186	Eq (2.32)	0.05331	2.48007	0.1132
		Eq (2.33)	0.00575	8.98674	0.1171
VL2	388	Eq (2.32)	0.29446	2.03435	0.0835
		Eq (2.33)	0.25437	7.32211	0.0966
VL1 & VL2	574	Eq (2.32)	0.23418	2.23702	0.0956
		Eq (2.33)	0.18513	8.12949	0.1056

Units to be used in Eqs. (2.32) and (2.33): the daily variation of surface pressure Δp (hPa), the standard deviation of daily surface pressure $\sigma_{p,day}$ (hPa). Mean square error (MSE) is defined as the residual sum of squares divided by the number of recordings

The accuracy of both Eqs. (2.32) and (2.33) is good (compare the MSE values from Table 2.8 with those from the other tables). This recommends Δp as the best optical depth regression parameter during both clear sky and dust storm seasons. Figure 2.14d shows the best-fit in Eq. (2.32) applied to VL1 data (which is more difficult to be fitted than the VL2 data, because of the larger optical depth values during the dust storm season). The curve is well fitted to the data for both small and large optical depth values.

The accuracy of fitting the data by using the daily standard deviation of the surface pressure instead of its daily variation is slightly worse (Table 2.8). More details are given in Badescu (2001a).

2.3.2 Forecasting the Daily Mean of Atmospheric Pressure

The atmospheric surface pressure on Mars was intensively studied. An important source of information are the archives of the two Viking expeditions. They contain values related to the daily variation of pressure as well as daily averaged values. These data have been used to study the Martian transient planetary waves (Barnes 1980, 1981; Tillman 1988), for comparison with the results of seasonal energy balance models (Paige and Ingersoll 1985; Paige and Wood 1992; Wood and Paige 1992; Pollack et al. 1993) and for direct comparisons with general circulation models (GCM) simulations (Hourdin et al. 1993). Also the daily average statistics and the associated atmospheric phenomena have been studied (Tillman 1988). Now we know that the seasonal oscillations of the pressure are mainly

caused by the variation of the atmospheric mass (during the sublimation/condensation of the carbon dioxide in the polar caps) and to the internal latitudinal mass redistribution associated with the atmospheric circulation (Hourdin et al. 1995). Also, fast surface pressure oscillations with periods of a few sols are due to the transient planetary waves, which are present mainly during autumn and winter in the Northern Hemisphere.

The short-term time series of the daily average pressure on Mars's surface have specific sequential and distribution characteristics, which have been rarely studied. In Badescu (2001b) both features were analyzed by using data from the Viking Lander Meteorology and Atmospheric Opacity Data Set Archive. As a first step, the surface pressure was separated into a slow varying term (the "average" pressure) and a fast varying term (the "excess" pressure). The average time variation of the daily pressure was fitted by using a formula retaining the first eight harmonics of the annual cycle. In order to obey as much as possible the stationarity criterion the analysis was performed separately for each of the four seasons of the Martian year. A number of intervals without data gaps was selected for both VL1 and VL2 and the first autocorrelation coefficients were derived. They were used to develop autoregressive (AR) models of various orders. The excess pressure time series generated by AR models keep the sequential properties of the measured data quite well. Also, there is reasonably good agreement between the first statistical moments of the measured and synthetic excess pressure data series.

In this section we describe in more detail the sequential characteristics of the short-term time series of the daily averaged pressure. Our first objective is to find appropriate transforms to allow us to obtain statistical equilibrium for the data sets. This is a central feature in the development of time series models. A somewhat restricted form of statistical equilibrium is stationarity, which means that the time series is adequately described by its mean, variance and spectral density (or, equivalently autocorrelation) function. All these three statistical indicators are used here. Our second objective is to obtain forecast models, which are such that the mean square of the deviations between the actual and forecasted values (i.e. the residuals) is as small as possible. A classical technique developed by Box and Jenkins (1970) for time series modeling will be applied (see Boland (2008) for a recent presentation). It uses a combination of autoregressive (AR), integration (I) and moving average (MA) terms in the general auto-regressive integrated moving average (ARIMA) models.

2.3.2.1 Information and Data about the Martian Surface Pressure

The seasonal variation of the surface pressure at a particular point on the planet can differ significantly, under the effect of latitudinal mass redistribution, from that of the mean surface pressure (Hourdin et al. 1993). A meteorological factor was introduced to quantify this effect (Hourdin et al. 1995). The horizontal variation of the meteorological factor can be decomposed into an orographic and a dynamical component. The orographic component accounts for the pressure variation along slopes and varies in time because of the strong seasonal variation of temperature. The dynamic component is due to the geostrophic balance between

the surface pressure and the near surface longitudinal winds (for detailed considerations see Houdin et al. 1995).

A large-amplitude seasonal oscillation in the daily-average atmospheric pressure was observed at both VL sites. There was a deep pressure minimum near the areocentric longitude $L_s=150^\circ$ signaling a maximum CO₂ accumulation onto the south polar cap at the end of southern winter. A secondary, much shallower minimum occurred near $L_s=350^\circ$ signaling a maximum CO₂ accumulation onto the north polar cap at the end of northern winter. The daily-averaged pressure at both sites remained relatively constant from $L_s=340^\circ$ to 360° suggesting that the north polar cap was stable during this period (Santee and Crisp 1995). Note that the aphelion occurs around $L_s=270^\circ$ and 80° , respectively. There appears to be some correlations between polar cap recessions and major dust activity. In the Northern Hemisphere, the Martian atmosphere varies greatly from year to year during the dust storm season while at the other seasons most phenomena are remarkably similar from year to year.

The daily standard deviation of the surface pressure is about 0.1 hPa (during the “clear” days) and 0.3-0.4 hPa (during the dust storm days). The periods with “clear” and “storm” days correspond approximately to the seasons spring-summer and autumn-winter, respectively, when the average daily surface pressure is about 7-8 hPa and 9-10 hPa, respectively. The diurnal cycle and the associated thermal tides (known to be very strong on Mars) are not discussed here. We are just using the diurnally averaged pressure data.

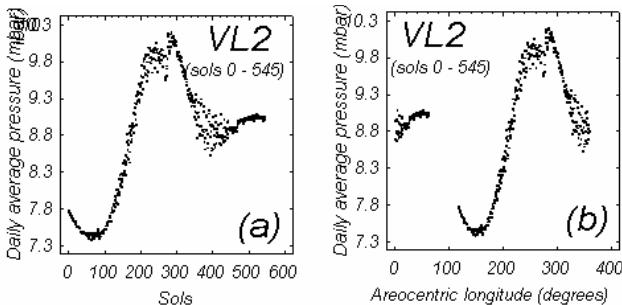


Fig. 2.15 Daily average surface pressure on Mars at VL2 site (sols 0 to 545). (a) Dependence on sol number. (b) Dependence on areocentric longitude

Due to gaps in the measured data, information for 1633 sols for VL1 and 971 sols for VL2 is available only. Figure 2.15 shows a set of daily averaged surface pressure at VL2 site as a function of both sol number and areocentric longitude. The first 545 sols with continuous data are only shown. The available data do not cover completely the first spring on Mars of Viking Lander 2 (Fig. 2.15b).

2.3.2.2 Analysis

An important class of stochastic processes is the *stationary* process. Stationary time series are assumed to be in a specific form of statistical equilibrium, and, in

particular, vary about a fixed mean. Daily average surface pressure time series on Mars behave as though they have no fixed mean (see Fig. 2.15a). Even so, they could exhibit homogeneity in the sense that, apart from local level and trend, one part of the series behaves much like any other. Models, which describe such homogeneous non-stationary behavior, can be obtained by supposing some suitable transform of the process to be stationary.

A number of transforms were tested in order to make the daily averaged pressure data sets stationary at both VL1 and VL2 sites. They include:

(i) A differencing procedure was performed for the pressure (p) data up to order seven. In all cases the mean after differencing becomes constant in time. However, the variance keeps its dependence on the time interval.

(ii) The Viking Lander Meteorology and Atmospheric Opacity Data Set Archive includes values of the daily pressure standard deviation σ_{day} for all sols with available data. A new dimensionless variable was defined as p/σ_{day} . The differencing procedure of (i) was applied to this new variable with similar results.

(iii) Another dimensionless variable was defined as $(p - p_{ave})/\sigma_p$ where p_{ave} and σ_p are the mean and standard deviation of the available pressure data. Differencing this new variable leads to similar results as those mentioned in (i).

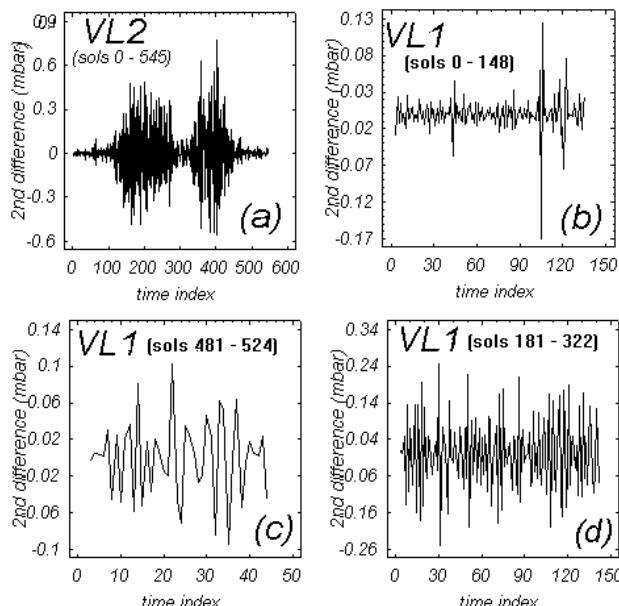


Fig. 2.16 Second difference of daily average surface pressure at VL1 and VL2 sites. **(a)** VL2: sols 0 to 545. **(b)** VL1: sample with small data variance (time interval A1: sols 0-148); **(c)** VL1: sample with medium data variance (interval F1: sols 481-524) **(d)** VL1: sample with large data variance (interval C1:sols 181-323)

From (i) to (iii) above we conclude that stationarity can be obtained for relatively short time intervals only. Figure 2.16 shows the 2nd difference of the daily average pressure at VL2 site for sols 0 to 545. Differencing is not enough to ensure stationarity. However, one can identify shorter intervals with relatively constant data variance. Table 2.9 shows some time intervals with continuous data recording almost satisfying stationarity for VL1 site. They can be classified as intervals with small, medium and large pressure variance.

Table 2.9 Time intervals with quasi-stationary time series of daily averaged surface pressure (2nd difference) on Mars. At VL1 site

Interval Sol name	Areocentric numbers	longitude (degrees)	data variance characteriza- tion	Comments
A1	0-148	97.039-171.046	Small	<i>summer</i> Clear sky + transients
B1	149-180	171.610-189.605	Medium	<i>summer-autumn</i> 1977-A dust storm
C1	181-322	190.202-280.050	Large	<i>autumn-winter</i> 1977-A dust storm beginning of 1977-B dust storm
F1	481-524	10.354-31.093	Medium	<i>spring</i> Clear sky
G1	525-820	31.563-172.970	Small	<i>spring-summer</i> Clear sky+transients

Large pressure variance intervals are usually associated with the dust storm season. Medium and small pressure variance intervals can usually be found during clear sky periods. Exceptions exist, however (Badescu 2001c). Figures 2.16b to 2.16d show examples of time intervals with rather constant daily pressure variance. The excursion of the second difference of the daily average pressure is about 0.05, 0.1 and 0.3 hPa for small, medium and large variance intervals, respectively.

Useful equivalent theoretical tools for describing the behavior of stationary processes are the autocorrelation function and the periodogram. The first was extensively studied in connection with daily average pressure on Mars in Badescu (2001b). The last concept is based on the assumption that a time series is made up of sine and cosine waves with different frequencies (for details see Box and Jenkins (1970, p. 36)). The *sample* spectrum is a sort of periodogram for which the frequency is allowed to vary continuously in the range 0 – 0.5 cycles. We used it here to provide a check on the randomness of a series (or of a series of residuals after fitting a particular model). Note that the highest frequency is 0.5 cycles per unit time interval because the smallest period is 2 intervals.

For convenience we remind that if a time series of N values were truly random with zero mean and variance σ^2 , then the intensities of all frequencies would have expectation $2\sigma^2$ and would be distributed as $\sigma^2 \chi(2)$, independently of all the other components. By contrast, if the series contained a systematic sine component of amplitude A , then its expected value would be $2\sigma^2 + NA^2/2$. *Stationary* time series are characterized by random changes of frequency, amplitude and phase. For this type of series, the sample spectrum normally fluctuates widely and is not capable of any meaningful interpretation. An intuitive explanation of this fact is that the sample spectrum corresponds to such an interval, in the frequency domain, whose width is too small. One can interpret $p(f)df$ [$p(f)$ = the spectrum] as an approximation of the variance of the process in the frequency range f to $f+df$. When the variance of the series is concentrated at high frequencies, this is due to rapid oscillations in the original series. When the variance of the series is concentrated at low frequencies the series is relatively smooth.

Figure 2.17 shows the daily averaged pressure's sample spectrum for two time intervals at VL1 site. The time interval with small daily pressure variance corresponds to a clear sky period (Fig. 2.17a). No systematic behavior is observed. The variance of the time series is concentrated at high frequency. During dust storm intervals the data variance is large and systematic behavior at particular frequencies are obvious (Fig. 2.17b). The 2nd differencing procedure was not enough to eliminate this systematic behavior.

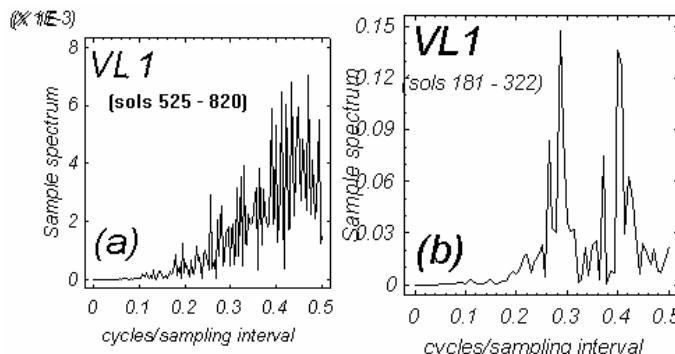


Fig. 2.17 Sample spectrum of the 2nd difference of daily average surface pressure at VL1 site for two time intervals. (a) sample with small data variance (time interval G1: sols 525 to 820); (b) sample with large data variance (time interval C1: sols 181 to 323).

The general auto-regressive integrated moving average model ARIMA (p,d,q) allows us to evaluate the variable z (here identifies the daily average pressure) at the discrete time t (here identifies the sol number) as a function of its values at previous times. Its general form is (Box and Jenkins 1970 p. 11):

$$w_t = \varphi_1 w_{t-1} + \dots + \varphi_p w_{t-p} + a_t + \theta_0 - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (2.34)$$

where the new variable w is obtained by differencing d times the variable z :

$$w_t = \nabla^d z_t . \quad (2.35)$$

Here φ_i ($i=1,p$) are the autocorrelation coefficients, while θ_i ($i=1,q$) are the moving-average coefficients. Also, θ_0 is an adjustment constant and a is a Gaussian white noise of zero mean and standard deviation σ_a . The explicit form of all the ARIMA models used in this section is given below:

$$\text{ARIMA}(3,0,0): \quad z_t = \varphi_1 z_{t-1} + \varphi_2 z_{t-2} + \varphi_3 z_{t-3} + a_t + \theta_0, \quad (2.36-39)$$

$$\text{ARIMA}(0,2,3): \quad z_t = 2z_{t-1} - z_{t-2} + a_t + \theta_0 - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3},$$

$$\begin{aligned} \text{ARIMA}(3,0,3): \quad z_t = & \varphi_1 z_{t-1} + \varphi_2 z_{t-2} + \varphi_3 z_{t-3} + a_t + \theta_0 - \theta_1 a_{t-1} \\ & - \theta_2 a_{t-2} - \theta_3 a_{t-3}, \end{aligned}$$

$$\begin{aligned} \text{ARIMA}(3,2,3): \quad z_t = & (2 + \varphi_1) z_{t-1} + (-1 - 2\varphi_1 + \varphi_2) z_{t-2} \\ & + (\varphi_1 - 2\varphi_2) z_{t-3} + (\varphi_2 - 2\varphi_3) z_{t-4} + \varphi_3 z_{t-5} + a_t + \theta_0 - \theta_1 a_{t-1} \\ & - \theta_2 a_{t-2} - \theta_3 a_{t-3}. \end{aligned}$$

The coefficients φ_i and θ_i as well as the standard deviation of the noise, σ_a , were obtained here by using the maximum likelihood method (Box and Jenkins 1970, their Chap. 7).

The procedure to find the “appropriate” model for a given time series was as follows. First, the level of differencing the data was gradually increased. We tried to obtain a sample spectrum with as small as possible systematic tendencies. Differencing at level higher than $d=2$ did not improve significantly the results. Figure 2.18a shows the cumulative distribution of residuals for the 2nd difference of a daily average pressure time series with no obvious systematic trend.

Second, the significance of the fitted φ_i and θ_i coefficients was estimated. The following criterion was adopted: we tested if the t -statistics greater than 2 in magnitude correspond to p -values less than 0.05. If they were not, we tried to refit the model with the least significant variable excluded. One has to remind that the t -statistic is just the estimated coefficient divided by its own standard error. Thus, it measures “How many standard deviation from zero” the estimated coefficient is, and it is used to test the hypothesis that the true value of the coefficient is non-zero, in order to confirm that the coefficient really belongs in the model. The p -value is the probability of observing a given value (or a value larger than that given value) of the t -statistic under the null hypothesis that the true coefficient value is zero. If the p -value is greater than 0.05 (say) – which occurs roughly when the t -statistic is less than 2 in absolute value – this means that the coefficient may be only “accidentally” significant. To obey the parsimony principle the number of coefficients was kept less than seven.

Every “suitable” model was checked by using a method employing the cumulative spectrum of the residuals. As a series length increases, the residuals become close to the white noise a_t values. Therefore, one expects that study of the residuals could indicate the existence and nature of model inadequacy (Box and Jenkins 1970, their Chap. 8). Generally, a white noise spectrum $p(f)$ has a constant value

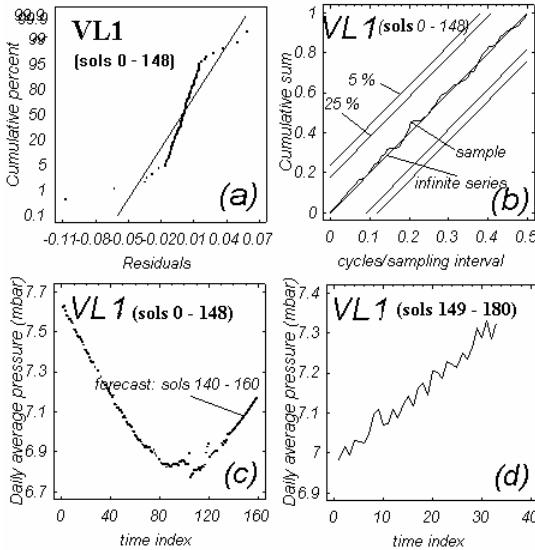


Fig. 2.18 (a) Normal probability plot of residuals for an ARIMA (3,2,3) model associated to time interval A1 (sols 0-148) at VL1 site. (b) Normalized cumulative sample spectrum for residuals of the model in (a). (c) Observed values of daily average surface pressure for sols 0-139 and forecast for sols 140-160. The model in (a) was used. (d) Observed time series for sols 149-180 (time interval B1).

$2\sigma_a$ over the frequency domain $0 - 0.5$ cycles. Consequently, the cumulative spectrum $P(f)$ for the white noise plotted against frequency f is a straight line running from $(0,0)$ to $(0.5, \sigma_a^2)$, that is, $P(f)/\sigma_a^2$ is a straight line running from $(0,0)$ to $(0.5, 1)$. For a white noise series, the plot of the normalized cumulative spectrum against frequency f is scattered about the straight line joining the points $(0,0)$ and $(0.5,1)$. On the other hand, model inadequacy would produce nonrandom values, whose cumulative sample spectrum could show systematic deviation from this line. The Kolmogorov-Smirnov test was used to assess deviation of the sample spectrum from the case in which the residuals were white noise. Using this test, we placed limit lines about the theoretical line. The limit lines are such that if the residual series were white noise, then the cumulative spectrum would deviate from the straight line sufficiently to cross these limits only with the stated probability. Figure 2.18b shows an example of how the normalized cumulative sample spectrum looks in case of a “suitable” model. The probability limits there correspond to 5% and 25 %. These values were used during all calculations.

The ARIMA models found to be “appropriate” are shown in Table 2.10 for all the time intervals together with the fitted coefficients. Generally, a memory effect is always present as practically all models contain autoregressive terms. Shorter time series usually do not need differencing. The time interval B1 (32 sols) is quite short and belongs to a dust storm period. It is characterized by a small white

Table 2.10 ARIMA models at VL1 site associated with the time intervals from Table 2.9

Int. Model Eq.	Φ_1	Φ_2	Φ_3	Θ_0	Θ_1	Θ_2	Θ_3	σ_a
A1 (3,2,3) (2.39)	-0.0894	-0.1395	-0.211	0.00014	1.0713	-0.2430	0.03341	0.0169
B1 (3,0,0) (2.36)	0.4562	0.2289	0.3965	-0.5696	-	-	-	0.0268
C1 (0,2,3) (2.37)	-	-	-	-0.00016	1.3845	-0.1545	-0.2453	0.0559
F1 (3,0,3) (2.38)	0.9549	-0.8910	0.9838	-0.3857	0.4559	-0.8645	0.3547	0.0248
G1 (3,2,3) (2.39)	-0.368	0.3677	0.1742	0.00002	1.1775	0.3090	-0.5353	0.0137

noise influence (i.e. an ARIMA(3,0,0) model looks “appropriate” for this time interval). ARIMA(3,2,3) seems to be the most appropriate model to forecast the daily average surface pressure on Mars. It is a model with seven coefficients. This is in agreement with Badescu (2001b), if one takes into account the equivalence between the AR and MA coefficients. In that work AR models were tested only. The number of terms was large (up to ten), in order to allow a good modeling of the “excess” pressure time series.

Figure 2.18c shows an example of forecasting the daily average surface pressure at VL1 site for sols 140 to 160. The exhibited trend is in good agreement with observed values and tendency for sols 149 to 180 (see Fig. 2.18d). More details can be found in Badescu (2001c).

2.3.3 Forecasting Hourly Mean of Meteorological Parameters

Predicting the values of meteorological parameters on short time periods (hours, say) is of considerable interest for future manned missions on Mars. Figure 2.19a shows an example of the time variation of the near-surface atmospheric temperature at VL1 site during time period C1 (see Table 2.9). The data refer to autumn/winter conditions during a dust storm. One can easily see the daily cyclical pattern of temperature. However, this pattern is rather strongly affected at the end of the interval, which coincides with the beginning of the 1977-B dust storm.

The cyclical pattern of meteorological parameters recommends the usage of seasonal ARIMA $(p,d,q)(P,D,Q)_s$ models for forecasting purposes (Box and Jenkins 1970). The general form of these models is:

$$\Phi_p(B_1)\phi_P(B_s)(\nabla_1)^d(\nabla_s)^D z_t = \Theta_q(B_1)\Theta_Q(B_s)a_t, \quad (2.40)$$

where z_t and a_t have the same meaning as in Eq. (2.34). The difference operator ∇_s and the backward shift operator B_s are defined as

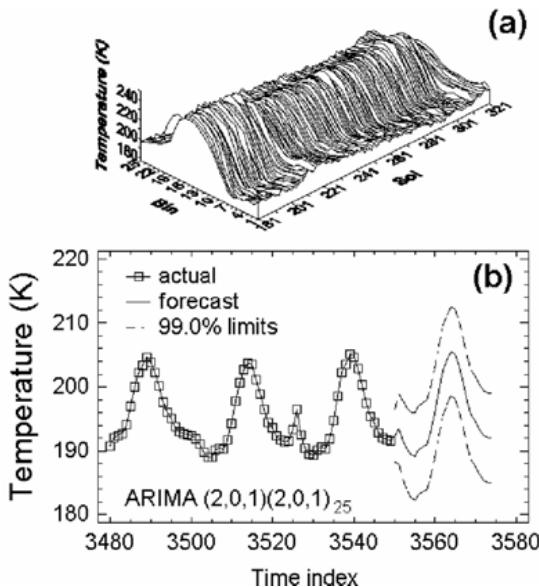
$$\nabla_s z_t = z_t - z_{t-s}, \quad B_s z_t = z_{t-s}, \quad (2.41,42)$$

while the polynomials Φ_p, Φ_P, Θ_q and Θ_Q are generally defined by

$$Y_\alpha(X) = 1 - c_{Y,1}X - c_{Y,2}X^2 - \dots - c_{Y,\alpha}X^\alpha, \quad (2.43)$$

where Y stands for Φ_p, Φ_P, Θ_q and Θ_Q ; α stands for p, P, q and Q ; X stands for B_1 and B_s while the coefficients $c_{Y,i}$ ($i = 1, \alpha$) are obtained by fitting the models Eq. (2.40) to the data.

As an example, we give here the ARIMA models obtained by using data from interval C1. In the case of ambient temperature (in Kelvin) the following $(2,0,1)(2,0,1)_{25}$ model was obtained: $c_{p,1} = 1.7686$, $c_{p,2} = 0.7694$, $c_{q,1} = 0.9213$, $c_{p,1} = 0.9270$, $c_{p,2} = 0.0614$ and $c_{Q,1} = 0.5865$ while the standard deviation of the residuals is $\sigma = 1.6842\text{K}$. For the atmospheric pressure (in hPa) the following $(2,0,1)(2,0,1)_{25}$ model was obtained: $c_{p,1} = 1.4651$, $c_{p,2} = -0.4983$, $c_{q,1} = 0.6367$, $c_{p,1} = 0.9051$, $c_{p,2} = 0.0092$ and $c_{Q,1} = 0.3421$ while the standard deviation of the residuals is $\sigma = 0.03661\text{ hPa}$. Finally, the mean zonal wind speed (in m/s) could be predicted by using the following $(2,0,1)(2,0,1)_{25}$ model: $c_{p,1} = 1.1389$, $c_{p,2} = -0.2627$, $c_{q,1} = 0.1827$, $c_{p,1} = 0.9880$, $c_{p,2} = 0.00001$ and $c_{Q,1} = 0.8241$ while the standard deviation of the residuals is $\sigma = 1.1930\text{ m/s}$.



2.4 Conclusions

The amount of incident solar energy depends on the tilt and orientation of the receiving surface. Due to the apparent motion of the Sun the receiving surface must continuously change its orientation if maximum direct (or global) solar radiation has to be collected. The best results are obtained when a tracking mechanism with two rotation axes is used. Other simpler technical solutions consist of using a mechanism based on a single rotation axis. In this case both continuous and discrete adjustment of the surface tilt angle can be envisaged while the orientation of the receiving surface should always be towards the equator of the planet (i.e. the surface is South oriented in the Northern Hemisphere and North oriented in the Southern Hemisphere). All the above ways of collecting solar radiation were analyzed in this chapter in the case of a receiving surface placed on the surface of Mars. For this aim we developed a procedure to compute solar direct, diffuse and global radiation on surfaces of arbitrary tilt and orientation. The core of the model consists of the results shown in Appelbaum and Flood (1990), Appelbaum et al. (1993) concerning solar direct, diffuse and global radiation on a horizontal surface. These results were derived by using measurements performed by the Viking lander's equipment. A simple model was used to convert the horizontal data to data corresponding to surfaces of arbitrary tilt and orientation.

The main outputs of the model are the direct, diffuse and global solar irradiances. Generally, they are functions of day, hour and tilt and azimuth angles of the receiving surface. The last two parameters can be optimized so that the global irradiance be maximum. Also, the model evaluates direct, diffuse and global irradiations on different time intervals, such as day, season or whole year. Any of these irradiations is maximum for a certain optimum tilt angle.

The recommended strategies to collect the maximum amount of solar radiation depend on the dust content in the atmosphere. Thus, during the "clear" days (i.e. when the atmospheric optical depth is around 0.5) the best strategy is to keep the receiving surface perpendicular to the Sun's rays (P-strategy). This requires a tracking mechanism with two rotation axes. The P - strategy allows the collection of the maximum *direct* solar radiation and it is recommended in the case of solar concentrators. In the Northern Hemisphere a simpler procedure (called InOp here) is to use a South oriented surface whose tilt angle is continuously adjusted to collect the maximum *global* irradiance. During the dust storms the diffuse component exceeds the direct component of solar radiation. As a consequence, the dependence of the incident irradiance upon tilt and orientation becomes less important and even simpler strategies are more convenient. Examples are the H - strategy (use of horizontal receiving surfaces) and the DOp - strategy (the surface tilt angle is adjusted once a day).

One of the objectives of this chapter was to derive regression relations between the optical depth and common near-surface meteorological parameters as pressure and temperature. These relations could provide the entries for the solar energy computing models used during the simulation of various devices operating on Mars, as solar cells or solar thermal power plants. The accuracy of all types of regressions is lower during the dust storm season while during "clear" sky periods

the accuracy is usually good. Atmospheric pressure is more recommended than ambient temperature as a regression parameter. The influence of the areocentric longitude and local solar time on the regression's accuracy is important during the dust storm season mainly. Generally, the regression accuracy is very good at low atmospheric pressure (i.e. on clear sky). The regression accuracy is better at VL2 site than at VL1 site, especially at higher atmospheric pressure. This is a consequence of the higher optical depth at VL1 site during the dust storm season.

The optical depth is well correlated to the daily variation of the surface pressure Δp and to the standard deviation of the daily surface pressure $\sigma_{p,day}$. The following simple linear relations apply to both VL1 and VL2 sites (here both Δp and $\sigma_{p,day}$ enter in hPa):

$$\tau = 0.2342 + 2.237\Delta p, \quad \tau = 0.1851 + 8.1295\sigma_{p,day}. \quad (2.44,45)$$

Use of these correlations for other places of similar latitudes should be made, however, with caution because there are potentially significant longitudinal variations of local opacity and regional meteorological fields. Also, the optical depth varies seasonally. The largest changes are due to non-monotonous increases and declines associated with dust storm onset. The relationships could be different in years without major dust storms. Larger Martian data sets should be measured by future missions in order to increase the reliability of the estimates.

A second objective of this review was to develop time series models for weather forecasting on Mars. In this context, a central feature is to find appropriate transforms allowing to obtain stationarity of the data. This was performed here for the daily average surface pressure time series on Mars using the mean, the variance and the data spectrum as statistical indicators. Preliminary tests proved that stationarity can be obtained for relatively short time intervals only. These time intervals were classified as intervals with small, medium and large pressure variance. Large pressure variance intervals are usually associated with the dust storm season. Medium and small pressure variance intervals can be found usually during clear sky periods. The excursion of the 2nd difference of the daily average pressure is about 0.05, 0.1 and 0.3 hPa for small, medium and large variance intervals, respectively. The variance of the clear sky daily pressure time series is concentrated at high frequency showing rapid oscillations in the original series. No systematic behavior is observed. During dust storm intervals the data variance is large and systematic behavior at particular frequencies is sometimes seen.

General auto-regressive integrated moving average (ARIMA) models were used to obtain forecast models for daily average pressure on Mars. A procedure to find the "appropriate" model for a given time series was developed. It was based on differencing the original time series and a *t*-statistic assisted estimation of the significance of the autoregressive and moving-average coefficients. Every selected model was checked by using a method employing the cumulative spectrum of the residuals. ARIMA(3,2,3) seems to be the most appropriate model to forecast the daily average surface pressure on Mars. However, exceptions exist.

Seasonal ARIMA models were used to forecast various weather parameters (such as atmospheric pressure and temperature and wind speed) on short time intervals (less than one day). Few results were deployed here.

Finally, a few comments will give perspective to our results. The present techniques could be used as candidates for preparing databases of the atmosphere of Mars. In this respect, Viking sites could provide good test cases. Another approach would be to simply devise a method for interpolation of existing measured series. High-order ARIMA models could also be used in “forecast” mode taking into account the few observations before a data gap occurs. If the objective is to provide data for testing instruments and machines, the analysis should be performed separately under the two major Martian weather situations (i.e. clear sky and dust storm). Using one or another of these approaches depends, of course, on the objectives and the user.

References

- Andersen, P.: Comment on calculation of monthly average insolation on tilted surfaces. *Solar Energy* 25, 287 (1980)
- Appelbaum, J., Flood, D.J.: Solar radiation on Mars. *Solar Energy* 45, 353–363 (1990)
- Appelbaum, J., Landis, G.A., Sherman, I.: Solar radiation on Mars - Update 1991. *Solar Energy* 50, 35–51 (1993)
- Badescu, V.: Dynamic solar space power system. *Space Technol.* 14, 331–337 (1994)
- Badescu, V.: Optimization of a solar space power system based on thermodynamic cycles. *Int. J. Solar Energy* 16, 263–275 (1995)
- Badescu, V.: Optimum design and operation of a dynamic solar space power system. *Energy Conv. Mgmt.* 37, 151–160 (1996)
- Badescu, V.: Use of autoregressive models to generate a series of daily averaged point cloudiness values. *Renewable Energy* 12, 71–82 (1997)
- Badescu, V.: Simulation of solar cells utilization on the surface of Mars. *Acta Astronautica* 43, 443–453 (1998)
- Badescu, V.: Inference of atmospherical optical depth from near-surface meteorological parameters on Mars. *Renewable Energy* 24, 45–57 (2001a)
- Badescu, V.: Synthesis of short-term time series of daily averaged surface pressure on Mars. *Environmental Modelling & Software* 16, 283–295 (2001b)
- Badescu, V.: Time-series analysis and forecasting of daily averaged surface pressure on Mars (VL1 and VL2 sites). *Meteorol. Atmos. Phys.* 78, 195–204 (2001c)
- Badescu, V. (ed.): *Modeling Solar Radiation at the Earth’s Surface*, 517 p. Springer, Berlin (2008)
- Badescu, V., Dinu, C.: Solar energy flux and maximum efficiency of solar thermodynamic power generation in our planetary system. *Space Power* 12, 151–164 (1993)
- Badescu, V., Dinu, C.: Maximum performance of omnicolor photothermal and photovoltaic converters in our planetary system. *Renewable Energy* 6, 765–777 (1995)
- Badescu, V., Dinu, C.: Optimization of multicolor photothermal power plants in the solar system: a finite-time thermodynamic approach. *J. Phys. III France* 6, 143–163 (1996)
- Badescu, V., Popescu, G., Feidt, M.: Design and optimisation of a combination solar collector-thermal engine operating on Mars. *Renewable Energy* 21, 1–22 (2000)
- Barnes, J.R.: Time spectral analysis of the midlatitude disturbances in the Martian atmosphere. *J. Atmos. Sci.* 37, 2002–2015 (1980)
- Barnes, J.R.: Midlatitude disturbances in the Martian atmosphere: A second Mars year. *J. Atmos. Sci.* 38, 225–234 (1981)

- Beish, J.D., Parker, D.C.: Meteorological survey of Mars. *J. Geophys. Res.* 95(B9), 14657–14675 (1990)
- Boland, J.: Time series modeling of solar radiation. In: Badescu, V. (ed.) *Modeling solar radiation at the Earth's Surface. Recent Advance*, pp. 283–312. Springer, Berlin (2008)
- Box, G., Jenkins, G.: *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco (1970)
- Chowdhury, H., Rahman, S.: Comparative assessment of plane-of-array irradiance models. *Solar Energy* 39, 391–398 (1987)
- Clancy, R.T., Lee, S.W.: A new look at dust clouds in the Martian atmosphere. Analysis of emission-phase-function sequences from global Viking IRTM observations. *Icarus* 93, 135–158 (1991)
- Colburn, D.S., Pollack, J.B., Haberle, R.M.: Diurnal variations in optical depth at Mars. *Icarus* 79, 159–189 (1989)
- Crisp, D.: Infrared radiative transfer in the dust-free Martian atmosphere. *J. Geophys. Res.* 95(B9), 14577–14588 (1990)
- Festa, R., Ratto, C.F.: Solar radiation statistical properties. Report No IEA-SCHP=9E-4, IEA Task 9, Solar radiation and pyranometer studies, University of Genova (1993)
- Hourdin, F.: A new representation of the absorption by the CO₂15-μm band for a Martian global circulation model. *J. Geophys. Res.* 97(E11), 18319–18335 (1992)
- Hourdin, F.P., Van Le, F., Forget, F., Talagrand, O.: Meteorological variability and the annual surface pressure cycle on Mars. *J. Atmos. Sci.* 50, 3625–3640 (1993)
- Hourdin, F., Forget, F., Talagrand, O.: The sensitivity of the Martian surface pressure and atmospheric mass budget to various parameters: A comparison between numerical simulations and Viking observations. *J. Geophys. Res.* 100, 5501–5523 (1995)
- Iqbal, M.: The influence of collector azimuth on solar heating of residential buildings and the effect of anisotropic sky-diffuse radiation. *Solar Energy* 26, 249–257 (1981)
- Jones, R.E.: Effects of overhang shading of windows having arbitrary azimuth. *Solar Energy* 24, 305–312 (1980)
- Kahn, R.: Temperature measurements of a Martian local dust storm. *J. Geophys. Res.* 100(E3), 5265–5275 (1995)
- Kambezidis, H.E.: Estimation of sunrise and sunset hours for locations on flat and complex terrain: review and advancement. *Renewable Energy* 11, 485–494 (1997)
- Lee, S.W.: *Viking Lander Meteorology and Atmospheric Opacity Data Set Archive*, Volume VL-1001, Laboratory for Atmospheric and Space Physics, Campus Box 392, University of Colorado, CO 80309-0392 (1995)
- Leovy, C.B., Tillman, J.E., Guest, W.R., Barnes, J.: Interannual variability of Martian weather. In: Hunt, G. (ed.) *Recent Advances in Planetary Meteorology*, pp. 69–84. Cambridge University Press, New York (1985)
- Lindner, B.L.: The Martian, polar cap: radiative effects of ozone, clouds and airborne dust. *J. Geophys. Res.* 95(B9), 1367–1379 (1990)
- Martin, L.J., Zurek, R.W.: An analysis of the history of dust activity on Mars. *J. Geophys. Res.* 98(E2), 3221–3246 (1993)
- Nijegorodov, N., Devon, K.R.S., Jain, P.K., Carlsson, S.: Atmospheric transmittance models and an analytical method to predict the optimum slope of an absorber plate, variously oriented at any latitude. *Renewable Energy* 4, 529–543 (1994)
- Paige, D.A., Ingersoll, A.P.: Annual heat balance of Martian polar caps: Viking observations. *Science* 228, 1160–1168 (1985)
- Paige, D.A., Wood, K.D.: Modeling the Martian CO² cycle, 2, Interannual variability. *Icarus* 99, 15–27 (1992)

- Pal, A., Ureche, V.: Astronomie (in Romanian). Editura Didactica si Pedagogica, Bucharest, Appendix 5 (1983)
- Pollack, J.B., Colburn, D., Flasar, F., Kahn, R., Carlson, C.D., Pidek, D.: Properties and effects of dust particles suspended in the Martian atmosphere. *J. Geophys. Res.* 84, 2929–2945 (1979)
- Pollack, J.B., Haberle, R.M., Murphy, J.R., Shaeffer, J., Lee, H.: Simulation of the general circulation of the Martian atmosphere, 2, seasonal pressure variations. *J. Geophys. Res.* 98(E2), 3149–3181 (1993)
- Santee, M., Crisp, D.: Thermal structure and dust loading of the Martian atmosphere during late Southern summer: Mariner 9 revisited. *J. Geophys. Res.* 98(E2), 3261–3279 (1993)
- Santee, M.L., Crisp, D.: Diagnostic calculations of the circulation in the Martian atmosphere. *J. Geophys. Res.* 100(E3), 5465–5484 (1995)
- Tillman, J.E.: Martian meteorology and dust storms from Viking observations. In: McKay, C.P. (ed.) *The case for Mars II. Science and Technology Series*, vol. 62. American Astronautical Society, Univelt (1985)
- Tillman, J.E.: Mars global atmospheric oscillations: Annually synchronized transient normal-mode oscillations and triggering of global dust storms. *J. Geophys. Res.* 93(D8), 9433–9451 (1988)
- Van Hemelrijck, E.: The influence of global dust storms on the mean seasonal daily insolation at the Martian surface. *Earth, Moon and Planets* 33, 157–162 (1985)
- Van Hemelrijck, E.: The oblateness effect on the mean seasonal daily insolation at the Martian surface during global dust storms. *Earth, Moon and Planets* 38, 209–216 (1987)
- Wood, S.E., Paige, D.A.: Modeling the Martian seasonal CO₂ cycle: Fitting the Viking lander pressure curves. *Icarus* 99, 1–14 (1992)
- Zurek, R.W., Leovy, C.B.: Thermal tides in the dusty Martian atmosphere: A verification of theory. *Science* 213, 437–439 (1981)
- Zurek, R.W., Martin, L.J.: Interannual variability of planet-encircling dust storms on Mars. *J. Geophys. Res.* 98(E2), 3247–3259 (1993)

Chapter 3

Weather Influence on PV Solar Cells Operation on Mars

Viorel Badescu

Polytechnic University of Bucharest, Romania

3.1 Introduction

Among the different solar space power systems to be used on Mars a leading position is played by the photovoltaic (shortly, PV) solar cells. Due to their relatively low cost power, high reliability and the lack of moving parts they powered the space program from the very beginning (Landis and Appelbaum 1991). As a result the array manufacturing technology is now well developed, and the technology is well characterized for vibration, thermal-cycling, and other environmental loads of the space environment. A number of studies concerning the usage of PV cells on the Mars surface have been performed during the years and we quote here some of them. A rover powered by a PV array was designed in Hibbs (1989). To add robustness, a fixed array rather than one which tracks the sun was envisaged, because of the modest power requirements (275 Watt average power and peak power capability of up to 2000 W during climbing over large boulders). The array surface was as large as 7 m^2 . Another study refers to a long-endurance, remotely piloted aircraft capable of flight within the Martian environment (Colozza 1990). There is a variety of mission scenarios which would be possible with this aircraft ranging from magnetic and gravity field mapping to surveillance / reconnaissance missions. The flight duration would be on the order of one year. The power and propulsion systems consist of solar PV array panels, a regenerative fuel cell and an electric motor. Depending on the solar cell type the required PV array surface could be quite large (from 336 to 405 m^2 in case of silicon-based solar cells or from 118 to 166 m^2 when gallium arsenide PV arrays are used). A solar power system for a 40 day manned Mars surface scientific expedition was studied in McKissick et al. (1990). The electrical power requirements were assumed to be 40 kW for life support and experiment power during the Martian day and 20 kW for life support during the Martian night. The solar energy system consisted of roll out amorphous silicon arrays and a hydrogen - oxygen regenerative fuel cell energy storage system. The total array covers 2822 m^2 when deployed and has a blanket mass of 177.6 kg.

Early works were performed to study the influence of tilt and orientation on the power supplied by a PV array. So, in Colozza (1991) a tent-shaped array was designed and optimized for both Lunar and Mars applications. A tent shape was chosen because it can produce a nearly constant power profile throughout the day time period, if the proper tent angle is selected. It was concluded that a tent angle of 60 degrees will produce the lowest amount of output power variation. A single axis tracking PV array able to operate on Moon or Mars was analyzed in Colozza (1992). Detailed studies dealt with both stationary PV cells (Appelbaum and Landis 1993; Appelbaum et al. 1995) and tracking PV arrays (Appelbaum and Landis 1993; Appelbaum et al. 1996). As a result of the growing knowledge and experience in the field, PV cells were used in the past decade by NASA to run unmanned missions needed to investigate the geology and meteorology of Mars, starting with the Pathfinder's Sojourner Rover in 1997 and continuing with Spirit and Opportunity Rovers in 2004 and Phoenix Lander in 2008.

In this chapter we report simulation results for the operation of different sorts of solar cells on Mars surface. Measurements performed by the Viking Landers are used during the simulation. They refer to ambient temperature, wind speed and the optical depth of the Martian atmosphere. [At the moment of this writing (January 2009) important experience has been accumulated by NASA during the five year operation of Spirit and Opportunity Rovers, but the data is not yet fully publicly available]. The incoming global solar radiation flux (direct, diffuse and ground reflected) is computed by using the model we described in Chap. 2 (see also Badescu (1998)). A thermodynamic model of solar cell was developed in Badescu and Landsberg (1995). The same simple model is used here with a number of changes allowing to simulate the incoming *diluted* radiation fluxes. Simulations were performed for both Viking Landers (VL) 1 and 2 sites during all the four Martian seasons. Note that both Viking Landers were powered by nuclear sources (see Chap. 1). Solar cells based on silicon and gallium arsenide are envisaged under the assumption of both ideal and present state-of-the-art performance. Information about the optimum voltage and cell temperature is provided as well as details about the electric power supply and solar cell energy conversion efficiency. Different sorts of stationary and tracking solar cells arrays are studied.

The results we report are upper bound for real solar cell performances. They are, however, useful in comparing various operation regimes and weather conditions.

3.2 Model of Solar Cell

An empirical model of solar cell operation on Mars was proposed in Appelbaum and Flood (1989). The main advantage of that model is that experiments made on Earth (but simulating the Martian conditions) allow to derive the appropriate values of the adjusting coefficients. Here we use the second thermodynamic model for photovoltaic energy conversion developed in Badescu and Landsberg (1995) and Badescu et al. (1996). This is a flexible model which does not require empirical adjusting coefficients. One starts by assuming that the PV cell receives photon *number* fluxes of direct, diffuse and ground reflected solar radiation and radiation

emitted by the atmosphere. The PV cells provide electrical power and emit radiation because their temperature grows up as a result of photon absorption. The model optimizes both the PV cell voltage and temperature. This model is adapted here to be used under the Martian surface conditions.

The electrical power flux P_{el} provided by a (band-gap material) solar cell, under the assumption that the flow of charge in both the conduction and valence bands is a reversible process, is given by Eq. (3.4) of Badescu et al. (1996):

$$P_{el} = eVA \left(\gamma_{sc} + \gamma_{ac} - \frac{1}{\eta_{rec}} \gamma_{ca} \right), \quad (3.1)$$

where e is the electron charge, V is the voltage across the cell, A is the area of the solar cell, γ_{sc} , γ_{ac} and γ_{ca} are net photon number fluxes on the cell surface while η_{rec} is the radiative recombination efficiency. Thus γ_{ca}/η_{rec} is the total interband recombination rate, which increases by decreasing η_{rec} . Subscripts sc, ac and ca denote the following transfer directions: Sun - cell, ambient - cell and cell - ambient, respectively. In deriving Eq. (3.1) a perfect photon absorption was assumed while the photon recycling effect and the net trap ionization rate were neglected.

The solar cells envisaged in this chapter are based on silicon (band-gap $E_g=1.12$ eV) and gallium arsenide (band-gap 1.42 eV). Both materials have the peak spectral response wavelength well below 15 μm (Appelbaum et al. 1995) and, consequently, are practically transparent for the radiation emitted by the carbon dioxide from the Mars atmosphere. Then, one could assume $\gamma_{ac}=0$ and Eq. (3.1) becomes:

$$P_{el} = eVA \left(\gamma_{sc} - \frac{1}{\eta_{rec}} \gamma_{ca} \right). \quad (3.2)$$

At the top of the Mars atmosphere the solar radiation flux can be very well modeled as black-body radiation at the temperature of the Sun's photosphere (say $T_s \approx 5760$ K). However, because of the absorption and scattering in the atmosphere the PV cells receive fluxes of *diluted* radiation. Consequently, a model of diluted black-body radiation was used here (Landsberg and Tonge 1980; Badescu 1991).

Generally, a photon number flux of diluted radiation γ incident on the PV cell surface is given by $\gamma = \varepsilon \gamma_{bb}(T_s)$ where $\gamma_{bb}(T_s)$ is the photon *number* flux of black-body of temperature T_s at the level of Mars orbit while ε is the wavelength independent dilution factor. ε was evaluated by using the relationship $\varepsilon = \varphi_{optical}(T_s)/\varphi_{bb}(T_s)$ where $\varphi_{optical}(T_s)$ and $\varphi_{bb}(T_s)$ are photon *energy* fluxes, computed, in the first case, by using the atmospheric optical depth value as input data (see Chap. 2) and, in the second case, under the assumption of black-body radiation. The above model was applied for all the three components of the

incident solar radiation (i.e. direct (dir), diffuse (dif) and ground reflected (refl)). To be explicit, the three dilution factors are given by:

$$\epsilon_{dir} = \frac{\varphi_{optical,dir}}{\varphi_{bb}(T_s)}, \quad \epsilon_{dif} = \frac{\varphi_{optical,dif}}{\varphi_{bb}(T_s)}, \quad \epsilon_{refl} = \frac{\varphi_{optical,refl}}{\varphi_{bb}(T_s)}, \quad (3.3)$$

where the energy flux of direct (beam) solar radiation at the top of Mars atmosphere is (Badescu and Landsberg 1995):

$$\varphi_{bb}(T_s) = \frac{B_{sc}}{\pi} \sigma T_s^4. \quad (3.4)$$

Here σ is Stefan-Boltzmann constant while B_{sc} is a geometric factor given in case of symmetrical radiation sources by:

$$B_{sc} = \pi \sin^2 \theta \quad (3.5)$$

where θ is the half-angle subtended by the Sun when viewed from Mars (Landsberg and Baruch 1986). One easily finds that $\sin \theta = R_s/r_{SM}$ where R_s (= 695500 km) and r_{SM} are Sun radius and the distance from Sun to Mars, respectively. The last quantity is a function of the areocentric longitude through $r_{SM} = < r_{SM} > / (1 + e_{SM} \cos L_s)$ (see e.g. Hourdin et al. (1995)) where $< r_{SM} >$ (= 1.5236915 AU) is the mean distance Sun-Mars while e_{SM} (= 0.093377) is the eccentricity of Mars orbit. During computations the areocentric longitude values from the file VL_OPAC.DAT were used (see Chap. 2).

Now, the photon number flux γ_{sc} is given by:

$$\gamma_{sc} = \sum_{i=dir,dif,refl} \epsilon_i (1 - \rho_{s,i}) \gamma_{bb}(T_s), \quad (3.6)$$

where $\rho_{s,i}$ ($i = dir, dif, refl$) refer to the reflectance for the three components of solar global radiation entering the cell. $\gamma_{bb}(T_s)$ and γ_{ca} are given by Eqs. (3.7a) and (3.7c) of Badescu et al. (1996), i.e.:

$$\begin{aligned} \gamma_{bb}(T_s) &= \frac{15}{\pi^5} \frac{\sigma}{k} B_{sc} T_s^3 F_2(x_s, 0), \\ \gamma_{ca}(T_s) &= \frac{15}{\pi^5} \frac{\sigma}{k} (1 - \rho_c) B_T T_c^3 F_2\left(\frac{x_s}{b}, \frac{v_s}{b}\right). \end{aligned} \quad (3.7,8)$$

Here k is Boltzmann's constant, ρ_c refers to the reflectance for the radiation emitted by the cell itself upon leaving the material while T_c is cell temperature. B_T is a geometric factors which refer to the radiation emitted by the cell. Here

solar cells emitting on a hemispherical solid angle are considered and $B_T = \pi$. The following dimensionless factors were used in Eqs. (3.7) and (3.8):

$$b \equiv \frac{T_c}{T_s}, \quad x_s \equiv \frac{E_g}{kT_s}, \quad v_s \equiv \frac{eV}{kT_s}. \quad (3.9)$$

The function F_2 entering Eqs. (3.7) and (3.8) is defined by:

$$F_2(u, v) \equiv \int_u^\infty \frac{x^2 dx}{e^{x-v} - 1}. \quad (3.10)$$

The radiation emitted by the cell is assumed to be unpolarized ($P_c = 0$). Actually, most efficient solar cells use a variety of techniques to improve the optical efficiency: antireflection coatings, textured surface treatment, increase of back-surface reflection, etc. They lead to reflectance value as small as 2-4 % (Hamakawa 1992; Green et al. 2008). Here we assumed the PV cell reflectance to be 0.05 for all kind of incident radiation ($\rho_{s,i} = 0.05; i = \text{dir, dif, refl}$).

The usual silicon and gallium arsenide PV cells have a recombination efficiency η_{rec} around 10^{-4} and $1.6 \cdot 10^{-4}$, respectively (Badescu et al. 1996). Here we assume $\eta_{rec} = 10^{-4}$ in both cases. However, the ideal case $\eta_{rec} = 1$ will be considered, too, to give perspective to the results. One denotes by φ_{total} the total solar energy flux incident on the PV cell:

$$\varphi_{total} = \varphi_{optical, dir} + \varphi_{optical, dif} + \varphi_{optical, refl}. \quad (3.11)$$

The flux of thermal energy \dot{Q} generated by the solar cell is given by Eq. (2.4) of (Badescu et al. 1997):

$$\dot{Q} = \varphi_{total} - P_{el}. \quad (3.12)$$

In practice the thermal flux \dot{Q} is lost to the Martian atmosphere by a convective mechanism. The following Newton-like relationship is used here:

$$\dot{Q} = UA(T_c - T_a), \quad (3.13)$$

where the convection heat loss coefficient U is mainly controlled by the wind speed value w . The pressure of the Martian atmosphere was measured at both VL1 and VL2 sites during the whole year. Usually, it is as low as 6 to 9 hPa. Under these conditions the atmosphere can be modeled as an ideal gas (see e.g. Sante and Crisp (1995)). This enables us to use for the convection heat losses coefficient U (in W/(m² K)) the following relationship specific to the flow of an ideal gas along a flat plate (Danescu et al. 1985):

$$U = 6.16 + 4.18w. \quad (3.14)$$

Here the unit of the wind speed w is m/s. Equation (3.14) is used whatever the tilt and orientation of the solar cell array is. More accurate models of evaluating the convection heat transfer U exist (see Chap. 5) but they are not needed at this stage.

The electrical power P_{el} supplied by the PV array can be maximized in respect to the voltage V and cell temperature T_c . The maximum electric power ($P_{el,max}$) is reached at the optimum voltage and cell temperature obtained by solving the equations $\partial P_{el} / \partial V = 0$ and $\partial P_{el} / \partial T_c = 0$ under the constraint Eq. (3.13). Then, the maximum PV cell efficiency ($\eta_{cell,max}$) is simply given by $\eta_{cell,max} = P_{el,max} / \varphi_{total}$.

All the following optimizations were performed by using a quasi-Newton algorithm with line search (Kahaner et al. 1989). At any iteration, the objective function (e.g. the solar energy conversion efficiency) is approximated by a quadratic function derived from a Taylor series. The quadratic function is then maximized to obtain a search direction, and an approximate maximum of the objective function along the search direction is found using a line search. Note that the above procedure determines sometimes approximate maxima when the objective function is relatively complicated, as is the case here.

3.3 Performance of Solar Cells on Mars Surface

Different sorts of solar cells for space applications are in use or in stage of laboratory research. They include silicon thin-film or single crystal cells, GaAs-based cells in both V-grooved and point contact technology and for both unconcentrated and concentrated solar radiation, as well as InP cells, among others (Flood 1989; Appelbaum and Landis, 1993; Reddy 2003). For the near future use on Mars it is likely that silicon and GaAs -based solar cells will be taken into consideration. Also, concentration systems seem to be of smaller interest, as direct solar irradiance is very small or even vanishes during dust storm periods (see Chap. 2).

A number of studies analyzed the optimum tilt and orientation of both stationary collectors and tracking photovoltaic arrays operating on Mars surface (Appelbaum and Landis 1993; Appelbaum et al. 1995, 1996; Badescu 1998). In case of the *stationary* collectors operating on "clear" skies the daily optimum tilt angle of the receiving surface was computed in Appelbaum and Landis (1993) and Appelbaum et al. (1995) by minimizing the incidence angle of the *direct* (beam) solar radiation. When "cloudy" skies are considered, the optimum tilt angle was obtained by maximizing the incoming *global* solar irradiation. In order to obtain analytical relationships for the optimum tilt angle, the sunrise and sunset hour angles on the tilted surface were approximated by the corresponding values for a horizontal receiver (Appelbaum et al. 1995).

The optimum tilt angle for *tracking* collectors was obtained in Appelbaum et al. (1996) by minimizing the incidence angle of the *direct* radiation. A variety of single and two axis tracking collectors were studied in the quoted paper. Different strategies to collect the maximum amount of *global* solar radiation on Mars surface were analyzed in Badescu (1998). As no approximation was made for the

sunrise and sunset angles, the procedure of computing the optimum tilt angle was numerical. The results obtained in Badescu (1998) are in good concordance with those from Appelbaum et al. (1995, 1996). As an example, the yearly optimum tilt angle computed in Badescu (1998) for a South oriented stationary collector placed at the Viking Landers 1 and 2 sites is 5 degrees and 23 degrees, respectively, as compared to 6.5 degrees and 22 degrees, respectively, which are the values reported in Appelbaum et al. (1995). In this chapter we report results obtained by using the procedures proposed in Badescu (1998) to optimize the tilt of the PV arrays so that the incoming flux of solar global radiation be a maximum. The following strategies to collect solar radiation are envisaged here:

- (H) the receiving surface is horizontal;
- (InOp) the receiving surface is South oriented and the tilt angle is permanently adjusted as the incident flux of global solar radiation be a maximum and
- (P) the tilt and orientation are continuously adjusted so that the receiving surface be perpendicular on Sun's rays.

The first strategy refers to a stationary collector while the last two are associated to tracking PV arrays.

The simulation of solar cell operation was performed for all available *complete* records of meteorological and actinometric data (see Table 2.2 in Chap. 2). Two performance indicators and two operating parameters are reported here. They are the maximum PV conversion efficiency and electrical power supplied per unit cell area ($\eta_{cell,max}$ and $P_{el,max}$, respectively) and the optimum cell voltage and temperature (V_{opt} and $T_{c,opt}$, respectively). Among the two performance indicators, $P_{el,max}$ is strongly dependent on the level of the incoming solar energy flux (or, in other words, on the local solar time). $\eta_{cell,max}$ is less dependent on the solar radiation flux.

3.3.1 VL1 Site

Figure 3.1 shows the simulation results obtained for the VL1 site by using all the available data during autumn, year 1. Numbering the sols (i.e. Martian days) started when the VL1 touch down the soil. The case of PV array perpendicular on the Sun's rays was considered (strategy P). Autumn is a quite "cloudy" season, with most records associated to an optical depth value between 1 and 2. There is no obvious dependence of the PV conversion efficiency on the optical depth. When the case of ideal solar cells is considered ($\eta_{rec} = 1$) both the silicon and gallium arsenide solar cells have generally a similar performance, with the silicon cells slightly better. For the actual solar cells ($\eta_{rec} = 10^{-4}$) the performance has a stronger dispersion and the GaAs - based solar cells perform generally better. The conversion efficiency is obviously higher than that obtained under terrestrial conditions.

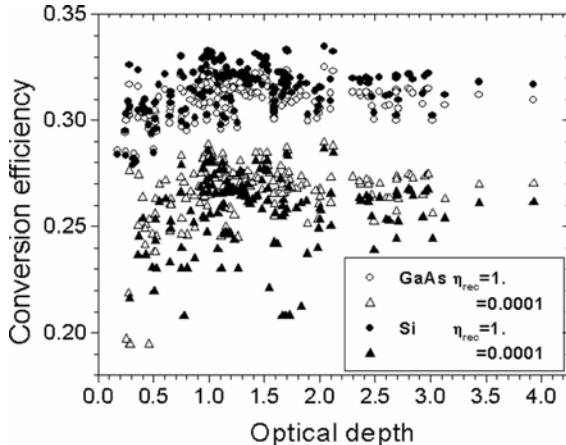


Fig. 3.1 Maximum solar cell conversion efficiency ($\eta_{cell,max}$) values at VL1 site during autumn of year 1 (sols 108 to 306) as a function of the atmospheric optical depth. A PV array perpendicular on the Sun's rays was considered (strategy P). Silicon and gallium arsenide solar cells with both ideal and actual values of recombination efficiency ($\eta_{rec}=1$ and 10^{-4} , respectively) were analyzed.

As an example, the same model was applied in Landsberg and Tonge (1980) to silicon and gallium arsenide PV arrays operating on Earth (AM1.5 and an ambient temperature of 300 K). Better optical properties were assumed in the quoted paper (the reflectance there is 0.02, compared to 0.05 used here). In case of a realistic recombination efficiency ($\eta_{rec}=10^{-4}$) the maximum conversion efficiency of well ventilated solar cells was found to be in good concordance with practical evidence, namely 15-17 % in case of silicon and 20-22 % in case of gallium arsenide (see Fig. 3a of Landsberg and Tonge (1980)). When ideal solar cells were considered ($\eta_{rec}=1$) the maximum efficiency was 25-27 % and 27-28 % for Si and GaAs, respectively (see Fig. 4a of Landsberg and Tonge (1980)). The increase in efficiency on Mars is mainly due to the low ambient temperature there, which in the case analyzed in Fig. 3.1 varies between 180 and 220 K (see e.g. Fig. 3.2c in this chapter). One could evaluate the increase in efficiency due to the decrease in the ambient temperature, if one compares the data of Figs. 4a and 5a from Landsberg and Tonge (1980), on one hand, and the data of the present Fig. 3.1, on the other hand. In case of usual solar cells ($\eta_{rec}=10^{-4}$) the relative increase in efficiency is 30-80 % for silicon cells and 10-30 % for GaAs cells. This is in good concordance with previous results which state that at an operating temperature of 200 K the output power of a silicon cell will increase by nearly 45 % whereas a GaAs cell will only increase by 20 % (Appelbaum et al. 1995). Our results show that when ideal solar cells are considered ($\eta_{rec}=1$), the relative increase in efficiency is 10-25 % for silicon and 5-15 % for GaAs. So, one confirms that the low operating temperatures will increase the cell conversion efficiency by a greater

amount for cells with a higher temperature coefficients, and thus reduce the advantage in efficiency of GaAs solar cells over silicon (Appelbaum et al. 1995).

The input meteorological data file contains only few sols with *complete* records covering appropriately the whole daylight time (see Chap. 2). Most sols have one to three records only associated to time periods close to sunrise, noon or sunset. In particular, no summer sol with complete records covering the whole day exists for VL1 or VL2 sites. For each season, we selected those sols which are richer in available data. In case of VL1, these sols are 301 (autumn) and 328 (winter). Note that no available data refers to spring. In case of VL2, the sols selected are 420 (spring) and 872 (autumn). The sol 406 (winter) is also available but no reference to it will be made here as the winter conditions are relatively similar for both Viking Landers sites.

The sol 301 is associated to autumn at VL1 site (latitude 22.3 °N). The global irradiance on a surface perpendicular on the Sun's rays is relatively high and has a bell-shape time variation (Fig. 3.2). It mainly consists in diffuse radiation. The meridional wind speed is higher during the morning while the zonal wind velocity increases during the afternoon. The ambient temperature has a maximum around 14.00 local solar time (LST) and is generally higher in the afternoon (Fig. 3.3c).

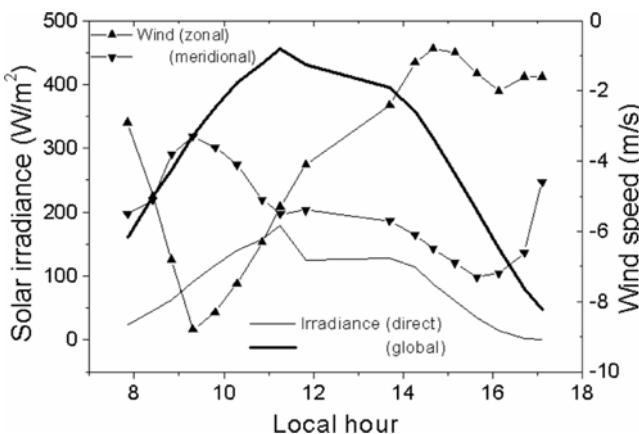


Fig. 3.2 Global and direct solar irradiance on a surface perpendicular on Sun's rays during Sol 301 (autumn) at VL1 site. Speed of zonal and meridional winds is also shown.

The performance of a PV array operating during sol 301 at VL1 site is shown in Fig. 3.3. The electrical power delivered by the array (Fig. 3.3a) has a time variation similar to that of the global solar irradiance (see Fig. 3.2). Under ideal conditions ($\eta_{rec}=1$) the silicon cell provides a slightly higher electrical power than the GaAs - based device. However, when the more realistic case $\eta_{rec}=10^{-4}$ is considered the gallium arsenide cell performs slightly better. The PV efficiency (Fig. 3.3b) is obviously weakly related to the time variation of the solar irradiance. It has a monotonous decrease during the morning till around 14.00 LST. This decrease can be

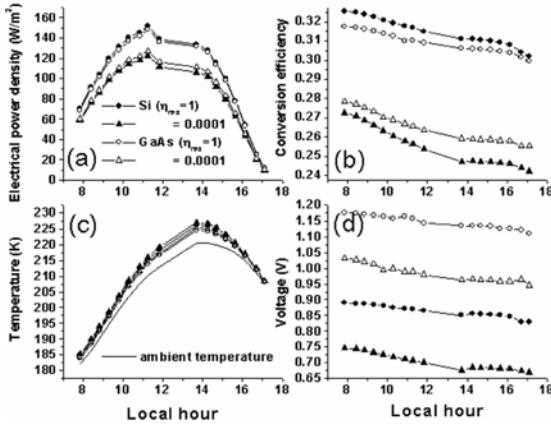


Fig. 3.3 Solar PV array performance during sol 301 (autumn) at VL1 site. The array is perpendicular on Sun's rays (strategy P). Both silicon and gallium arsenide cells were considered under ideal and realistic operating conditions ($\eta_{rec}=1$ and 10^{-4} , respectively). **(a)** the maximum electrical power per unit cell area, $P_{el,max}$; **(b)** the maximum solar cell efficiency $\eta_{cell,max}$; **(c)** optimum cell temperature $T_{c,opt}$ and ambient temperature T_a , **(d)** optimum voltage V_{opt} .

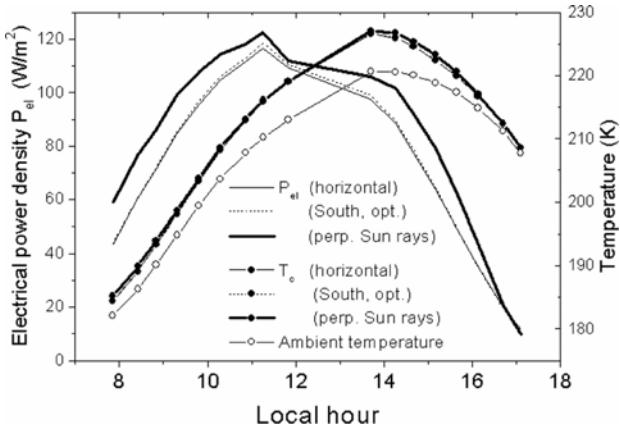


Fig. 3.4 Maximum electrical power per unit PV array area ($P_{el,max}$) and optimum cell temperature ($T_{c,opt}$) during sol 301 (autumn) at VL1 site. Silicon solar cells with $\eta_{rec}=10^{-4}$ were considered. Three strategies of collecting solar radiation were analyzed as the PV array is: (i) horizontal (H), (ii) South oriented and tilted to the instantaneous optimum slope angle which maximizes the incoming flux of global radiation (InOp) and (iii) perpendicular on Sun's rays (P).

related to the monotonous increase of the ambient temperature T_a (see Fig. 3.3c). Also, the optimum cell temperature $T_{c,opt}$ is closely related to the time variation of the ambient temperature (Fig. 3.3c). Generally, $T_{c,opt}$ exceeds T_a with a few

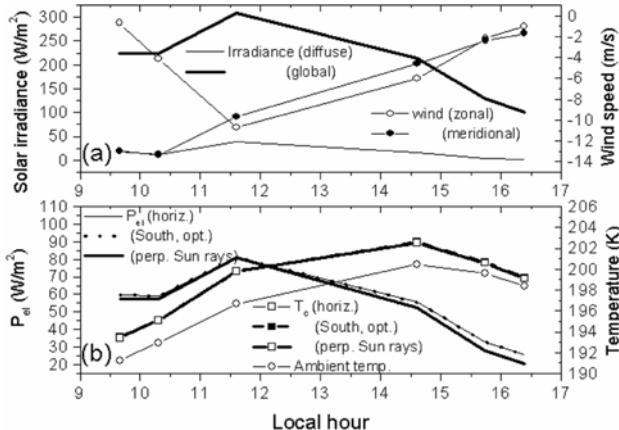


Fig. 3.5 Time variation of (a) meteorological and actinometric parameters and (b) maximum electrical power per unit PV array surface area ($P_{el,max}$) and optimum cell temperature ($T_{c,opt}$) during sol 328 (winter) at VL1 site. Silicon solar cells with $\eta_{rec}=10^{-4}$ were considered. Three strategies of collecting solar radiation were analyzed as the PV array is: (i) horizontal (H), (ii) South oriented and tilted to the instantaneous optimum slope angle which maximizes the incoming flux of global radiation (InOp) and (iii) perpendicular on Sun's rays (P).

degrees. The cell temperature slightly augments by decreasing the material band-gap and the recombination efficiency. On the contrary, the optimum voltage V_{opt} augments by increasing the material band-gap and the recombination efficiency (Fig. 3.3d). The time variation of V_{opt} is similar to the ambient temperature time variation, as already shown in case of cell temperature. Different strategies to collect solar radiation were tested in case of a PV array based on silicon cells with radiative recombination efficiency $\eta_{rec} = 10^{-4}$. As expected, the P strategy gives the higher power output (Fig. 3.4). There is little difference between the performance of the strategies H and InOp, respectively. The optimum cell temperature is similar for all these three strategies.

The sol 328 is associated to winter (the dust storms season) at VL1 site. As most of the design procedures for the Martian solar devices are based on the "worst conditions" assumption (Colozza 1990; Landis and Appelbaum 1991), this sol deserves special attention even if only six complete records are associated to it. Figure 3.5a shows the time variation during sol 328 of the principal meteorological and actinometric parameters. Solar radiation mainly consists in diffuse radiation. This is in good concordance with previous studies concerning dust storm conditions (Appelbaum and Flood 1990; Badescu 1998). The ambient temperature T_a exhibits a maximum of about 200 K around 15.00 LST. The daylight time excursion of T_a is about 10 K, which should be compared to the 40 K excursion during an autumn sol (see Fig. 3.3c). The speed of both the zonal and

meridional winds has an important variation during the day. There is little difference between the electrical output power of the PV array when different strategies to collect solar radiation are considered (Fig. 3.5b). As usual, the daily profile of $P_{el,max}$ follows that of the incident global radiation (compare Figs. 3.5b and 3.5a, respectively). The optimum cell temperature is weakly dependent on the tilt and orientation of the PV array.

3.3.2 VL2 Site

At the VL2 site (latitude 47.7 °N) the sol 420 is associated to spring, a season with relatively low atmospheric optical depth (Fig. 3.6a). The global solar irradiance incident on a surface perpendicular on Sun's rays (strategy P) is fairly higher than on a horizontal surface, especially during the afternoon, when the atmospheric optical depth decreased. This is mainly a consequence of the increase in the direct component, as the diffuse irradiance shows a weak dependence on the tilt and orientation of the receiving surface. As a result, the tracking arrays operating under the strategy P have a considerably higher electrical power output than the horizontal PV cells (strategy H) (Fig. 3.6b). The other tracking strategy, namely InOp, gives only slightly better results as compared to the horizontal arrays. The optimum cell temperature is similar for all the three strategies, being slightly higher in case of the strategy P.

The sol 872 is associated to autumn at the VL2 site. This sol is characterized by a strong time variation of the atmospheric optical depth (Fig. 3.7a). It becomes

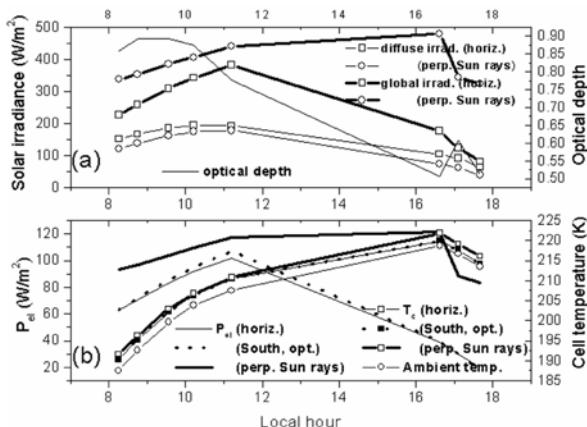


Fig. 3.6 Time variation of (a) atmospheric optical depth, global and diffuse solar irradiance and (b) maximum electrical power per unit PV array surface area ($P_{el,max}$) and optimum cell temperature ($T_{c,opt}$) during sol 420 (spring) at VL2 site. Gallium arsenide solar cells with $\eta_{rec}=10^{-4}$ were considered. Three strategies of collecting solar radiation were analyzed as the PV array is: (i) horizontal (H), (ii) South oriented and tilted to the instantaneous optimum slope angle which maximizes the incoming flux of global radiation (InOp) and (iii) perpendicular on Sun's rays (H)

evident that in the early morning and late afternoons of "clear" days a receiving surface perpendicular on Sun's rays collects much more solar energy than a horizontal surface. As usual, the time variation of the maximum electrical power follows the features of the incoming flux of global radiation (Fig. 3.7b). The South oriented tracking array gives better performance than the horizontal array but both of them perform obviously worse than the solar cells perpendicular on Sun's rays.

Finally, note that a number of factors neglected here would decrease in practice the PV arrays performance. When taking into account the array packing factor of 91 %, the interconnection efficiency of 95 % and the power conditioning efficiency of 93 %, the overall array efficiency decreases from 13.8 % to 11.9 % (McKissick et al. 1990). Also, dust settles out of the atmosphere and deposits onto the solar arrays. This could result in degradation of performance. The PV array obscuration rate depends, of course, on many factors, such as the array tilt and orientation, the wind speed and the number and intensity of the dust storms. A detailed analysis of this effect was made in Landis (1996) in case of the Pathfinder probe, which was designed for a 30 day primary mission. As an example, one of the scenarios estimated a possible obscuration rate for the 30 days mission of 5.8 %. Also, in the worst case dust deposition (landing the Pathfinder probe during a major dust storm) the PV array power by the end of the mission could be roughly two-thirds of the initial power (Landis 1996).

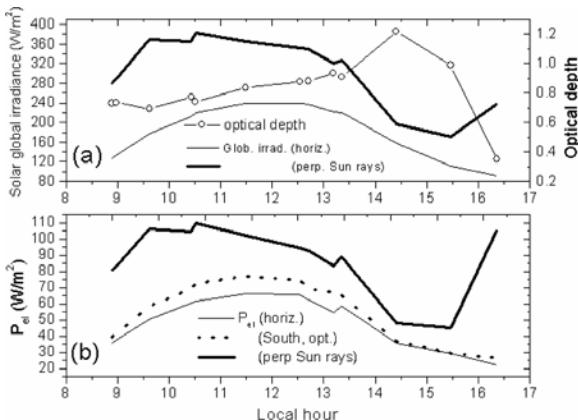


Fig. 3.7 Time variation of (a) atmospheric optical depth and global solar irradiance and (b) maximum electrical power per unit PV array surface area ($P_{el,max}$) during sol 872 (autumn) at VL2 site. Silicon solar cells with $\eta_{rec}=10^{-4}$ were considered. Three strategies of collecting solar radiation were analyzed as the PV array is: (i) horizontal, (H) (ii) South oriented and tilted to the instantaneous optimum slope angle which maximizes the incoming flux of global radiation (InOp) and (iii) perpendicular on Sun's rays (P)

3.4 Conclusions

Results concerning the simulation of solar cells operation on Mars surface are reported in this chapter. PV arrays based on silicon and gallium arsenide solar cells,

respectively, are analyzed. Two cases are considered: cells exhibiting ideal properties and cells associated to present-day technology. Results obtained at Viking Lander 2 site in autumn show that there is no obvious dependence of the PV efficiency on the optical depth. When the case of ideal solar cells is considered both the silicon and gallium arsenide solar cells have generally a similar performance, with the silicon cells performing slightly better. For the more realistic solar cell model, the performance has a stronger dispersion and the GaAs - based cells perform generally better. The PV efficiency is obviously higher than that obtained under terrestrial conditions due to the lower ambient temperature on Mars. In case of usual solar cells the relative increase in efficiency is 30-80 % in case of silicon and 10-30 % in case of GaAs. When ideal solar cells are considered, the relative increase in efficiency is 10-25 % for silicon and 5-15 % for GaAs. So, one confirms that the low operating temperatures will reduce the advantage in efficiency of GaAs solar cells over silicon.

The input meteorological data file contains only few sols with complete records covering appropriately the whole daylight time. In case of VL1 these sols are 301 (autumn) and 328 (winter) while in case of VL2 the sols are 420 (spring) and 872 (autumn). The following strategies to collect solar radiation were envisaged: (H) - the receiving surface is horizontal ; (InOp) - the receiving surface is South oriented and the tilt angle is permanently adjusted as the incident flux of global solar radiation be a maximum and (P) - the receiving surface is perpendicular on Sun's rays. The first strategy refers to a stationary collector while the last two are associated to tracking PV arrays.

The global irradiance during sol 301 (autumn) at VL1 site has a bell-shape time variation and mainly consists in diffuse radiation. The electrical power delivered by the array has a time variation similar to that of the global solar irradiance. Under ideal conditions the silicon cell provides a slightly higher electrical power than the GaAs - based device. However, when the more realistic case $\eta_{rec}=10^{-4}$ is considered the gallium arsenide cell perform slightly better. The PV efficiency has a monotonous decrease during the morning which can be related to the monotonous increase of the ambient temperature T_a . Generally, T_a exceeds cell temperature with a few degrees. Among the different strategies to collect solar radiation, the P strategy gives the higher output power, as expected. There is little difference between the performance of the strategies H and InOp. The optimum cell temperature is similar for all these three strategies.

During the winter sol 328 at VL1 site the solar radiation mainly consists in diffuse radiation. The daylight time excursion of the ambient temperature is about 10 K, which should be compared to the 40 K excursion during the sol 301 (autumn). There is little difference between the electrical output power of the PV array when different strategies to collect solar radiation are considered. The optimum cell temperature is weakly dependent on the tilt and orientation of the PV array.

At VL2 site the sol 420 is associated to spring, a season with relatively low atmospheric optical depth. The tracking arrays which follow the strategy P have a considerably higher electrical output power than the horizontal PV cells. The other strategy based on tracking PV arrays, namely InOp, gives only slightly improved

results. The optimum cell temperature is similar for all the three strategies, being slightly higher in case of the strategy P.

The sol 872 is associated to autumn at VL2 site. This sol is characterized by a strong time variation of the atmospheric optical depth. It becomes evident that in the early morning and late afternoons of "clear" days a receiving surface perpendicular on Sun's rays collects much more solar energy than a horizontal surface. As usual, the time variation of the maximum electrical power follows the features of the incoming flux of global radiation. The South oriented tracking array gives better performance than the horizontal array but both of them perform obviously worse than the solar cells perpendicular on Sun's rays.

The results reported in this chapter should be seen as upper bounds for real solar cell performances. A number of disturbing phenomena, such as dust deposition on the surface of the panels, ageing and temperature-induced degradation, were neglected. When included into the analysis, predicted performance decrease significantly.

References

- Appelbaum, J., Flood, D.J.: The Mars climate for a photovoltaic system operation. NASA Technical Memorandum 101994, Lewis Research Center, Cleveland, Ohio (1989)
- Appelbaum, J., Flood, D.J.: Solar radiation on Mars. *Solar Energy* 45, 353–363 (1990)
- Appelbaum, J., Landis, G.A.: Solar radiation for Mars power systems. In: Proceedings of the European Space Power Conference, Florence, Italy (ESA SP-320), 2-6 September 1991, pp. 707–712 (1991)
- Appelbaum, J., Landis, G.A.: Photovoltaic arrays for Martian surface power. *Acta Astronautica* 30, 127–142 (1993)
- Appelbaum, J., Sherman, I., Landis, G.A.: Solar radiation on Mars: stationary photovoltaic array. *J. Propulsion and Power* 11, 554–561 (1995)
- Appelbaum, J., Flood, D.J., Crutchik, M.: Solar radiation on Mars: tracking photovoltaic array. *J. Propulsion and Power* 12, 410–419 (1996)
- Badescu, V.: Maximum conversion efficiency for the utilization of multiple scattered solar radiation. *J. Phys. D: Appl. Phys.* 24, 882–1885 (1991)
- Badescu, V.: Different strategies for maximum solar radiation collection on Mars surface. *Acta Astronautica* 43, 409–421 (1998)
- Badescu, V., Landsberg, P.T.: Statistical thermodynamic foundation for photovoltaic and photothermal conversion. II. Application to photovoltaic conversion. *J. Appl. Phys.* 78, 2793–2802 (1995)
- Badescu, V., Landsberg, P.T., Dinu, C.: Thermodynamic optimization of non-concentrating hybrid solar converters. *J. Phys. D: Appl. Phys.* 29, 246–252 (1996)
- Badescu, V., Landsberg, P.T., De Vos, A.: Statistical thermodynamic foundation for photovoltaic and photothermal conversion. III Application to hybrid converters. *J. Appl. Phys.* 81, 3692–3699 (1997)
- Colozza, A.J.: Preliminary design of a long-endurance mars aircraft. NASA CR185243, Sverdrup Technology Inc, Aerospace Technology Park, Brookpark, Ohio 44135, prepared for Lewis Research Center under Contract NAS3-25266 (1990)
- Colozza, A.J.: Design and optimization of a self-deploying PV tent array. NASA CR 187119, Prepared for Lewis Research Center under contract NAS3-25266 (1991)

- Colozza, A.J.: Design and optimization of a self-deploying single tracking PV array. NASA CR189132, prepared for Lewis Research Center under contract NAS3-25266 (1992)
- Danescu, A., Popa, B., Radcenco, V., Carbunaru, A., Iosifescu, C., Marinescu, M., Petrescu, S., Silasi, C., Stefanescu, D., Aradau, D., Dinache, P., Madarasan, T.: Termotehnica si masini termice, p. 303. Table 12.7. Tehnica, Bucharest (1985)
- Flood, D.J.: The NASA space solar cell advanced research program. NASA Technical Memorandum 102020, Lewis Research Center, Cleveland, Ohio (1989)
- Green, M.A., Emery, K., Hishikawa, Y., Warta, W.: Solar cell efficiency tables (Version 31). *Prog. Photovoltaics: Res. Appl.* 16, 61–67 (2008)
- Hamakawa, Y.: Solar photovoltaics: recent progress and new roles. In: Krupp, H. (ed.) Energy politics and Schumpeter Dynamics, pp. 206–230. Springer, Tokyo (1992)
- Hibbs, B.D.: Mars rover feasibility study, Final Report, AeroVironment, Inc, Report AV-FR-89/7011 (1989)
- Hourdin, F.: A new representation of the absorption by the CO₂ 15- μm band for a Martian general circulation model. *J. Geophys. Res.* 97(E11), 18319–18335 (1992)
- Hourdin, F., Forget, F., Talagrand, O.: The sensitivity of the Martian surface pressure and atmospheric mass budget to various parameters: A comparison between numerical simulations and Viking observations. *J. Geophys. Res.* 100(E3), 5501–5523 (1995)
- Kahaner, D., Moler, C., Nash, S.: Numerical methods and software. Prentice Hall, New Englewood Cliffs (1989)
- Landis, G.A.: Dust obscuration of Mars solar arrays. *Acta Astronautica* 38, 885–891 (1996)
- Landis, G.A., Appelbaum, J.: Photovoltaic power options for Mars. *Space Power* 10, 225–237 (1991)
- Landsberg, P.T., Tonge, G.: Thermodynamics energy conversion efficiencies. *J. Appl. Phys.* 51, R1–R20 (1980)
- Landsberg, P.T., Baruch, B.: The thermodynamics of the conversion of radiation energy for photovoltaics. *J. Phys. A* 22, 1911–1926 (1989)
- McKissock, B.I., Kohout, L.L., Schmitz, P.C.: A solar power system for an early Mars expedition. NASA Technical Memorandum 103219, Lewis Research Center, Cleveland, Ohio, American Institute of Chemical Engineers summer National Meeting (1990)
- Reddy, M.R.: Space solar cells: tradeoff analysis. *Sol Energy Mat Solar Cells* 77, 175–208 (2003)
- Santee, M.L., Crisp, D.: Diagnostic calculations of the circulation in the Martian atmosphere. *J. Geophys. Res.* 100(E3), 5465–5484 (1995)

Chapter 4

Electric Power Generation on Mars Using Photovoltaic Helium Balloons

Pini Gurfil and Joseph Cory

Technion - Israel Institute of Technology, Haifa, Israel

4.1 Introduction

Solar energy is one of the most promising clean energy sources. For Earth applications, numerous technologies utilizing the photovoltaic effect, ranging from cellular phones to geostationary satellites, have been developed. The solar radiation reaches the Earth's upper atmosphere at a rate of $1,366 \text{ W/m}^2$ (NREL 2006, 2009). While traveling through the atmosphere, 6% of the incoming solar radiation (insolation) is reflected and 16% is absorbed, resulting in a peak irradiance at the equator of $1,020 \text{ W/m}^2$ (NASA 2006). Average atmospheric conditions (clouds, dust, pollution) reduce insolation by 20% through reflection and 3% through absorption. In North America the average insolation lies between 125 and 375 W/m^2 (3 to 9 kWh/m²/day).

On Mars, the insolation is reduced by the inverse squared Sun-Mars distance (about 1.5 AU) compared to the Earth insolation. Some studies indicated (Van Hemelrijck 1983) that the mean summer insolation on Mars lies between 150 W/m^2 and 240 W/m^2 , out of about 600 W/m^2 of incident insolation. Using photovoltaic panels, this energy can be transformed into electricity; one should keep in mind, however, that photovoltaic panels currently convert about 15-25% of sunlight into electricity. In addition, if DC/AC conversion is required, it would incur an additional energy penalty of 4-12%. Nevertheless, solar energy seems promising as a prospective energy source on Mars (see Chap. 2).

The main disadvantage of using solar electricity on Mars surface is its limited power density, due to seasonal dust storms (see Chap. 3). In this chapter, we propose to mitigate this deficiency by designing a balloon system for collecting solar electricity in the dust-free part of the atmosphere. This concept may be used to backup planned Martian power plants or as a primary energy sources where possible.

Earth balloon technology has been proven useful for various commercial, military and civil applications, including meteorological balloons, intelligence blimps, and stratospheric observatories (Gandorfer et al. 2006). To date, however, there are no existing balloons comprising embedded solar cells for electricity generation; existing technologies, developed over 25 years ago, propose either connecting an *exterior* solar array attached to airborne platforms such as balloons, kites and general aviation aircraft (Wenzel 1980), or to use a *ground* system comprising

a balloon with an embedded solar array (Hall 1978). Others proposed using a lighter-than-air airship to collect solar power and to beam it back to Earth using microwave radiation (Stark 1982).

Balloons have been proposed as means for continuous observations in future Mars exploration missions (Kerzhanovich et al. 2004), in particular imaging, magnetic field mapping and sub-surface radar mapping. Kerzhanovich et al. (2009) commented that due to the low Martian density (0.015 kg/m^3), which limits the buoyancy achievable per unit balloon volume, very lightweight balloon envelopes must be used, and balloons sizes of over 10 meters in diameter are required.

In this chapter, we propose using either Martian observation balloons or dedicated balloons as means for generating solar energy. The large surface area resulting from the large balloon diameter required for buoyancy can be coated with thin-film photovoltaic cells, and the generated electricity can be conducted to the Martian surface using cables.

We suggest a number of designs for generating electrical power using helium-filled Martian balloons carrying embedded solar cells on the balloon envelope. These balloons are to be strapped to the ground using dual-use insulated cables, carrying helium to the balloon and transporting electric charge to a ground segment. The choice of helium stems from its unique properties: Low boiling point, low density, low solubility, high thermal conductivity, and inertness. In addition, pressurized helium is available in large quantities. Because of its extremely low index of refraction, the use of helium reduces the distorting effects of temperature variations in the space between lenses in some telescopes. We shall subsequently see that this property is useful for one of our proposed designs.

In particular, we develop an aerodynamically-shaped balloon capable of mitigating wind-induced lift and drag, as well as maximizing the surface area (and hence the generated power) while minimizing volume. The balloon contains low-cost, off-the-shelf components such as solar arrays and wires. The balloon is anchored using a dual-use insulated cable, capable of conducting electricity and providing helium to the balloon. The collected power is delivered to the Martian surface using the balloon cable. It is then converted to AC (if needed) and regulated to provide stable user-defined voltage and current. The bulk of the generated heat will be radiated from the balloon surface and hence no additional radiators will be required (the specific heat capacity of helium permits passive cooling).

4.2 The Basic Concept: A Spherical Martian Balloon

Consider a spherical helium-filled balloon whose outer surface is covered by a photovoltaic array (PVA), as shown in Figure 4.1. For a given sphere radius, R , the Cartesian equation is

$$x^2 + y^2 + z^2 = R^2 \quad (4.1)$$

The surface area of the sphere is given by

$$S = 4\pi R^2 \quad (4.2)$$

while its volume is

$$V = 4\pi R^3 / 3 \quad (4.3)$$

If the balloon is filled with helium, the maximal mass that can be lifted depends on the balloon volume, the Martian “air” density and the helium density:

$$m = [\rho_{air}(H, T_a) - \rho_{He}(T_i, T_a)] V_{balloon} \quad (4.4)$$

where m is the total balloon mass, $\rho_{air}(H, T_a)$ is the altitude- and ambient temperature-dependent density of “air” and $\rho_{He}(T_i, T_a)$ is the balloon-temperature- and ambient temperature-dependant density of helium. The maximal allowable mass can be therefore obtained by substituting the Martian density at surface, $\rho_{air} = 0.015 \text{ kg/m}^3$, and the standard density of helium, $\rho_{He} = 0.00218 \text{ kg/m}^3$. For example, if $R = 5 \text{ m}$ (this particular value is chosen so as to facilitate the comparison with the alternative design discussed in Sect. 4.3), then $V \approx 523.6 \text{ m}^3$, and $m = 6.7 \text{ kg}$. The helium mass is 1.14 kg, so the total “dry” mass is 5.56 kg.

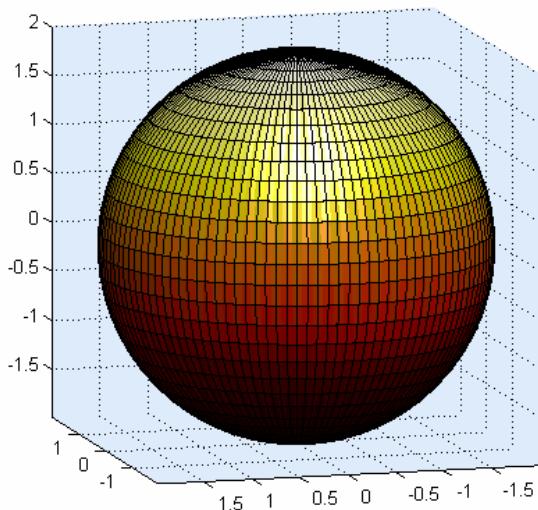


Fig. 4.1 A spherical lighter-than-“air” balloon covered by a photovoltaic array

This mass includes the light-weight structure, wiring, solar panels and radiators. If the total mass is less than 41.84 kg, then the balloon will rise. The maximum altitude can be obtained by first calculating the “air” density,

$$\rho_{air}(H, T_a) = \frac{m}{V_{balloon}} + \rho_{He}(T_i, T_a) \quad (4.5)$$

and then using dome ways to convert the resulting “air” density into altitude.

The total power generated by this balloon, assuming that only the upper hemisphere is exposed to sunlight, can be obtained by means of the relationship

$$P = P_{\text{sun}} S \eta_{\text{PVA}} / 2 \quad (4.6)$$

where P_{sun} is the total energy flux from the sun, and η_{PVA} is the Silicone or Gallium-Arsenide semiconductors photovoltaic conversion efficiency (usually in the range 0.15-0.25).

Substituting $R = 5\text{m}$ into Eq. (4.2) yields $S = 314\text{m}^2$. The average incoming solar radiation is assumed to be $P_{\text{sun}} = 200\text{W/m}^2$. Let $\eta_{\text{PVA}} = 0.15$ (this value is typical for a standard design of solar arrays). Substituting these values into Eq. (4.6) yields $P = 4.7\text{ kW}$. Hence, the balloon in our example is capable of supplying about 4.7 kW of electrical power. The power per unit volume is then

$$P/V = 9 \text{ W/m}^3 \quad (4.7)$$

4.3 The Advanced Concept: A Paraboloidic Reflector - Collector

While the spherical balloon offers the simplest implementation of a lighter-than-air PVA, it is more susceptible to aerodynamic forces (lift and drag). More importantly, the ratio of surface area to volume can be optimized to yield a much better power return per unit volume. Consequently, we propose herein two alternative designs:

1. A balloon comprised of two paraboloids wherein the upper part is transparent and the bottom part contains embedded paraboloidic collectors to increase the insolation conversion efficiency. The resulting electric power is then collected and conducted to a ground system, containing a DC/AC inverter and power regulation and control unit.
2. A balloon comprised of two paraboloids wherein the upper part is transparent and the bottom part contains an embedded paraboloidic reflector. The incident light is reflected onto the transparent part, which contains a PVA at the focus length of the bottom reflector. The resulting electric power is then collected and conducted to a ground system, containing a DC/AC inverter and power regulation and control unit.

These ideas are inspired by the observation that the solar radiation can be focused using a parabolic mirror, similar to the method of concentrating streaming light in optical telescopes: A parallel beam of light incident on the paraboloid is concentrated at the focal point (this applies for other waves, hence parabolic antennae). Alternatively, a parabolic design of the solar cells can increase the insolation-to-electricity conversion efficiency. Keeping this fact in mind, we suggest using a paraboloidic lighter-than-air inflatable helium-filled structure in which the bottom paraboloid comprises a collector/reflector, and the upper paraboloid is transparent to visible wavelengths.

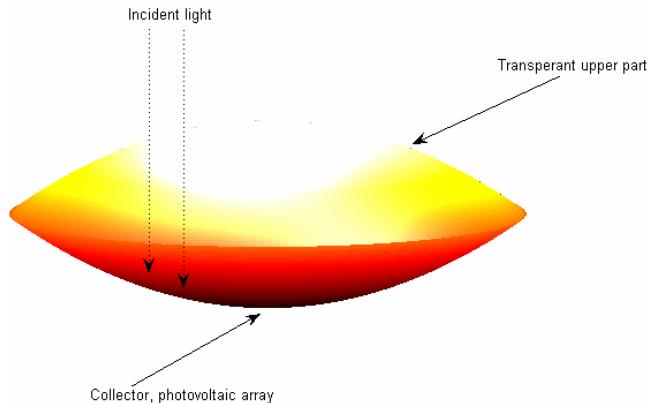


Fig. 4.2 The paraboloidic balloon. The upper part is transparent to visible wavelengths, while the bottom part is a parabolic photovoltaic collector.

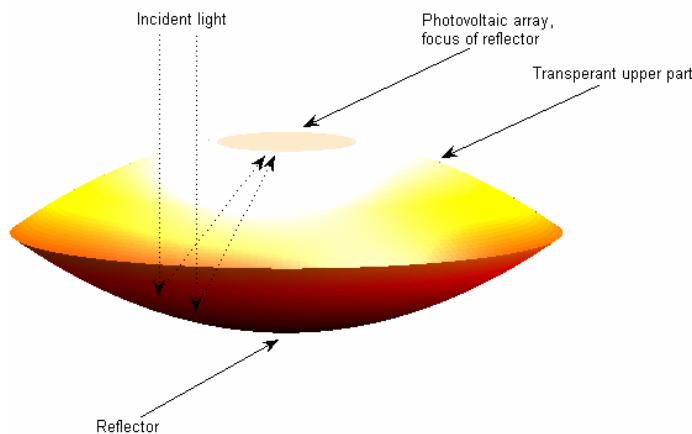


Fig. 4.3 The paraboloidic balloon. The upper part is transparent to visible wavelengths, while the bottom part is a parabolic mirror. The transparent part comprises a PVA at the focus length of the reflector, thus collecting the reflected light.

The resulting three-dimensional designs are depicted by Figs. 4.2 and 4.3. Additional geometric views are delineated in Fig. 4.4, showing the bottom and upper circular paraboloids, whose geometric properties are determined so as to guarantee that the focus length of the reflector coincides with the vertex of the upper paraboloid.

To quantify this concept, recall that a paraboloid is the surface of revolution of the parabola. The resulting quadratic surface satisfies the Cartesian equation

$$z = \frac{x^2 + y^2}{4a} \quad (4.8)$$

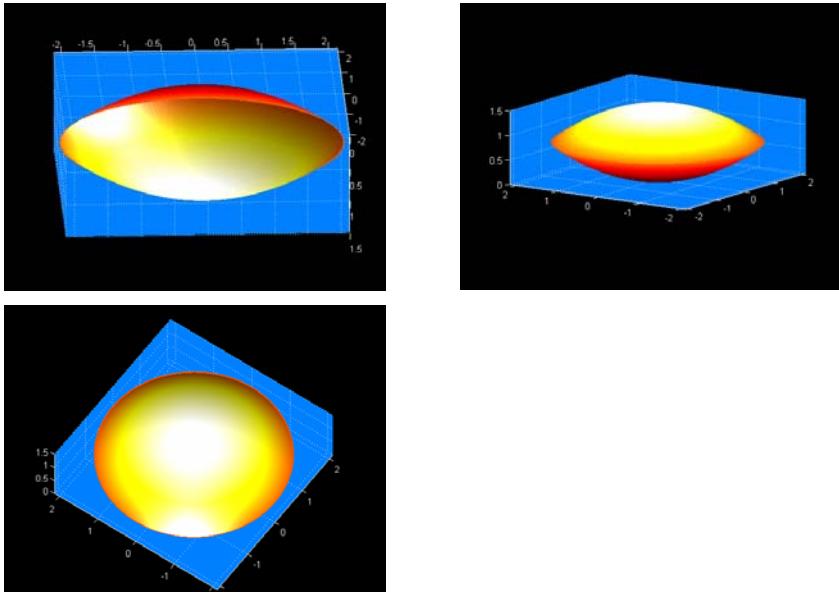


Fig. 4.4 Different viewpoints of the double-paraboloid balloon, comprising an upper and lower circular paraboloids

In this case, the focus is located at $z = a$. Since the focus is the *only* parameter defining the parabola (the distance from the focus to vertex is equal to the distance from the focus to the directrix, by definition), this constraint defines the maximal height of the reflector, the collecting surface area, the total volume, and the maximal airborne mass. These, in turn, together with given loss factors, determine the maximal power which can be generated using the balloon.

In general, a circular paraboloid that has radius α at height h is given by the parametric equations (Gray 1997)

$$x(u, v) = \alpha\sqrt{u/h} \cos v, \quad y(u, v) = \alpha\sqrt{u/h} \sin v, \quad z(u, v) = u \quad (4.9)$$

where $u \geq 0, v \in [0, 2\pi]$.

The surface area of the paraboloid satisfies (Harris and Stocker 1998)

$$S = \frac{\pi\alpha}{6h^2} \left[(\alpha^2 + 4h^2)^{3/2} - \alpha^3 \right] \quad (4.10)$$

and the volume is given by

$$V = \frac{1}{2}\pi\alpha^2 h \quad (4.11)$$

Thus, the total volume of the balloon and the total collecting area of the bottom collector/reflector satisfy, respectively,

$$V_{\text{balloon}} = 2V, \quad S_{\text{collector/reflector}} = S \quad (4.12)$$

In order for the focus of the reflector, a , to lie on the uppermost point of the collector, we must select a maximum paraboloid height of $h = a/2$. This, in turn, defines the reflector maximal radius:

$$\alpha = \sqrt{2}a \quad (\text{for } 0 \leq u \leq h = a/2) \quad (4.13)$$

To illustrate the idea, consider the balloon shown in Fig. 4.4. The reflector (bottom paraboloid) has a focal length of $a = 5.5$ m. This yields $h = 2.25$ m and $\alpha = 7.78$ m. Substituting into Eq. (4.12) yields $V_{\text{balloon}} = 523.6 \text{ m}^3$, similarly to the spherical balloon. To find the maximal allowable mass for this balloon, we shall use Eq. (4.4). The maximal allowable mass can be obtained by substituting $\rho_{\text{air}} = 0.015 \text{ kg/m}^3$, and $\rho_{\text{He}} = 0.00218 \text{ kg/m}^3$, into Eq. (4.4), yielding $m_{\text{max}} = 6.7 \text{ kg}$ – identical to the spherical balloon. The total power generated by this balloon can be obtained by means of the relationship

$$P = P_{\text{Sun}} S \eta_{\text{PVA}} \quad (4.14)$$

We assume that the total collecting area equals to the area of the reflector/collector. Substituting the focal length into Eq. (4.10) yields $S = 212 \text{ m}^2$. The average insolation is assumed to be $P_{\text{sun}} = 200 \text{ W/m}^2$ and $\eta_{\text{PVA}} = 0.25$ (this value is larger than the spherical case due to the specialized solar cell design). Substituting these values into Eq. (4.14) yields $P = 10.6 \text{ kW}$. The power per unit volume is then

$$P/V = 20.2 \text{ W/m}^3 \quad (4.15)$$

which is over twice better than the power to volume ratio of a spherical balloon.

4.4 Materials, Mechanical Design, and Aerodynamics

4.4.1 Proposed PVA Assembly

In the spherical configuration discussed above, the approximate power output is 2 kW. We shall now estimate the number of solar cells required to generate 2 kW, assuming a voltage output of 1000 V (DC) at 2 A.

The voltage produced by a solar cell is typically 0.6 V (DC). If an electrical power system requires a voltage supply of 1000 V, and has 0.6 V cells connected in series, it will need $1000V/0.6V/\text{cell} = 1667$ cells connected in series.

Since the current supplied by a single solar photovoltaic cell is on the order of 0.01 A, the cells must be connected in parallel to combine the electron flow equivalent to the required current, which, in this example, is 2 A. The total number of cells in parallel would be $2.0 \text{ A}/0.01 \text{ A}/\text{cell} = 200$ cells. The total array would

then be 1667 x 200 cells. This would develop 1000V at 2.0A. This amounts to 2A x 1000V = 2 kW of power.

4.4.2 Candidate Balloon Materials

One candidate for the inflatable structure is a silicon-impregnated balloon. The material DT891, specifically developed by Linstrandt (2009) is characterized by an increased tear strength without an increase in weight. The Silicone-impregnation technique is particularly useful for manufacturing a fabric with an embedded flexible solar arrays, since the most commonly available PVA is silicone-based. We envision two production options:

1. The bottom paraboloid is designed using a photovoltaic fabric;
2. The PVA is embedded onto the skin by adhesives.

The second option offers a simpler and less costly manufacturing process while the first option offers better durability. In both options, it is guaranteed that helium leaks will be kept to the bare minimum through a judicious selection of the fabric, maintaining minimum permeability.

Traditionally, manufacturers have used PVC to create inflatable materials. Unfortunately, PVC is heavy and has a high permeability factor. To create inflatable fabrics, urethane, which is lighter, more durable, and has lower permeability than PVC, may be used. Urethane-based inflatable fabrics are substantially lighter, offer greater durability and provide an excellent low-permeable barrier to better retain the helium.

4.4.3 Candidate PVA Materials

PVA comes in many flavors, though the bulk of the material in use today is silicon-based. In general, PVA materials are categorized as either thick crystalline or thin film (deposited in thin layers on a substrate), polycrystalline or amorphous. The PVA of interest in this study is thin film, since it considerably facilitates integration with the balloon skin. The following types are the most common thin-film PVA materials:

Amorphous Silicon (a-Si): A non-crystalline form of silicon, first used in photovoltaic materials in 1974. In 1996, amorphous silicon constituted more than 15 percent of the worldwide PV production. Small experimental a-Si modules have exceeded 10-percent efficiency, with commercial modules in the 5-7-percent range. Used mostly in consumer products, a-Si technology holds great promise in building-integrated systems, replacing tinted glass with semi-transparent modules.

Cadmium Telluride (CdTe): A thin-film polycrystalline material, deposited by electrodeposition, spraying, and high-rate evaporation, holds the promise of low-cost production. Small laboratory devices approach 16-percent efficiency, with commercial-sized modules (7200-cm^2) measured at 8.34-percent (measured total-area) efficiency and production modules at approximately 7 percent.

Copper Indium Diselenide (CuInSe₂, or CIS): A thin-film polycrystalline material, which has reached a research efficiency of 17.7 percent, in 1996, with a

prototype power module reaching 10.2 percent. The difficulty in taking this technology to a production level lies in the difficulty of avoiding defects formation during deposition. These defects prevent the formation of uniform layers.

4.4.4 Mechanical Design

The cutting-edge PVA and skin materials will be used to create a simple yet efficient mechanical setup as illustrated in Fig. 4.5. This figure shows the following system components: (1) Upper transparent paraboloid of the inflatable balloon; (2) Helium gas; (3) Rigid bend; (4) Lower opaque paraboloid, wherein the inner surface contains an embedded PVA; (5) Strapping cables for statically stabilizing the balloon; (6) Pressure valve connecting the exterior part of the coaxial cable to the bottom part of the balloon, used for initial helium inflation and occasional helium refill/discharge (for altitude control); (7) Central coaxial isolated cable, wherein the inner part conducts the DC current to the DC/AC inverter and the exterior

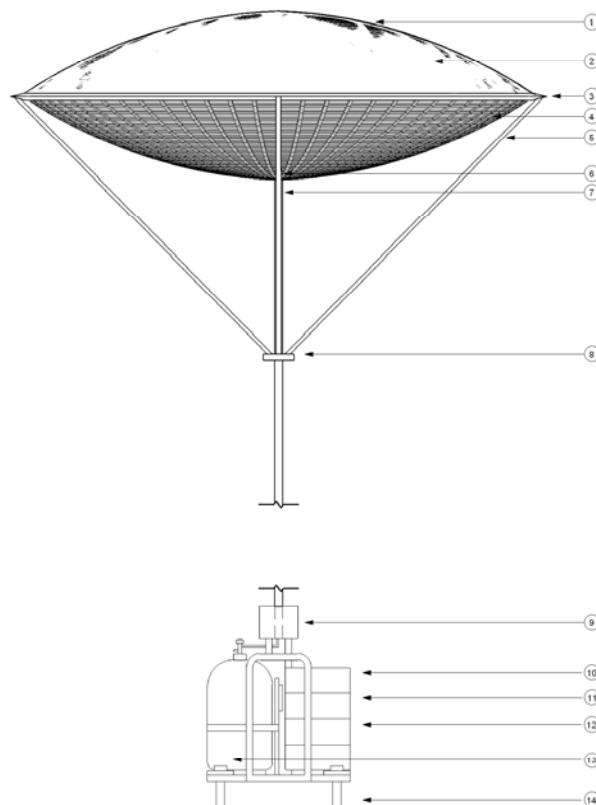


Fig. 4.5 Mechanical design of the paraboloidal balloon. This front-view diagram shows the mechanical setup including the balloon, strapping cable system, central coaxial cable and the ground segment.

isolated part is helium-filled; (8) Ring bearing connecting the wires to the central cable for wind shear mitigation; (9) Charge controller for regulating and controlling voltage; (10) DC/AC inverter transforming the DC power generated by the PVA to an AC power (optional); (11) Battery for night/clouds operation; (12) Rectifier for on-grid operation (optional); (13) Helium tank; (14) Docking platform.

The coaxial cable is depicted in Fig. 4.6. Helium is transported to the balloon as a means for altitude control. The pressure valve connected to the cable controls helium refill or discharge through activating the helium tank. The inner part of the cable is a conducting wire, used to transport the electric charge to a charge controller and to the DC/AC inverter.

Fig. 4.6 A coaxial isolated cable, wherein the outermost part is helium filled and the innermost part is a conducting wire

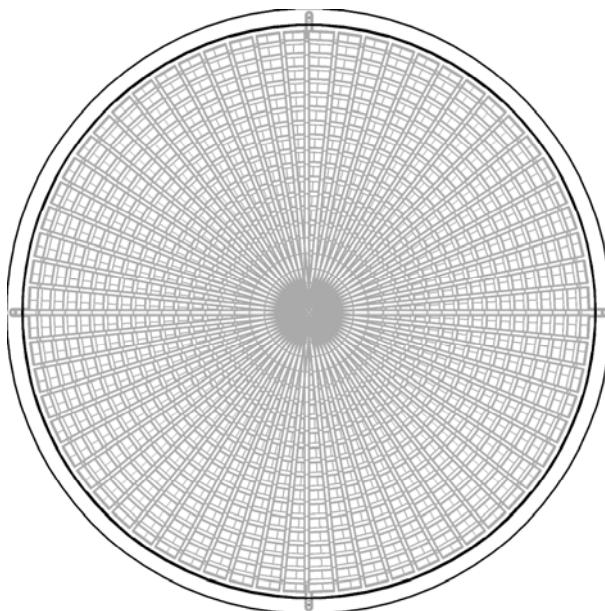
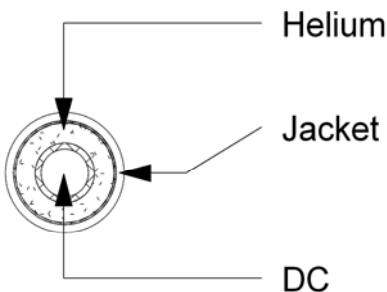


Fig. 4.7 A horizontal top-cross section of the balloon showing the thin-film PVA

Figure 4.7 shows a horizontal cross section of the balloon. The PVA array is arranged so as to produce the required voltage and current. One option for assembling the PVA is to use a thin-film silicone, as explained above. This results in a considerable weight reduction and to a marginal loss of efficiency.

4.4.5 Aerodynamics in the Pitch Plane

The balloon is subjected to a number of forces during operation. In order to guarantee durability and robustness of the mechanical design, these forces must be calculated. To that end, we shall now derive the total force balance in the pitch (vertical) plane (a similar analysis may be performed in the yaw plane, but is omitted here for the sake of conciseness), computed in a body-fixed reference frame as discussed below.

Consider a body-fixed coordinate system, centered at the balloon's center of mass, whose \hat{x} -axis points rightward along the horizontal symmetry plane and whose \hat{z} -axis points upward along the vertical symmetry plane. Let \vec{V}_w denote the Martian wind velocity vector, and α be the angle of attack, as shown in Fig. 4.8. The drag force, \vec{D} , is then given by

$$\vec{D} = \frac{1}{2} \rho V_w^2 S_{ref} C_D \hat{v} \quad (4.16)$$

where ρ is the atmospheric density, S_{ref} is a reference area, C_D is the drag coefficient and \hat{v} is a unit vector along the wind velocity vector, as shown in Fig. 4.8. The lift force due to wind is given by

$$\vec{L}_w = \frac{1}{2} \rho V_w^2 S_{ref} C_L \hat{n} \quad (4.17)$$

where C_L is the lift coefficient and \hat{n} is a unit vector normal to wind velocity direction. In addition to the aerodynamical lift, a gas buoyancy lift force, \vec{L}_B , acts upon the balloon due to the lighter-than-“air” medium. This force is given by (cf. Eq. (4.4))

$$\vec{L}_B = g (\rho - \rho_{He}) V_{balloon} \hat{z} \quad (4.18)$$

Finally, the balloon weight is

$$\vec{W} = -mg\hat{z} \quad (4.19)$$

The above forces are balanced using the strapping cables tension forces, $\vec{T}_1, \vec{T}_2, \vec{T}_3, \vec{T}_4$. If equilibrium is assumed, then writing the moment equation about the center of mass will yield

$$\vec{T}_1 \approx \vec{T}_2 \approx \vec{T}_3 \approx \vec{T}_4 \quad (4.20)$$

In reality, the aerodynamic forces act at the aerodynamic center, and not at the center of gravity. We assumed here that the distance from the aerodynamic center to the center of gravity is negligible relative to the balloon size. In practice, this effect will cause a slightly different tension force in each cable.

Let δ be the angle between the strapping cable and the balloon horizontal cross section, as shown in Fig. 4.8. The equilibrium force equation in the \hat{z} direction under the constraint Eq. (4.20) is given by

$$L_w \cos \alpha + D \sin \alpha + L_B = -4T_1 \sin \delta - W \quad (4.21)$$

Substituting the expressions in Eqs. (4.16)-(4.19) yields

$$\begin{aligned} & \frac{1}{2} \rho V_w^2 S_{ref} C_L \cos \alpha + \frac{1}{2} \rho V_w^2 S_{ref} C_D \sin \alpha + g(\rho - \rho_{He}) V_{balloon} \\ & = -4T_1 \sin \delta + mg \end{aligned} \quad (4.22)$$

Similarly, the equilibrium force equation in the \hat{x} direction is

$$L_w \sin \alpha + T_2 \cos \delta = D \cos \alpha + T_1 \cos \delta \quad (4.23)$$

from which we find that

$$\frac{1}{2} \rho V_w^2 S_{ref} C_L \sin \alpha = \frac{1}{2} \rho V_w^2 S_{ref} C_D \cos \alpha \quad (4.24)$$

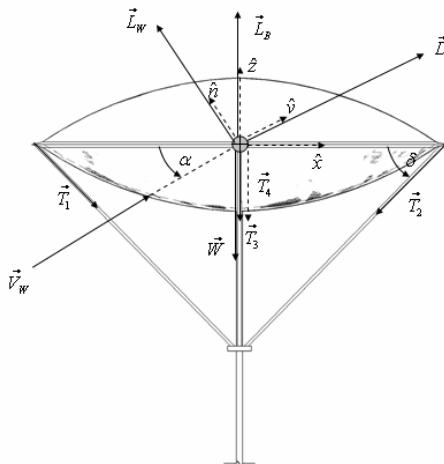
For small angles of attack, we may use the approximation $\cos \alpha \approx 1, \sin \alpha = \alpha$. Under this assumption, Eq. (4.24) simplifies into

$$\alpha \approx \frac{C_D}{C_L} \quad (4.25)$$

Substituting Eq. (4.25) into Eq. (4.22) yields an estimate of the tension force acting on each cable:

$$T_C \approx -\frac{1}{4 \sin \delta} \left[\frac{1}{2} \rho V_w^2 S_{ref} C_L + \frac{1}{2 C_L} \rho V_w^2 S_{ref} C_D^2 \right] + g(\rho - \rho_{He}) V_{balloon} - mg \quad (4.26)$$

A negative tension implies that the balloon operates at some designated altitude, while a positive tension implies that the balloon is below the desired altitude. For example, using the numerical values from Sect. 3, and assuming that $C_L = 0.5, C_D = 0.1, \delta = 45^\circ$ yield a tension force of about 13 kg in each cable for a wind speed of 30 m/s.

Fig. 4.8 Forces acting on the balloon

4.5 Impact on Future Mars Exploration

Our proposed balloons constitute an efficient, infrastructure-free energy source for potential Martian exploration missions, including surface exploration by astronauts and robotic vehicles. The balloons can be delivered by an orbiter from the outside of the Martian atmosphere onto pre-designated landing sites, thus facilitating energy utilization for future Martian explorers. The balloons are highly portable and can thus be mobilized in compact backpacks by astronauts and ground vehicles.

The proposed design is versatile and adaptable, so it can be utilized in diverse applications, ranging from local lighting through communication antennae to emergency power generation.

The idea to use balloons instead of covering vast areas of the Martian surface with thin-film photovoltaic cells offers three main advantages:

1. There is no need to mobilize explorers throughout the unknown, potentially dangerous, Martian surface.
2. The balloons constitute a point-source of energy, which can be assembled in-situ next to the explorers' camp.
3. Raising the thin-film photovoltaic cells above the surface mitigates the effect of dust-storms and topography.

Moreover, the proposed balloons may be used in a variety of compound design concepts:

- Balloons can be assembled to form a spatial structure such as the one shown in Fig. 4.9. This assembly will increase the power return and will enable to increase the buoyancy of the system.
- Balloons may be interconnected to form lighter-than-“air” structures, thus maximizing the power return while minimizing the environmental signature. Such a structure is shown in Fig. 4.10.

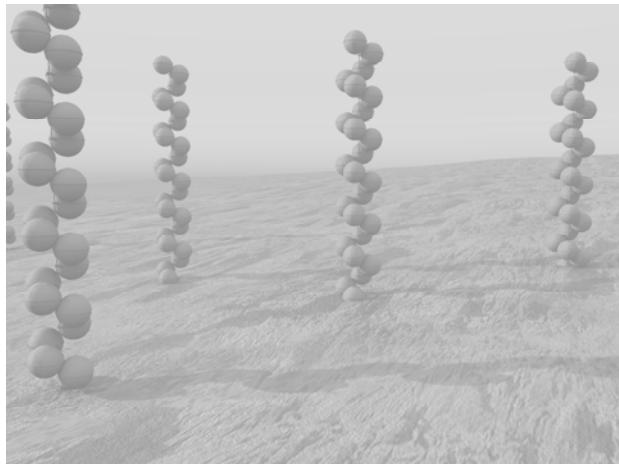


Fig. 4.9 Structures of interconnected balloons with embedded photovoltaic arrays for generating electricity on Mars

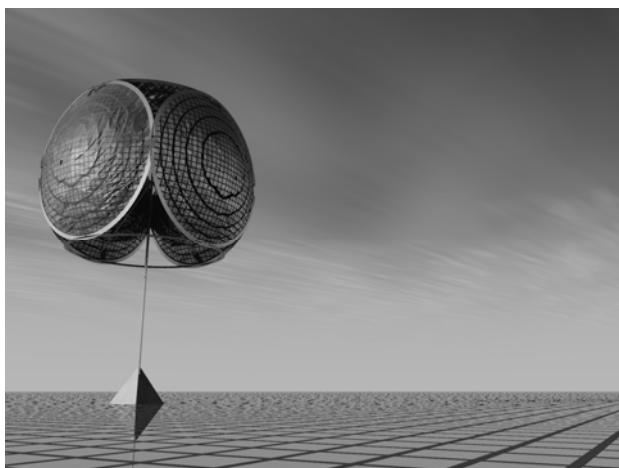


Fig. 4.10 A compound concept wherein 5 balloons are arranged in a spatial structure to maximize power return and sun exposure over Martian Polar Regions

We have created a number of renderings of the balloons, showing the potential applicability of the project as well as possible environments for locating balloons. These are shown in Figs. 4.9-4.10.

4.6 Discussion and Conclusions

This chapter presented a feasibility study of a novel platform for transforming solar power into electricity. It was proposed to design a helium balloon with an

embedded photovoltaic array (PVA) in order to collect the incoming solar radiation on Mars. Three main design alternative were proposed: A spherical balloon containing an embedded PVA on its outer skin; a double-paraboloidic balloon with a transparent upper part and a PVA on the inner surface of its bottom part; and an advance design, wherein the inner PVA is replaced by concentrating mirror, reflecting the incoming solar radiation on a PVA located on the upper paraboloid.

The new design developed herein has several clear advantages: It constitutes an accessible, portable and infrastructure-free access to electric power on Mars that could be used in diverse locations.

Basic balloons can be used as either a primary or a secondary reliable electric energy source in Martian exploration missions. The advanced balloons, comprising a paraboloidic balloon with a bottom collector and an upper transparent part or a bottom reflector and an upper collector provides a smooth aerodynamical design. Future applications may include flocks of balloons which are inter-connected and/or inter communicating. Moreover, an ensemble of balloons floating in various altitudes can create a Martian lighter-than-“air” mini power station.

Future research will be conducted to bridge some of the gaps arising in the proposed technology. One issue is icing, which may have an impact on the efficiency of the photovoltaic array.

References

- Gandorfer, A.M., Solanki, S.K., Barthol, P., Lites, B.W., Martinez-Soltau, D., Tile, A.M.: SUNRISE: High-Resolution UV/VIS Observatories of the Sun from the Stratosphere. *Adv. Space Research* 42, 70–77 (2006)
- Gray, A.: The Paraboloid. Chap. 13.5 in *Modern Differential Geometry of Curves and Surfaces with Mathematica*, 2nd edn, pp. 307–308. CRC Press, Boca Raton (1997)
- Harris, J.W., Stocker, H.: Paraboloid of Revolution. In: *Handbook of Mathematics and Computational Science*, ch. 4.10.2, p. 112. Springer, New York (1998)
- Hall, F.F.: Solar Energy Collector Including a Weightless Balloon with Sun Tracking Means. US Patent 4,126,123 (1978)
- Kerzhanovich, V.V., Cutts, J.A., Cooper, H.W., Hall, J.L., McDonald, B.A., Pauken, M.T., White, C.V., Yavrouian, A.H., Castano, A., Cathey, H.M., Fairbrother, D.A., Smith, I.S., Shreves, C.M., Lachenmeier, T., Rainwater, S.E.: Breakthrough in Mars Balloon Technology. *Adv. Space Research* 33, 1836–1841 (2004)
- Linstradt (2009), <http://www.linstradt.com>
- NASA, Earth Radiation Budget. NASA Langley Research Center (2006),
<http://asd-www.larc.nasa.gov/erbe/ASDerbe.html>
- NREL, Solar Spectra: Standard Air Mass Zero. National Renewable Energy Laboratory Resource Data Center (2006), <http://www.nrel.gov/>
- NREL, Dynamic Maps, GIS Data, and Analysis Tools - Solar Maps. National Renewable Energy Laboratory Resource Data Center (2009),
<http://www.nrel.gov/gis/solar.html>
- Stark, V.: Apparatus for Collecting Solar Energy at High Altitudes and on Floating Structures. US Patent 4,364,532 (1982)
- Van Hemelrijck, E.: The Effect of Orbital Element Variations on the Mean Seasonal Daily Insolation on Mars. *Moon and Planets* 28, 125–136 (1983)
- Wenzel, J.: Solar Power Station. US Patent 4,361,295 (1980)

Chapter 5

Weather Influence on Solar Thermal Power Plants Operation on Mars

Viorel Badescu

Polytechnic University of Bucharest, Romania

5.1 Introduction

For many years space researchers realized that in some specific situations solar dynamic power systems can provide significant savings in life cycle costs when compared with conventional photovoltaic power systems with battery storage (Menetrey 1963; Secunde et al. 1989; Prisnjakov 1991; Prisnjakov et al. 1991). A standard solar dynamic power system uses a mirror to concentrate solar radiation onto an absorber structure. By conduction through a solid material or circulation of a working fluid, the absorber heat is transferred to a thermal engine (i.e. a turbogenerator, Stirling engine, thermocouple or thermionic-emitter). Alternators coupled to these thermal engines may generate electrical energy. Previous practice proved that three different cycles can be used for thermodynamic conversion of solar radiation: Brayton, Rankine and Stirling (Menetrey 1963; Prisnjakov 1991; Prisnjakov et al. 1991). For continuous operation during dark periods the use of melted materials to store thermal energy is being considered. Among the advantages of dynamic power systems one could quote their ability to provide electrical energy and heat simultaneously, the fact that the power plant may be unified by using either solar or nuclear energy or their relative invulnerability to corpuscular particles and to electromagnetic radiation and the possibility of power control according to a given power consumption schedule (Prisnjakov 1991).

Earlier studies reported solar dynamic systems based on thermodynamic cycles and thermionic converters, respectively. Two systems of the former type were described in Menetrey (1963). One of them is a 3 kW solar dynamic system of 10 % overall efficiency consisting of a parabolic concentrator and a turbo-generator unit which uses mercury as working fluid in a Rankine cycle. The other one is a 15 kW system for space vehicle applications. Its working fluid is rubidium while the overall electrical efficiency is 21.7 %. The solar - thermionic system was designed to provide 135 W at Mars distance from the Sun with a system weighting about 14 kg. Other studies are using permanent gases (mainly mixtures of xenon and helium) as working fluid in Brayton cycles. They have output power of 3 to 5 kW, specific weights up to 40-45 kg/kW and overall efficiency of 31.2 % (Prisnjakov 1991; Prisnjakov et al. 1991, 1994). Significant efforts to develop solar dynamic power systems (25 kW) for application to a low Earth orbit space station have been conducted by NASA Lewis Research Center (LeRC) as an integral part of

the US space station development program (McLallian et al. 1988). The culmination of the LeRC programme demonstration program was a full system test in a thermal/vacuum tank of a Space Station Freedom type solar dynamic system at the 2 kWe size in February 1995. Related reports can be found in Secunde et al. (1989) and Weingartner et al. (1994). A comparison between the various power systems designed for the Space Station Freedom shows the sun-to-user efficiency is 5.7 to 6.8 % for the PV system and 18.7 to 19.8 % for the dynamic system. If, in addition, the storage subsystem is taken into consideration, the solar dynamic system becomes even more attractive, mainly because the efficiency of the latent heat storage is higher compared to the Ni-Cd batteries of the PV system (for details see Weingartner et al. (1994)).

A natural question arising about power systems to be used on Mars surface is: are properly designed dynamic systems comparable in performance with the PV – based power systems? Some of the above references seem to indicate an affirmative answer. However, they all refer to systems operating in the interplanetary space. A solar power system placed on Mars surface has to take into consideration one very important feature of the Martian weather. The period between the areocentric longitude $L_s=161 - 326^\circ$ and variants thereof have been referred to as the “dust storm season”. This period is nearly centered on perihelion, which is the time of maximum insolation on Mars (Martin and Zurek 1993; Hourdin et al. 1995). In the Northern hemisphere, the Martian atmosphere varies greatly from year to year during the dust storm season while during the other seasons, most phenomena are remarkably similar from year to year. A comprehensive description of Martian meteorology is provided by Zurek et al. (1992). Chapter 2 gives some information on the subject. The dust storm periods are characterized by an almost vanishing direct solar radiation (see e.g. Badescu (1998a) and Chap. 2). The main consequence is that isotropic diffuse solar radiation can not be concentrated and flat plate solar collectors have to be used in this case.

The structure of this chapter is as follows. Upper bounds for the efficiency of solar thermal power plants operating in the Martian environment are evaluated in section 5.2. A model of a solar power plant based on a Carnot cycle is presented in section 5.3 while section 5.4 refers to solar Stirling engines based on selective flat-plate converters. These models provide lower upper limits for the performance of actual thermal solar engines. Section 5.5 contains the conclusions.

5.2 Upper Limits for Solar Energy Conversion Efficiency into Work

The purpose of this section is to give information about the upper limits of the maximum efficiency of converting solar energy into mechanical work on the surface of Mars. This subject was already treated in a number of papers in case of solar energy conversion systems placed on Earth surface. Changing the conversion system on Mars does not require a different theoretical framework. However, the characteristics of Mars meteorology is expected to have considerable influence on the numerical results.

The upper bounds for the performance presented in this section are obtained by using highly idealized models. They have however the advantage of giving a perspective for the results obtained with the more realistic models developed in the next sections.

5.2.1 Exergy of Solar Radiation on Mars

The exergy of thermal radiation was studied in many papers, starting with the seminal work of (Petela 1964). Controversial arguments were presented in literature (see e.g. Bejan (1988) for an early review). More recently, the discussion was re-opened in relation with the exergy of solar radiation (Petela 2003; Candau 2003). Also, many authors studied the thermodynamics of converting radiation energy into mechanical work (for early reviews see Landsberg and Tonge (1979), Bejan (1988). Three different theories were proposed by Jeter (1981), by Landsberg, Petela and Press (Petela 1961; Landsberg and Mallison 1976; Press 1976) and by Spanner (1964), respectively. The three theories were unified in Bejan (1987) while a generalization was given in Badescu (1988). The exergy factor proposed by Landsberg-Petela-Press, which is very often used in literature, is:

$$\eta_{LPP} = 1 - \frac{4}{3} \frac{T_0}{T_s} + \frac{1}{3} \left(\frac{T_0}{T_s} \right)^4 . \quad (5.1)$$

where T_0 and T_s are the ambient and Sun temperatures, respectively.

This factor shows the fraction of thermal radiation energy that can be converted into work. Note that η_{LPP} depends on just two temperatures (namely the temperatures of the two heat reservoirs).

Let us apply the above result in case of solar radiation on Mars. The temperature of solar direct radiation is $T_s=5760$ K while the average temperature of Mars surface is $T_0=248$ K. Equation (5.1) allows to compute the value of the Landsberg-Petela-Press exergy factor: $\eta_{LPP} = 0.9246$. For a solar direct irradiance at ground level $\varphi_s=400$ W/m², the exergy flux is 377.04 W/m².

5.2.2 Accurate Upper Bounds for Solar Energy Conversion Efficiency

The theories shortly reviewed in the previous subsection predict efficiencies of converting radiation energy into work too high to be of practical interest. Much more accurate rigorous upper bounds for the real solar energy conversion efficiency η were reported in recent years (Badescu 1998b, 1998c). Some of these new results are listed below and applied to the special case of solar energy conversion on Mars surface.

(i) In case of a smaller than hemispherical blackbody radiation source and a blackbody receiver coupled to a reversible engine the upper bound of η is (Badescu 1998b):

$$\eta_i = 1 - \frac{4}{3} \left(f \frac{T_0}{T_s} \right) + \frac{1}{3} \left(f \frac{T_0}{T_s} \right)^4 > \eta , \quad (5.2)$$

where the geometric factor

$$f \equiv \left(\frac{\Omega}{\pi} \right)^{-1/4} \left(1 - \frac{\Omega}{4\pi} \right)^{-1/4} \quad (5.3)$$

is a function of the solid angle Ω subtended by the radiation source. In case of a spherical source of radiation (e.g. the Sun) the following simple relation applies:

$$f = \sin^{-1/2} \delta \quad (5.4)$$

where δ is the half-angle of the cone subtended by the source. Of course, when a hemispherical source of radiation is considered, $\delta = \pi/2$ and η_i given by Eq. (5.2) reduces to η_{LPP} given by Eq. (5.1). However, generally $\eta_i < \eta_{LPP}$. One can see that η_i depends on just two temperatures (namely the temperatures of the two heat reservoirs) in a way similar to the Carnot efficiency. But it also takes account of the geometry of the radiation source and this additional information increases its accuracy.

Let us apply Eqs. (5.2)-(5.4) to solar energy conversion on Mars. The case of direct solar radiation is considered. Then, δ is the half-angle of the cone subtended by the Sun when viewed from Mars. The mean distance Sun to Mars is $d_{Sun-Mars}=227.94 \cdot 10^6$ km while Sun radius is $R_s=696 \cdot 10^3$ km. Consequently, $\delta = \tan^{-1}(R_s/d_{Sun-Mars}) \approx \sin^{-1}(R_s/d_{Sun-Mars}) = 0.003053$. Equation (5.4) yields $f=18.096$. Use of Eq. (5.2) gives the upper bound $\eta_i = 0.08393$. This is obviously a *more accurate* upper bound than those predicted in section 5.2.1.

(ii) In case of smaller than hemispherical blackbody radiation source and a blackbody receiver coupled to a Chambadal-Novikov-Curzon-Ahlborn (CNCA) engine the upper bound of η is (Badescu 1998b)

$$\eta_{ii} = 1 - \frac{4}{3} \left(f \frac{T_0}{T_s} \right)^{1/2} + \frac{1}{3} \left(f \frac{T_0}{T_s} \right)^2 > \eta. \quad (5.5)$$

This is of course a more accurate estimate for η than η_i is, because part of the entropy generation related to heat engine operation is taken into account. As an application, using Eq. (5.5) in case of solar direct radiation on Mars yields $\eta_{ii} = 0.02542$ which is lower than the value derived previously for η_i , as expected.

Note that a formula corresponding to η_{LPP} , in case the system contains a CNCA engine, can be obtained from Eq. (5.5) by making $f = 1$. One finds

$$\eta'_{ii} \equiv 1 - \frac{4}{3} \left(\frac{T_0}{T_s} \right)^{1/2} + \frac{1}{3} \left(\frac{T_0}{T_s} \right)^2. \quad (5.6)$$

Equation (5.6) provides a more accurate estimate for η than η_{LPP} , but a less accurate estimate than η_{ii} . Using Eq. (5.6) in case of solar direct radiation on Mars yields $\eta'_{ii} = 0.72395$. This is a rather high upper bound for the real efficiency and this loss in accuracy is the cost of using a simpler upper limit formula.

(iii) Repeating the procedure from Bejan (1987) but including details about some irreversibilities one finds relationships generalizing Jeter, Spanner and Landsberg-Petela-Press efficiencies, respectively (Badescu 1999). For example, the relation generalizing the last efficiency is:

$$\eta_{ii} \equiv 1 - \frac{4}{4-n} \left(\frac{T_0}{T_s} \right)^n + \frac{n}{4-n} \left(\frac{T_0}{T_s} \right)^4, \quad (5.7)$$

where n is a parameter related to the type of thermal engine. In case of $n = 1$ (i.e. a Carnot engine) one finds of course from Eq. (5.7) the efficiency η_{LPP} . Let us consider a CNCA engine ($n = 1/2$) as part of a power system converting the energy of direct solar energy on Mars. Then Eq. (5.7) yields the upper limit $\eta_{ii} = 0.76286$. This rather high upper bound is comparable in size to η'_{ii} .

(iv) The case of gray thermal radiation is of special interest, as it is sometimes used as an approximation for diffuse solar radiation. Note that gray radiation is a particular case of diluted radiation. A general theory was first developed by Landsberg and co-workers (see for instance the review paper by Landsberg and Tonge (1979)). However, important results were derived more recently and they will be summarized now. An accurate upper bound for the conversion efficiency of solar diluted radiation into work is given by Badescu (1990)

$$\eta_d \equiv 1 - \frac{4}{3} \left(f_d \frac{T_0}{T_s} \right) + \frac{1}{3} \left(f_d \frac{T_0}{T_s} \right)^4. \quad (5.8)$$

Here the dimensionless factor f_d can be obtained from Eqs. (10) and (14) of Badescu (1990):

$$f_d \equiv \chi(\varepsilon_d) \left(\frac{\Omega_d}{\pi} \right)^{-1/4} \left(1 - \frac{\Omega_d}{4\pi} \right)^{-1/4}, \quad (5.9)$$

where Ω_d is the solid angle subtended by the source of diluted radiation source while $\chi(\varepsilon_d)$ is a function of the radiation dilution factor ε_d . This function was calculated in Landsberg and Tonge (1979). It is such that $\chi(1)=1$ and can be approximated for small ε_d by:

$$\chi(\varepsilon_d) = 0.9652 - 0.2777 \ln \varepsilon_d + 0.0511 \varepsilon_d. \quad (5.10)$$

The above general formulation yields elegant results in the particular case of solar scattered radiation. The upper bound for the efficiency of converting into

work the energy of multiply scattered solar radiation was derived in Badescu (1991) and is given by

$$\eta_{d,i} \equiv 1 - \frac{4}{3} \left(f_{d,i} \frac{T_0}{T_s} \right) + \frac{1}{3} \left(f_{d,i} \frac{T_0}{T_s} \right)^4, \quad (5.11)$$

where the factor $f_{d,i}$ depends on the number i of scatterings ($i=1,2,3,4$). It is computed by

$$f_{d,i} \equiv \left(\frac{\Omega_{d,i}}{\pi} \right)^{-1/4} \left(1 - \frac{\Omega_{d,i}}{4\pi} \right)^{-1/4} \prod_{j=1}^i \chi(\varepsilon_{d,i}) \quad (i=1,2,3,4). \quad (5.12)$$

In case the scattered radiation is not subsequently concentrated, the following relations apply for the dilution factors and solid angles of i -scattered solar radiation, respectively:

$$\begin{aligned} \varepsilon_{d,1} &= \frac{\Omega_s}{\pi} & \varepsilon_{d,i} &= \frac{\Omega_s}{\pi} \left[\prod_{j=1}^{i-1} \chi^4(\varepsilon_j) \right] \quad (i=2,3,4) \\ \Omega_{d,i} &= 2\pi \left\{ 1 - \left[1 - \frac{\Omega_s}{\pi} \prod_{j=1}^i \chi^4(\varepsilon_{d,i}) \right]^{\frac{1}{2}} \right\} \quad (i=1,2,3,4), \end{aligned} \quad (5.13,14)$$

where Ω_s is the solid angle subtended by the Sun. The half-angle $\delta_{d,i}$ of the cone subtended by the source of i -scattered radiation can be easily computed from:

$$\Omega_{d,i} = 2\pi(1 - \cos \delta_{d,i}). \quad (5.15)$$

The effective temperature $T_{d,i}$ of i -scattered solar radiation is given by:

$$T_{d,i} = \frac{T_s}{\prod_{j=1}^i \chi(\varepsilon_{d,i})}. \quad (5.16)$$

Table 5.1 shows results obtained in case of i -scattered solar radiation on Mars. One can see that after a single scattering (i.e. $i=1$) solar radiation is still strongly anisotropic (compare $\delta_{d,1} = 3.061^\circ$ with the half-angle $\delta_s = 0.1749^\circ$ of the cone subtended by the Sun). After four scatterings solar radiation is isotropic and an observer at the ground would see a uniformly brilliant sky. The effective temperature of the scattered radiation decreases by increasing the number of scatterings, as expected. However, the upper bound efficiency $\eta_{d,i}$ predicted by Eq. (5.11) yields the same value, whatever the number of scatterings. This is a consequence of keeping the energy flux constant during the scattering process.

Table 5.1 Upper bounds for the conversion efficiency of i -scattered solar radiation into work on Mars.

Number of scatterings, i	$\varepsilon_{d,i}$	$\Omega_{d,i}$	$\delta_{d,i}$ (deg)	$T_{d,i}$ (K)	$\eta_{d,i}$	$C_{max,i}$	$\eta_{d,conc,i}$
1	0.00000932	0.00897	3.06	1377.0	0.084	350.70	0.760
2	0.00285100	0.41880	21.04	531.2	0.084	7.76	0.393
3	0.12890000	2.99700	58.47	344.8	0.084	1.38	0.130
4	0.72650000	6.28300	90.00	316.0	0.084	1.00	0.084

$\varepsilon_{d,i}$ - dilution factor. $\Omega_{d,i}$ and $\delta_{d,i}$ - solid angle and half-angle of the cone subtending the source of radiation, respectively. $T_{d,i}$ - effective temperature. $\eta_{d,i}$ - upper bound efficiency. $C_{max,i}$ - maximum concentration ratio. $\eta_{d,conc,i}$ - upper bound efficiency for concentrated i -scattered radiation. Sun temperature $T_s=5760$ K; average temperature on Mars surface $T_0=248$ K.

For small i -values scattered solar radiation is generally anisotropic. Consequently, it can be concentrated. The maximum concentration ratio $C_{max,i}$ is given by:

$$C_{max,i} = \left[\frac{\Omega_{d,i}}{\pi} \left(1 - \frac{\Omega_{d,i}}{4\pi} \right) \right]^{-1}. \quad (5.17)$$

The upper bound efficiency $\eta_{d,conc,i}$ for fully concentrated scattered radiation is given again by Eq. (5.11) with $f'_{d,i}$ replaced by $f''_{d,i}$ given by:

$$f''_{d,i} \equiv \prod_{j=1}^i \chi(\varepsilon_{d,j}) \quad (i=1,2,3,4). \quad (5.18)$$

The upper bound efficiency $\eta_{d,conc,i}$ lies between 0.760 and 0.084 for single and four scatterings, respectively (Table 5.1). Generally, the effective temperature of scattered radiation is lower, and the upper bound efficiency is higher, on Mars than on Earth (compare present Table 5.1 with Table 1 of Badescu (1991)). This is mainly caused by the lower ambient temperature on Mars.

(v) The upper bounds at (i) to (iii) above apply to any source of blackbody undiluted or diluted radiation. Even more accurate (but also more complex) upper bounds of η were derived in case of solar energy applications (Badescu 1998c, 2000). These upper bounds do not involve the absorber temperature but are expressed just in terms of the two heat reservoir temperatures. However, they take into account a large number of details concerning the conversion process as for example, the dilution factors of solar and ambient radiation, the polarization degree of solar and absorber emitted radiation and the thermal and optical properties of the selective absorber (convective and conductive heat loss coefficients, concentration ratio, effective transmittance-absorptance product and emittances). These details increase of course the accuracy of the efficiency formula. As an example we shall use the accurate upper bound efficiency Eq. (36) of Badescu (1998c). For the purposes of this chapter we shall simplify it to take account just

on the selective properties of the absorber (i.e. its absorptance a and emittance e). Then, the upper bound efficiency is given by:

$$\eta_v = a \left[1 - \frac{4}{3} \left(f \frac{T_0}{T_s} \right) + \frac{1}{3} \left(f \frac{T_0}{T_s} \right)^4 \right], \quad (5.19)$$

where

$$f = \left(\frac{e}{a} \right)^{1/4} \sin^{-1/2} \delta. \quad (5.20)$$

For an usual selective flat-plate solar collector with $a = 0.9$ and $e = 0.1$, Eq. (5.19) yields $\eta_v = 0.3724$ which is considerably higher than the upper bound efficiency $\eta_i = 0.08393$ we previously found in case of black-body (non-selective) solar collectors.

Finally, note that the results presented in (i), (ii) and (v) correspond to solar direct radiation at normal incidence (i.e. Sun zenith angle $\theta_0 = 0$). In the general case the geometric factors in Eqs. (5.3), (5.4) and (5.20) should be multiplied by $\cos^{-1/4} \theta_0$ (Landsberg and Badescu 2000). For oblique radiation incidence ($\theta_0 > 0$) this procedure will yield lower estimates for the upper bound efficiency.

5.3 Carnot Solar Power Plants with External Irreversibility

In Sect. 5.2 we presented simple accurate upper bound formulas for the efficiency of solar power plants operating on Mars. Much more involved models are necessary to further increase the accuracy of results. A solar power plant model based on the Carnot cycle with external irreversibility is presented here. This approach keeps the advantage of generality but also takes into account the irreversibilities associated to the heat transfer at the hot and cold components of the plant.

5.3.1 Solar Power Plant Model

Generally, a solar power plant consists of a solar collector, a thermal engine and a heat storage system. A solar collector mainly consists in mirror, absorber (receiver), transparent cover(s) and thermal insulation. The mirror is missing in case of flat - plate collectors while for some concentration systems the receiver is not protected by transparent cover(s). The solar power plant analysed in this section is designed to operate during the whole Martian year. Consequently, it consists of a flat-plate *selective* solar collector coupled to a thermal engine. No heat storage system is considered in this approach.

5.3.1.1 Solar Collector Model

A selective flat - plate solar collector is shown in Fig. 5.1. Its main components are: the absorber plate (1) with pipes for the working fluid and a transparent cover

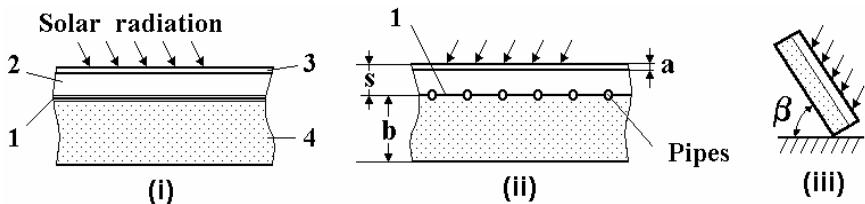


Fig. 5.1 Solar collector. (i) longitudinal section; (ii) transversal section; (iii) lateral view. 1 - absorber plat; 2 - carbon dioxide layer; 3 - transparent cover; 4 - bottom thermal insulation

(3) of thickness a . The gap between (1) and (3) (thickness s) is filled with carbon dioxide at the Martian atmospheric pressure (carbon dioxide represents 95 % of the Martian atmosphere). The bottom thermal insulation (4) has a thickness b . The collector tilt angle is β . One denotes by T_C and T_a the absorber plate and ambient temperatures, respectively. Other temperatures related to collector operation are defined in Table 5.2.

Table 5.2 Various temperatures related to solar collector operation

T_C	absorber plate temperature
T_{si}	temperature of lower surface of transparent cover
T_{ss}	temperature of higher surface of transparent cover
T_a	atmosphere temperature
$T_{C,si}=(T_C+T_{si})/2$	average fluid temperature between absorber plate and lower surface of transparent cover (see Fig. 5.2)
$T_{s,ave}=(T_{si}+T_{ss})/2$	average transparent cover temperature
$T_{ss,a}=(T_{ss}+T_a)/2$	average temperature of the atmosphere boundary layer near the collector transparent cover
$T_{iz}=(T_C+T_a)/2$	average bottom insulation temperature

Direct beam (b), diffuse (d) and ground reflected (r) fluxes of solar energy are incident on the collector. Their magnitude depends upon collector's tilt and orientation. Generally, the incident and absorbed solar energy density fluxes (φ_{inc} and φ_{abs} , respectively) are given by

$$\begin{aligned}\varphi_{inc} &= G_b + G_d + G_r \\ \varphi_{abs} &= (\tau\alpha)_b G_b + (\tau\alpha)_d (G_d + G_r).\end{aligned}\quad (5.21,22)$$

Here G 's denote solar irradiances at collector level while $(\tau\alpha)$'s are effective transmittance - absorptance products.

The collector heat losses are studied by using the thermal resistance method (Eaton and Blum 1975) (see Fig. 5.2 and Table 5.3). Then, the flux of heat losses from the collector towards the ambient is given by:

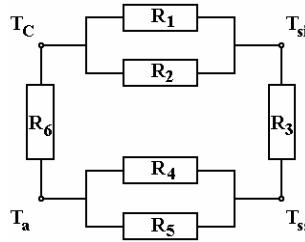


Fig. 5.2 Solar collector thermal resistances. T_a - ambient temperature; T_C - absorber plate temperature; T_{si} - temperature of lower surface of transparent cover; T_{ss} - temperature of upper surface of transparent cover

Table 5.3 Thermal resistances associated to solar collector operation (see Figs. 5.1 and 5.2)

Thermal Resistance	Process
R_1	Convection and conduction between absorber plate (1) and transparent cover (3)
R_2	Radiation between absorber plate (1) and the lower surface of the transparent cover (3)
R_3	Conduction through transparent cover (3)
R_4	Radiation between the higher transparent cover surface (3) and Martian atmosphere
R_5	Convection between the higher transparent cover surface (3) and Martian atmosphere
R_6	Conduction through the bottom thermal insulation(4)

$$q_{tot} = U_L (T_C - T_a) = (T_C - T_a) / R_{tot}, \quad (5.23)$$

where U_L is the overall heat losses coefficient while R_{tot} is the total thermal resistance. U_L is given by (see Fig. 5.2):

$$U_L = \frac{1}{R_{tot}} = \frac{1}{R_{1-5}} + \frac{1}{R_6}, \quad (5.24)$$

where the equivalent resistance R_{1-5} is computed from:

$$R_{1-5} = \frac{R_1 R_2}{R_1 + R_2} + R_3 + \frac{R_4 R_5}{R_4 + R_5}. \quad (5.25)$$

Details concerning the evaluation of thermal resistances follow. The thermal resistance R_1 is computed from

$$R_1 = \frac{s}{\langle \lambda_{CO_2} \rangle N u_{C,si}}, \quad (5.26)$$

Table 5.4 (a) Relationships for convection heat transfer between absorber plate and transparent cover (Eaton and Blum 1975); **(b)** General relationships for convection heat transfer between two plane parallel surfaces (Stefanescu et al. 1982, p. 160)

(a)	$Nu_{C,si} = \left(1 - \frac{\beta}{2\pi}\right) Nu_H + \frac{\beta}{2\pi} Nu_V$
	$Nu_H = 0.195 \cdot Gr^{0.25}$ for $2 \cdot 10^3 \leq Gr < 4 \cdot 10^5$
	$Nu_H = 0.068 \cdot Gr^{0.33}$ for $Gr \geq 4 \cdot 10^5$
	$Nu_V = 0.180 Gr^{0.25} (s/L)^{0.11}$ for $1 \cdot 10^4 \leq Gr < 2 \cdot 10^5$
	$Nu_V = 0.065 Gr^{0.33} (s/L)^{0.11}$ for $2 \cdot 10^5 \leq Gr < 11 \cdot 10^6$
(b)	$Nu_{C,si} = 1$ for $Gr \cdot Pr \leq 1 \cdot 10^3$
	$Nu_{C,si} = 0.105 \cdot (Gr \cdot Pr)^{0.30}$ for $1 \cdot 10^3 < Gr \cdot Pr \leq 1 \cdot 10^6$
	$Nu_{C,si} = 0.400 \cdot (Gr \cdot Pr)^{0.20}$ for $1 \cdot 10^4 \leq Gr < 2 \cdot 10^5$
	$Nu_{C,si} = 0.180 \cdot (Gr \cdot Pr)^{0.25}$ for $Gr \cdot Pr > 1 \cdot 10^{10}$

H and V denote a horizontal and vertical solar collector; β [rad] and L are tilt angle and collector length, respectively.

where $\langle \lambda_{CO_2} \rangle_{C,si}$ is thermal conductivity of carbon dioxide and $Nu_{C,si}$ is Nusselt number, both of them evaluated at the average fluid temperature $T_{C,si}$.

To compute $Nu_{C,si}$, a number of criterial relationships developed for solar collectors are used (Table 5.4a) (Eaton and Blum 1975). For cases not covered by Table 5.4a we use the general relationships from Stefanescu et al. (1982) (Table 5.4b). Prandtl and Grashoff numbers in Table 5.4 were computed with:

$$Pr = \frac{\langle c_{p,CO_2} \rangle_{C,si} \langle \eta_{CO_2} \rangle_{C,si}}{\langle \lambda_{CO_2} \rangle_{C,si}}, \quad Gr = \frac{g_M p_{CO_2}^2 s^3 (T_C - T_{si})}{\langle \eta_{CO_2} \rangle_{C,si} R_{CO_2} T_{C,si}} \quad (5.27)$$

where g_M is gravitational acceleration on Mars while p_{CO_2} and R_{CO_2} are the pressure and gas constant for the carbon dioxide between absorber plate and transparent cover. In Eq. (5.27) the specific heat at constant pressure ($\langle c_{p,CO_2} \rangle_{C,si}$), the dynamic viscosity ($\langle \eta_{CO_2} \rangle_{C,si}$) and the thermal conductivity ($\langle \lambda_{CO_2} \rangle_{C,si}$) of carbon dioxide were interpolated at the average temperature $T_{C,si}$. In computations the thermal properties of carbon dioxide from Pop et al. (1987) were used.

The thermal resistance R_3 is given by:

$$R_3 = a / \lambda_s. \quad (5.29)$$

The temperature dependence of transparent cover's thermal conductivity (λ_s) was described by using an interpolation parabolic function whose coefficients were obtained by using experimental data (the transparent cover is made of crystal; thermal

properties are found in Lide (1991)). The average transparent cover temperature $T_{s,ave}$ is used in computations.

The thermal resistance R_4 is evaluated with:

$$R_4 = \frac{1/\varepsilon_2}{2\sigma(T_{ss}^2 + T_a^2)T_{ss,a}}, \quad (5.30)$$

where $T_{ss,a}$ is the average temperature of the atmosphere boundary layer near the collector transparent cover.

Because of the wind, the heat transfer between the transparent cover and Martian atmosphere is mainly by forced convection. Thus, the thermal resistance R_5 was computed from:

$$R_5 = \frac{L}{<\lambda_{CO_2}>_{ss,a} N u_{ss,a}}, \quad (5.31)$$

where L is collector's length on wind direction and $<\lambda_{CO_2}>_{ss,a}$ is carbon dioxide thermal conductivity at temperature $T_{ss,a}$. In Eq. (5.31) $N u_{ss,a}$ depends on the critical length X_T which separates the laminar and turbulent wind flows over the collector. The relationships of Table 5.5 are used. The Reynolds and Prandtl numbers are evaluated with:

$$Re = \frac{p_{CO_2} W L}{<\eta_{CO_2}>_{ss,a} R_{CO_2} T_{ss,a}}, \quad Pr = \frac{<c_{p,CO_2}>_{ss,a} <\eta_{CO_2}>_{ss,a}}{<\lambda_{CO_2}>_{ss,a}}, \quad (5.32)$$

where W is wind speed and the thermal properties of the atmosphere ($<\eta_{CO_2}>_{ss,a}$, $<c_{p,CO_2}>_{ss,a}$ and $<\lambda_{CO_2}>_{ss,a}$) are interpolated at the average temperature $T_{ss,a}$.

The thermal resistance R_6 is evaluated with:

$$R_6 = b / \lambda_{iz} \quad (5.33)$$

where λ_{iz} is bottom insulation thermal conductivity, computed at the average temperature T_{iz} .

Table 5.5 Relationships for forced convection heat transfer between transparent cover and atmosphere (Stefanescu et al. 1982, p. 126)

$X_T = 5 \cdot 10^5 \frac{<\eta_{CO_2}>_{ss,a} R_{CO_2} T_{ss,a}}{p_{CO_2} W}$	
$N u_{ss,a} = 0.335 \cdot Re^{0.50} \cdot Pr^{0.50}$	for $L \leq X_T$ (laminar flow)
$N u_{ss,a} = 0.036 \cdot Re^{0.80} \cdot Pr^{0.75}$	for $L > X_T$ (turbulent flow)

5.3.1.2 Engine Model

A variety of working fluids were proposed to be used in space solar engines. They include mercury, potassium and rubidium in the early studies (Menetrey 1963) or organic fluids and mixtures of noble gases more recently (Prisnjakov et al. 1991; Angelino and Invernizzi 1993). The later seem to be more appropriate in case of Martian solar power plants. The largest and smallest working fluid temperatures are denoted T' and T'' , respectively (Fig. 5.3).

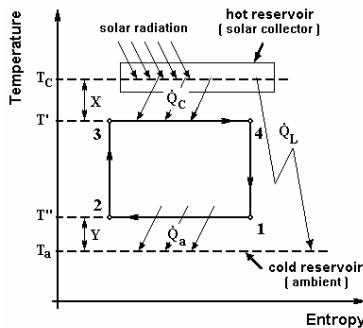


Fig. 5.3 Temperature - entropy diagram for a Carnot cycle with external irreversibility (for notations see text)

Generally, $T_c > T'$ and $T'' > T_a$. Hence, both

$$X \equiv T_c - T' \text{ and } Y \equiv T'' - T_a \quad (5.34)$$

are positive quantities. The flux of solar energy ϕ_{abs} absorbed by the collector is

$$\phi_{abs} = \phi_{abs} A_C, \quad (5.35)$$

where A_C is collector surface area. Part of this flux is transferred to the working fluid (\dot{Q}_c) while the other part constitutes the flux of thermal losses to the ambient (\dot{Q}_L). Then, the solar collector steady-state energy balance is:

$$\phi_{abs} - \dot{Q}_C - \dot{Q}_L = 0. \quad (5.36)$$

The heat fluxes \dot{Q}_C and \dot{Q}_L can be written as:

$$\dot{Q}_C = h_C A_C X \text{ and } \dot{Q}_L = U_L A_C (T_c - T_a), \quad (5.37)$$

where h_C and U_L are appropriately defined overall heat transfer coefficients between the solar collector and the working fluid and ambient, respectively.

The Carnot engine partially converts the flux \dot{Q}_C into power. Thus:

$$\dot{Q}_C = \dot{W} + \dot{Q}_a, \quad (5.38)$$

where \dot{W} is the output power and \dot{Q}_a is the thermal energy flux finally reaching the surroundings. There is no further increase of entropy during work production. Consequently

$$\frac{\dot{Q}_C}{T'} + \frac{\dot{Q}_a}{T''} = 0. \quad (5.39)$$

The heat flux \dot{Q}_a can be written as

$$\dot{Q}_a = h_a A_a Y, \quad (5.40)$$

where A_a and h_a are appropriately defined engine radiator heat transfer area and overall heat transfer coefficient between working fluid and ambient, respectively.

5.3.1.3 Performance Indicators

Four performance indicators may be considered in relation to Martian solar plant operation (Badescu et al. 1999, 2000a, 2000b). First, there is the Carnot engine efficiency defined as:

$$\eta_{engine} \equiv 1 - \frac{T''}{T'}. \quad (5.41)$$

Second, there is the output power \dot{W} , evaluated from:

$$\dot{W} = h_C A_C X - h_a A_a Y. \quad (5.42)$$

Here we used the Eqs. (5.37), (5.38) and (5.40). Third, there is the system efficiency, defined as:

$$\eta_{system} \equiv \frac{\dot{W}}{\phi_{abs} A_C}. \quad (5.43)$$

Finally, the solar (or sun-to-user) efficiency is defined as:

$$\eta_{solar} \equiv \frac{\dot{W}}{\phi_{abs} A_C}. \quad (5.44)$$

Information about some of these indicators is reported in the following. The exergetic efficiency has been not reported for reasons of space.

5.3.1.4 Optimisation

The output power \dot{W} is maximised now. The weight of space power plants should be kept at minimum. Following previous studies (see e.g. Badescu (1994)) we express the mass of the total heat transfer area (i.e. the mass M_C of the solar collector and the mass M_a of the engine radiator) as a k fraction from the total mass M_{tot} of the space system:

$$M_C + M_a = kM_{tot} . \quad (5.45)$$

One denotes d_c and d_a the superficial mass density of collector and radiator, respectively. One can simply write:

$$M_C = d_C A_C , \quad M_a = d_a A_a . \quad (5.46)$$

One can use $d_C = d_a$ as a reasonable assumption. By using Eqs. (5.45) and (5.46) one derives the following constraint to be fulfilled:

$$A_C + A_a = A \equiv k \frac{M_{tot}}{d_a} . \quad (5.47)$$

Other constraints (say F_1 and F_2) are the Eqs. (5.36) and (5.39). By using Eqs. (5.34), (5.37) and (5.40) they become:

$$\begin{aligned} F_1 &\equiv \frac{h_C A_C X}{T_C - X} - \frac{h_a (A - A_C)}{T_a + Y} = 0 \\ F_2 &\equiv \varphi_{abs} - U_L (T_C - T_a) - h_C X = 0 \end{aligned} \quad (5.48,49)$$

The power \dot{W} is considered here as a function of T_C , X , Y and A_C . The Lagrange function L associated to \dot{W} is:

$$L \equiv \dot{W} + \lambda_1 F_1 + \lambda_2 F_2 , \quad (5.50)$$

where λ_1 and λ_2 are multipliers. The maximum output power is obtained by solving the six equations shown in Table 5.6.

Table 5.6 Equations to be solved for power maximisation (for details see text)

-
- | | |
|----|--|
| 1. | $\frac{\partial L}{\partial T_C} = -\lambda_1 \frac{h_C A_C X}{(T_C - X)^2} - \lambda_2 U_L = 0$ |
| 2. | $\frac{\partial L}{\partial X} = h_C A_C + \lambda_1 \frac{h_C A_C T_C}{(T_C - X)^2} - \lambda_2 h_C = 0$ |
| 3. | $\frac{\partial L}{\partial Y} = -h_a (A - A_C) - \lambda_1 \frac{h_a (A - A_C) T_a}{(T_a + Y)^2} = 0$ |
| 4. | $\frac{\partial L}{\partial A_C} = h_C X + h_a Y + \lambda_1 \left(\frac{h_C X}{T_C - X} + \frac{h_a Y}{T_a + Y} \right) = 0$ |
| 5. | $\frac{\partial L}{\partial \lambda_1} = \frac{h_C A_C X}{T_C - X} - \frac{h_a (A - A_C) Y}{T_a + Y} = 0$ |
| 6. | $\frac{\partial L}{\partial \lambda_2} = \varphi_{abs} - U_L (T_C - T_a) - h_C X = 0$ |
-

Table 5.7 Optimum parameters for maximum output power (for details see text)

1.	$T_{C,opt} = \left(1 + f \cdot \frac{1 + f - (1 + f)^{1/2}}{K^{1/2} (1 + K^{1/2})} \frac{h_a}{U_L} \right) T_a$
2.	$X_{opt} = \frac{(1 + f) - (1 + f)^{1/2}}{K^{1/2} (1 + K^{1/2})} \frac{h_a}{h_c} T_a$
3.	$Y_{opt} = \frac{(1 + f)^{1/2} - 1}{1 + K^{1/2}} T_a$
4.	$A_{C,opt} = \frac{K^{1/2}}{1 + K^{1/2}} A$

Table 5.7 shows the solution that applies for $K > 1$. There, the dimensionless parameters K and f are defined as:

$$K = \frac{h_a}{h_c} \left(1 + \frac{h_c}{U_L} \right), \quad f = \frac{\varphi_{abs}}{U_L T_a}. \quad (5.51)$$

Table 5.7 and Eq. (5.42) allow writing the maximum power as:

$$\dot{W}_{\max} = \frac{h_a A}{T_a} Y_{opt}^2 \quad (5.52)$$

The optimum solar efficiency $\eta_{solar,opt}$ can be derived by using Eqs. (5.44), (5.52) and Table 5.7.

5.3.2 Results and Discussions

One of the goals of this chapter is to perform a rough comparison between the performance of PV cell power systems and dynamic solar power plants. In case of PV cell systems we had in mind the Pathfinder's Sojourner. Sojourner is a small (11.5 kg), six - wheel robotic vehicle built at Jet Propulsion Laboratories. Sojourner was landed on Mars aboard the Pathfinder spacecraft on July 4, 1997. In the same sol (i.e. Martian day) she began to traverse the Martian terrain, perform science and technology experiments, and transmit images and data back to the Lander space craft. Sojourner's equipment (computers, motors, radio modem) was mainly powered by a lightweight 0.34 kg solar array of receiving surface area 0.22 m². The PV cell system was designed to provide Sojourner with around 16 W of electric power at noon on Mars. The landing place was in Ares Vallis at 19.17° N and 33.21° W on the U.S. Geological Survey (USGS) cartographic network (Golombek et al. 1997). Pathfinder landed in late northern summer (areocentric longitude $L_s=143^\circ$) and operated for 83 sols.

To allow comparison with Sojourner performance an appropriately designed dynamic power plant equipped with a flat - plate selective solar collector is analyzed here (see Table 5.8). A number of assumptions are accepted as follows. A simple relation is used for the effective transmittance - absorptance products:

Table 5.8 Details about selective flat - plate solar collector design

Quantity	Notation	Value
Number of transparent covers	N	1
Transparent cover thickness	a	0.003 m
Distance between transparent cover and absorber plate	s	0.045 m
Bottom thermal insulation thickness	b	0.1 m
Short wavelengths transparent cover transmittance	τ_{cs}	0.82
Short wavelength transparent cover absorptance	α_{ps}	0.90
Long wavelengths (IR) absorber plate emittance	ε_1	0.10
Long wavelengths (IR) transparent cover emittance	ε_2	0.88
Thermal conductivity of bottom insulation (polyurethane)	λ_b	0.02 W/(mK)

($A_C = 0.22 \text{ m}^2$)

$$(\tau \alpha)_b = (\tau \alpha)_d = \tau_{cs} \alpha_{ps}, \quad (5.53)$$

where τ_{cs} and α_{ps} are the transparent cover transmittance and absorber absorptance, respectively, both of them for short (solar) wavelengths. The collector's transparent cover is made up of crystal (light flint glass). Its thermal conductivity lies between 0.691 W(mK)^{-1} at -100°C and 1.025 W(mK)^{-1} at $+100^\circ\text{C}$ (Lide 1991). The ratio $0.34/11.5=0.0296$ between the PV array weight and total Sojourner weight can be used as a first estimate for the coefficient k appearing in Eq. (5.47). Preliminary tests were performed to determine an optimum collector area $A_{C,opt}$ close to Sojourner PV cell surface area (0.22 m^2). Finally, the value $A=0.23 \text{ m}^2$ was used in computations. It corresponds to a radiator superficial mass density $d_a=1.48 \text{ kg/m}^2$ in Eq. (5.47). It is rather close to the value $d_a=2.3 \text{ kg/m}^2$ accepted in Angelo and Buden (1991) in case of the radiator of a nuclear power satellite but only half of the value $d_a=3 \text{ kg/m}^2$ accepted in Mozjorine et al. (1991) for a solar space power station.

The distance between VL1 and Pathfinder Lander is around 815 km. Thus, one expects quite similar meteorological and actinometric features in both places.

Four strategies of collecting solar energy were considered in preliminary tests (Badescu 1998a). The following two strategies will be used in this section: (i) horizontal collector - strategy H; (ii) the collector tilt and orientation are continuously adjusted to keep the receiving surface perpendicular on Sun's rays - strategy P. Strategy H is easier to use while strategy P gives the higher power output most time of the year.

5.3.2.1 Solar Collector Operation

Solar collector thermal losses mainly depend on the heat transfer regime between absorber plate and transparent cover (separated through the layer 2 of thickness s in Fig. 5.1). We considered the case of a horizontal solar collector. Three design

Table 5.9 Three design solutions for the system solar collector - thermal engine

Case	h_a [W/(m ² K)]	h_c [W/(m ² K)]	Thermal engine performance level
(a)	1	1	Low efficiency
(b)	10	10	Medium efficiency
(c)	500	50	High efficiency

h_a = heat transfer coefficient between the working fluid passing through the radiator of the heat engine and ambient (see Eq. (5.40)); h_c = heat transfer coefficient between solar collector and working fluid (see Eq. (5.37)).

solutions for the system solar collector - thermal engine were analysed (Table 5.9). They correspond to three different performance levels. For analysis we have chosen the middle of a clear sky autumn sol, when the thermal losses are rather high (sol 301 VL1, 14.04 local solar time (LST)). [We remind that sol numbering started when each Viking Lander touched down the Martian soil.]

The dependence of the overall heat loss coefficient U_L on the distance s between absorber plate and transparent cover is shown in Fig. 5.4 for the three cases of Table 5.9. The heat losses decrease by increasing s . This is specific to the conduction heat transfer regime. The thickness s is limited, however, by economic reasons. In the following we use $s=0.045$ m and the high efficiency engine (case (c) in Table 5.9). Note that the chosen value of s is about two times larger than the usual value for Earth placed solar collectors.

For Earth placed solar collectors the heat transfer between absorber plate and transparent cover is mainly by convection. This prompted various technical solutions meant to keep convection at local level (the honeycombs and Francia cells are examples of solutions for convection suppression) (see Meinel and Meinel

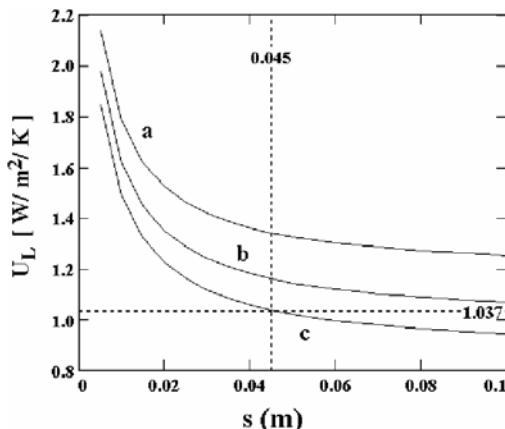


Fig. 5.4 Overall heat loss coefficient U_L as a function of parameter s (distance between absorber plate and transparent cover - see Fig. 5.1). Three design solutions (a), (b), (c) are considered (see Table 5.9). (VL1 site, autumn, sol 301, 14.04 LST)

(1976, p. 404)). The convective heat transfer is activated when the product $Gr \cdot Pr$ exceeds 1000 (see Table 5.4b). We analysed the heat transfer above the absorber plate by using the whole set of available meteorological data (Table 2.2 of Chap. 2). The results show that the product $Gr \cdot Pr$ varies between 9.65 and 211.2, well below the threshold value of 1000. We draw the important conclusion that on Mars the thermal losses between absorber plate and transparent cover are exclusively by conduction.

The convection heat losses from transparent cover to atmosphere are mainly controlled by wind speed. The change from laminar to turbulent convection occurs in case the heat transfer surface exceeds a certain critical length X_T (see Table 5.5). The flow regime above the transparent cover was analysed by using the whole set of available meteorological data. Results are shown in Fig. 5.5. The minimum value of the critical length X_T is 5.43 m. Consequently, the convection heat transfer above the transparent cover is laminar. The critical length X_T is smaller during winter and larger during summer (Badescu et al. 2000a). No significant difference between VL1 and VL2 sites is observed (note that a larger number of meteorological values are available for VL2 - see Table 2.2 of Chap. 2). The critical length obviously decreases when the optical depth increases. During the "clear sky" days ($\tau < 1$) the critical length values have an important dispersion. Scattering diminishes during the dust storm days ($\tau > 1$).

For the design solution we selected (case (c) in Table 5.9), the values of the overall heat loss coefficient U_L range between about 0.5 and 1 $\text{Wm}^{-2}\text{K}^{-1}$ (Fig. 5.6). Broadly, these values are comparable with those of Earth located vacuum solar

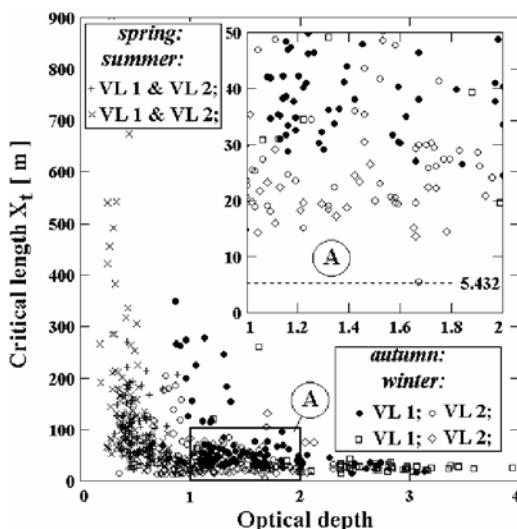


Fig. 5.5 Critical length X_T (see Table 5.5) as a function of optical depth. A high efficiency engine was considered (see Table 5.9). All the available meteorological data were used.

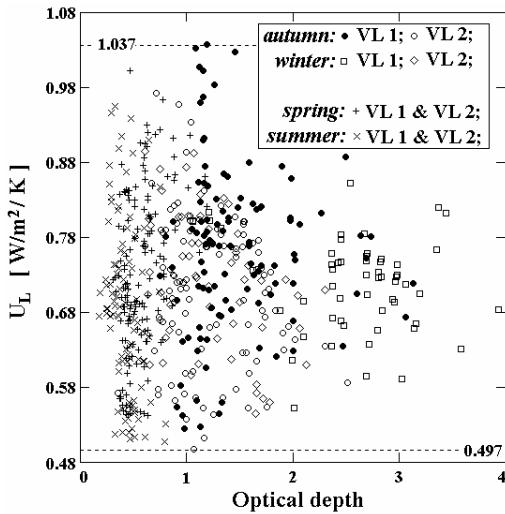


Fig. 5.6 Overall heat loss coefficient U_L as a function of optical depth. A high efficiency engine was considered (see Table 5.9). All the available meteorological data were used.

collectors (see e.g. Meinel and Meinel (1976, p. 387)). The dispersion of the U_L values is slightly smaller during winter and higher during autumn (Badescu et al. 2000a). No obvious dependence of U_L on the optical depth is observed. However, the dispersion of the U_L values is larger for values $\tau < l$, i.e. during the "clear" days. Also, Fig. 5.6 shows that U_L does not depend significantly on the latitude of the solar collector.

5.3.2.2 Thermal Engine Operation

The parameters h_a and h_C have an important influence on performance. For analysis we have chosen the middle of a clear sky autumn day (sol 301 VL1, 14.04 LST). Some results are shown in Fig. 5.7 for the maximum power provided by the thermal engine. The influence of both h_a and h_C is more important on the maximum power provided by the heat engine than on the overall efficiency (Badescu et al. 2000a).

Generally, the influence on performance is strong for low values of both heat transfer coefficients h_a and h_C (less than $100 \text{ W m}^{-2} \text{ K}^{-1}$). The performances are practically not dependent on the higher values of h_C . The influence of h_a on performance is more significant. The strategy of collecting solar radiation has a rather weak influence on the dependence of performance on h_a and h_C . As expected, strategy P leads to higher performance (Badescu et al. 2000a).

The maximum output power could be as high as 12 W and 13 W in case of strategies H and P of collecting solar radiation, respectively (Badescu et al. 2000a). This is lower than the designed power of Sojourner PV arrays (16 W at noon).

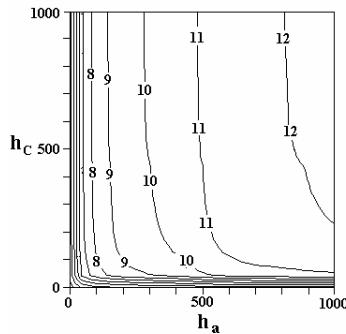


Fig. 5.7 Isolines of maximum power \dot{W}_{\max} provided by the thermal engine as a function of the overall heat transfer coefficients h_a and h_C . Horizontal solar collector (strategy H). Unconcentrated solar radiation was considered (VL1 site, autumn, sol 301, 14.04 LST).

However, this comparison put in disadvantage the solar dynamic system as we used input meteorological data from late autumn at VL1 site while Sojourner was designed to operate during Martian summer. This is a consequence of the fact that none of the 22 complete records available during summer at VL1 site (see Table 2.2 of Chap. 2) corresponds to solar noon.

5.3.2.3 Performance Dependence on Atmospheric Optical Depth

Previous results showed that the dynamic solar power plant performance depends significantly on the heat transfer coefficients h_a and h_C . Practical situations should normally lie between the cases (b) (i.e. medium efficiency engine) and (c) (i.e. high efficiency engine) in Table 5.9. These two cases will be used to provide lower and upper bounds for actual performance. To test the performance of the

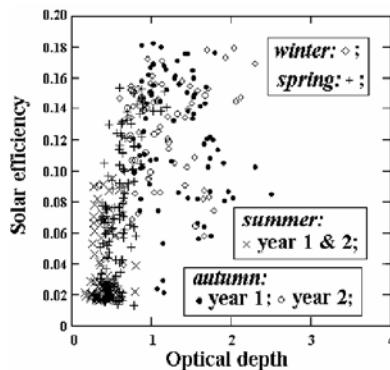


Fig. 5.8 Dependence of solar efficiency $\eta_{\text{solar},\text{opt}}$ on the atmospheric optical depth at VL2 site. High efficiency engine (see Table 5.9) and horizontal solar collector. All the available meteorological data for VL2 site were used.

thermal engine - solar collector combination we simulated its operation at both VL1 and VL2 sites. We analyzed both the horizontal collector (strategy H) and the collector oriented perpendicularly on sun's rays (strategy P), respectively (Badescu et al. 2001a). At VL1 site we used meteorological data from summer, autumn and winter, year 1. All available data were used to simulate the operation at VL2 site. Some results are shown in Fig. 5.8 (for solar energy conversion efficiency) and Fig. 5.9 (for output power).

The influence of latitude on performance is obvious. When the collector is horizontal, the solar energy conversion efficiency is generally smaller at VL1 site as compared to VL2 site (Fig. 5.8). When a P collector is considered, the efficiency increases (at small τ 's) but it still is smaller at VL1 site than at VL2 site (Badescu et al. 2001a). The maximum power provided by a horizontal collector at VL1 site shows some interesting features (Badescu et al. 2001a). Generally, it is smaller than in case of VL2 site. However, there are some situations (for $\tau \approx 1$) when the meteorological effects compensate the latitudinal effects and the output power is quite similar for both VL1 and VL2. If a P collector is used, the power increases for small τ 's but slightly at VL1 site than at VL2 site (Fig. 5.9).

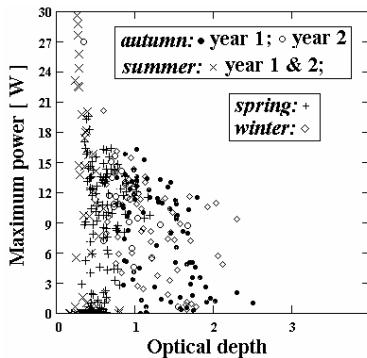


Fig. 5.9 The same as Fig. 5.8 in case of the maximum output power W_{\max} . High efficiency engine and P solar collector.

In case the medium efficiency engine is coupled to a horizontal collector, the efficiency does not exceed 0.13 at VL2 site (Badescu et al. 2001a). Generally, it is lower during summer, when the horizontal orientation is surely a good option (Badescu 1998a). It is higher during winter, when the optical depth is higher, indicating dust storms occurrence. During dust storms the incident solar radiation is mainly diffuse and the horizontal orientation is the best strategy for radiation collection (Badescu 1998a). In case a medium efficiency engine is coupled to a P collector, the maximum efficiency values lie around 0.13 (Badescu et al. 2001a). The efficiency increases at small optical depth values, which generally correspond to summer. The centre of the output data cloud is placed around 0.09-0.10. The P strategy does not diminish the performance at larger optical depth.

For a horizontal collector attached to a high efficiency engine, the solar efficiency is as high as 0.18 (Fig. 5.8). The qualitative features pointed out previously maintain. The lower sun-to-user efficiency corresponds to lower optical depth. If a P collector is considered, the solar efficiency increases during summer and spring but does not significantly exceed 0.18 (Badescu et al. 2001a). The centre of the output data cloud is placed around 0.15, higher than in case of the horizontal collector. Again, the P strategy does not diminish the performance. This strategy keeps constant the solar efficiency at higher optical depths but improves the performance as compared to the H strategy at smaller optical depths.

The power provided by a system consisting in a horizontal collector and a medium efficiency engine does not exceed 7 W (Badescu et al. 2001a). Generally, it is lower for small optical depth. Indeed, the flux of direct solar energy is higher at smaller optical depth and the horizontal orientation is not recommended in this case. At higher optical depth direct solar radiation diminishes and diffuse radiation increases. As a result, the horizontal orientation is close to optimum. This explains the higher values of \dot{W}_{\max} around $\tau = 1$. For further increase of τ the diffuse radiation flux diminishes. Consequently, \dot{W}_{\max} decreases in this case. If a P collector is considered, the solar efficiency is higher during summer and spring, when the optical depth is small (Badescu et al. 2001a). Generally, the efficiency increases as compared to the H strategy. At high optical depth the solar efficiency is comparable for both H and P strategies.

Using a better engine with a horizontal collector leads to output power up to 13 W for τ around 1 (during sols of spring, autumn and winter) (Badescu et al. 2001a). For small values of τ the output power is smaller. The P strategy is recommended in combination with high efficiency engines (Fig. 5.9). In this case the sun-to-user efficiency during sols with small optical depth increases strongly, exceeding 25 W. The centre of the output data cloud is placed around 13 W. The P strategy can be used at high optical depth, too.

5.3.2.4 Performance Dependence on the Level of Solar Irradiance

The dependence of solar efficiency on the level of incident global irradiance was studied in (Badescu et al. 2001a). In case of a horizontal collector operating at VL2 site, a slightly non-linear relationship exists between solar efficiency and incident global irradiance for both the medium and the high efficiency engines. When the P collector is considered, the dependence is more complicated.

When a combination medium efficiency engine - horizontal collector is considered, the efficiency is higher (whatever the incident irradiance is) during autumn and winter and lower during summer. This rather surprising fact is valid for the P collector, too. It is probably due to the influence of the ambient temperature, which is lower during autumn and winter. The situation is different in case of the high efficiency engine. Then, the solar efficiency is higher during summer, even if it corresponds to lower incident irradiance. This remark is even more obvious when a P collector is considered. A more uniform solar efficiency values distribution is obtained in this case.

In case of a horizontal collector, the efficiency increases by increasing the irradiance level, whatever the type of engine is. For the P collector, the efficiency depends on the irradiance level in a more interesting manner. Thus, the lower bound of the efficiency increases with the irradiance level, but the upper bound practically does not depend on the irradiance level. This is valid for both the medium and high efficiency engines.

During autumn and winter the dependence of η_{solar} on the level of incident global irradiance is nearly similar for both the horizontal and the P collector operating at VL1 site. Important differences exist during summer. The P collector is more effective, as expected. The influence of irradiance level on solar efficiency decreases drastically beyond a certain threshold value (about 300 W/m^2). This applies to both types of solar collectors.

Figure 5.10 shows some results for the output power at VL1 site. Meteorological data from year 1 were used. In case of a horizontal collector the dependence of the maximum output power on the level of global irradiance is not differentiated upon season (Badescu et al. 2001a). When a P collector is considered, the same incident irradiance value leads to an output power obviously higher during summer. For a horizontal collector the output power increases when the input irradiance increases (Badescu et al. 2001a). This remark maintains for the P collector except for summer results (Fig. 5.10).

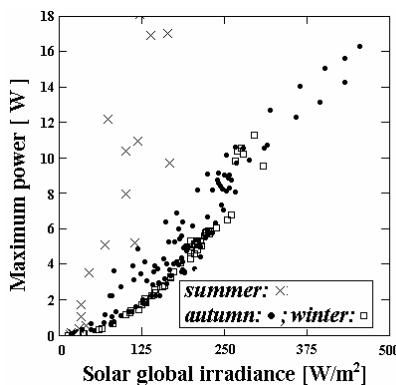


Fig. 5.10 Dependence of maximum output power \dot{W}_{max} on global irradiance at VL1 site. High efficiency engine (see Table 5.9) and P solar collector. Meteorological data for summer, autumn and winter year 1 were used.

The power provided by a system based on a horizontal collector increases at VL2 site with the level of solar irradiance, as expected (Badescu et al. 2001a). The increase is almost proportional in case of the medium efficiency engine and slightly non-linear for the high efficiency engine. Generally, there is no obvious dependence on season. In case of the P collector, the output power obviously depends on season, with a maximum during spring and summer and a minimum during autumn and winter. This is true for both the medium and high efficiency engines (Fig. 5.11).

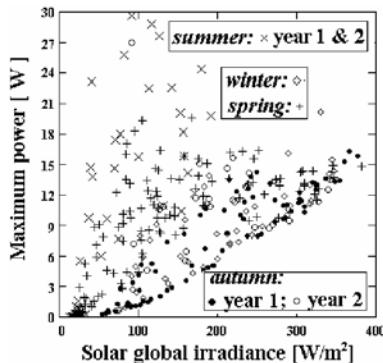


Fig. 5.11 Dependence of the maximum output power \dot{W}_{\max} on the level of solar global irradiance incident on a horizontal surface at VL2 site. High efficiency engine (see Table 5.9) and P solar collector. All the available meteorological data for VL2 site were used.

The seasonal influence is similar for both solar energy conversion efficiency and output power in case of a medium efficiency engine - horizontal collector combination (Badescu et al. 2001a). This is still valid when the collector is connected to a high efficiency engine. However, in this case the power values have a larger dispersion, which increases slightly with increasing the irradiance.

In case of a medium efficiency engine and a P collector, the influence of the season is stronger on power than on solar efficiency (Badescu et al. 2001a). This is even more obvious if one connects a P collector to a high efficiency engine. The high solar efficiency values during spring and autumn at smaller irradiance values are not the main cause of the higher output power (which is obtained in summer (Fig. 5.11)). For a medium efficiency engine, a P collector is recommended mainly during summer and spring (which are dust storm - free seasons) (Badescu et al. 2001a). One finds higher output power values as compared to the case when a horizontal collector is used. If the summer and spring values would be neglected, one can see that choosing a P instead of an H collector leads to a relatively small increase in power during the other two seasons (autumn and winter). This is valid for the high efficiency engine, too (Fig. 5.11). However, the increase in power by using a P collector during seasons without dust storms is more spectacular. Compare the power values as high as 20 W obtained at lower input irradiance by using a P collector (Fig. 5.11) with the power values of at most 12-13 W obtained at high irradiance values by using a horizontal collector (Badescu et al. 2001a).

5.3.2.5 Diurnal Variation of Performances

The input data file contains only a few sols with complete records covering appropriately the whole daylight time. The most part of the sols has one to three records only associated to time periods close to sunrise, noon or sunset. In particular, no summer sol with complete records covering the whole day was found either for VL1 or for the VL2 site. For each season, we selected those sols that are richer in

available data. In case of VL1 these sols are 301 (autumn) and 328 (winter). Note that no available data refers to spring (see Table 2.2 of Chap. 2). In case of VL2 the sols selected are 420 (spring) and 872 (autumn). The sol 406 (winter) is also available but no reference to it will be made here, as the winter conditions are relatively similar for both Viking Lander sites.

Various efficiencies defined in Section 5.1.3 were reported in Badescu et al. (2000b) in case of sol 301 (Martian autumn) at VL1 site. During most of the time the optimum engine efficiency $\eta_{\text{engine,} \text{opt}}$ lies around 33 %. The efficiency is decreasing suddenly near sunset. Generally, the optimum engine efficiency is less dependent on the type of engine and on the strategy of collecting solar radiation. In exchange, the system efficiency $\eta_{\text{system,} \text{opt}}$ is strongly dependent on engine quality and less dependent on the strategy of solar collection. It lies around 0.2 for the high efficiency engine and around 0.15 for the medium efficiency engine. When the high efficiency engine is considered, the decrease in performance is more abrupt near sunset.

The optimum solar efficiency $\eta_{\text{solar,} \text{opt}}$ is smaller than the optimum system efficiency (Fig. 5.12) but exhibits the same qualitative features. Generally, it doesn't exceed 0.15 for the high efficiency engine and 0.10 for the medium efficiency engine. The strategy of collecting solar radiation has not a strong influence on $\eta_{\text{solar,} \text{opt}}$. In exchange, the strategy significantly influences the maximum output power \dot{W}_{max} (Fig. 5.13). If the optimum solar efficiency is almost constant during the day, the maximum power's time variation is closely related to the temporal variation of solar irradiance. This is more obvious for the high efficiency engine. The maximum output power can exceed 15 W for the high efficiency engine and doesn't exceed 8 W for the medium efficiency engine. This proves that properly designed thermal power plants are comparable in performance with PV-based power systems (one reminds that Sojourner's design output power is 16 W at solar noon). Moreover, the previous analysis puts in disadvantage the solar dynamic

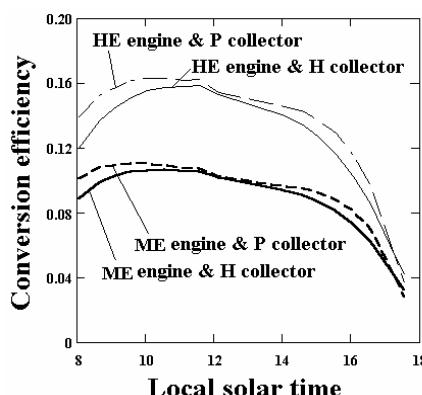


Fig. 5.12 Optimum solar efficiency $\eta_{\text{solar,} \text{opt}}$ during sol 301 (autumn) at VL1 site. Medium efficiency engine (ME) and high efficiency engine (HE) (see Table 5.9). Strategy H - horizontal collector; Strategy P - collector permanently kept perpendicular on Sun's rays.

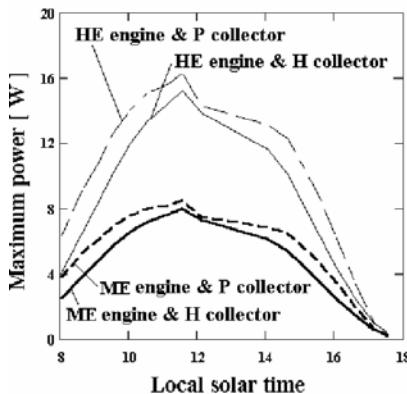


Fig. 5.13 The maximum output power \dot{W}_{\max} during sol 301 (autumn) at VL1 site. For details see Fig. 5.12.

system as we used here input meteorological data from late autumn at VL1 site while Sojourner was designed to operate during Martian summer.

The hourly variation of solar collector's optimum temperature is shown in Fig. 5.14. It is less dependent on the strategy of collecting solar radiation and a little more dependent on engine quality. The high efficiency engine leads to a smaller optimum collector temperature. Generally, the collector temperature exceeds with more than 100 K the ambient temperature. The hourly variation of $T_{C,\text{opt}}$ is stronger than that of the ambient temperature. This proves that $T_{C,\text{opt}}$ is mainly controlled by the level of incident solar radiation.

The optimum temperature of the working fluid in the cold part of the engine (T_{opt}) is well correlated with the time variation of the ambient temperature (Badescu et al. 2000b). T_{opt} is smaller in case of the high efficiency engine, when the influence of the strategy of collecting solar radiation is obvious. This influence is less

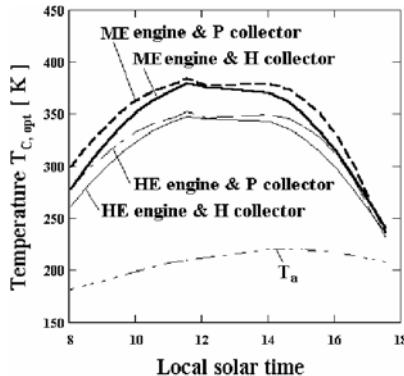


Fig. 5.14 Optimum solar collector temperature $T_{C,\text{opt}}$ associated to solar thermal power plant operation during sol 301 (autumn) at VL1 site. For details concerning the power plant see Fig. 5.12.

important in case of the medium efficiency engine. In exchange, the optimum temperature of the working fluid in the worm part of the engine (T''_{opt}) is better correlated with the hourly variation of the solar radiation (Badescu et al. 2000b). Both the engine quality and the strategy of collecting solar radiation are important factors. Generally, T''_{opt} is close to $T_{C,\text{opt}}$.

Sol 328 belongs to winter at VL1 site. The main performance indicators of the solar thermal power plant are studied in Badescu et al. (2000b). The optimum solar efficiency values are rather close to the autumn values. The quality of the engine is very important and the strategy of collecting solar radiation is less important, as far as their influence on the performance indicators is considered. This is due to the fact that sol 328 belongs to a dust storm period (Badescu 1998a). The maximum output power is, however, much smaller than during autumn. At noon it reaches 10 W in case of the high efficiency thermal engine but doesn't exceed 5 W in case of the medium efficiency engine. The maximum output power time variation is stronger than that of solar efficiency but this is less obvious than in autumn. The daily variation of the optimum collector temperature is rather weak. Generally, the high efficiency engine leads to a lower optimum collector temperature. Also, the optimum temperature is lower for a collector perpendicular on Sun's rays.

Sol 420 belongs to springtime at VL2 site. In this case, the optimum solar efficiency depends significantly both on engine quality and on the strategy of collecting solar radiation (Badescu et al. 2000b). In case of a P collector, the optimum solar efficiency is nearly constant during the day. It varies rather much for the horizontal collector (with a maximum at noon). The optimum solar efficiency values are quite close to those from autumn and summer at VL1 site. The maximum output power shows features similar to those of the optimum solar efficiency. Generally, the P collector assures a more constant maximum output power during the day. The engine quality is very important. The high efficiency engine connected to a P collector can ensure a maximum output power around 15 W during most part of the day. This is close to the design power of the PV system of Mars Pathfinder's Sojourner. When a P solar collector is considered, the optimum collector temperature shows a remarkable constancy during the day.

Sol 872 corresponds to autumn at VL2 site. The optimum solar efficiency shows an interesting variation during the day. This is more obvious for a P collector (Badescu et al. 2000b). In case of the horizontal collector, the optimum solar efficiency has a maximum around the noon. The optimum solar efficiency is higher in case of the high efficiency engine, as expected. It ranges from 8 to 16 %. The maximum output power shows similar features. Generally, the better performance in the morning corresponds to the stronger solar irradiance during that part of day (Badescu 1998a). Note that a horizontal collector assures a constant power during the day. The high efficiency engine should be used in combination with a P collector. A maximum output power of 15 W can be provided only a short time period during the day. In case of a P collector, the optimum temperature is rather constant until 14.00 LST. When a horizontal collector is considered, the optimum temperature has a maximum reached between 12.00 and 14.00 LST. Generally, $T_{C,\text{opt}}$ is lower than in the cases above and this can be correlated to the smaller ambient temperature value.

5.3.2.6 Correlation between Power and Conversion Efficiency

The dependence of the maximum output power on the optimum solar efficiency at VL1 site is shown in Badescu et al. (2000b). Meteorological data from year 1 were used there. Let us consider the case of the medium efficiency engine. The maximum output power increases when the optimum solar efficiency increases, as expected. There is little dependence on season, except for a few summer values in case of the P collector, when the solar efficiency is obviously higher. Broadly speaking, there is little difference between the maximum output power provided by systems based on H and P collectors, respectively. The highest optimum solar efficiency is 0.11 and the maximum output power doesn't exceed 10 W. When the high efficiency engine is considered, the influence of optimum solar efficiency on the maximum output power is important. The highest optimum solar efficiency is 0.16 and the output power can be as high as 15 W for an H collector and 17 W for a P solar collector.

Figure 5.15 shows some results obtained at VL2 site (Badescu et al. 2000b). All the available meteorological data for the two years of VL2 operation were used in computations. Generally, the dependence of the maximum output power on the optimum solar efficiency is non-linear. In case of a horizontal collector, the dependence seems to indicate a quadratic dependence ($\dot{W}_{\max} \propto \eta_{solar,opt}^2$). The coefficient of proportionality is smaller in case of the medium efficiency engine (Badescu et al. 2000b). When the P collector is considered, the dependence $\dot{W}_{\max}(\eta_{solar,opt})$ is more complicate, but keeps the quadratic feature. The P collector has a more compact dispersion of the optimum solar efficiency values. It ranges from 8 to 12 % in case of the medium efficiency engine and between 13 and 19 % in case of the high efficiency engine.

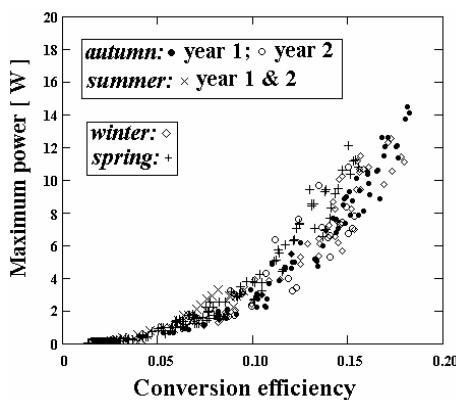


Fig. 5.15 Dependence of the maximum output power \dot{W}_{\max} on the optimum solar efficiency $\eta_{solar,opt}$ at VL2 site. High efficiency engine (see Table 5.9) and horizontal solar collector. Meteorological data for summer, autumn and winter (year 1) and spring, summer and autumn (year 2) were used.

In all cases a certain optimum solar efficiency threshold (around 5 %) must be exceeded in order the system provide useful power. The dependence of the maximum output power on optimum solar efficiency is stronger in case of the high efficiency thermal engine than in case of the medium efficiency engine, on one hand, and in case of the P collector than for the horizontal collector, on the other one hand. It seems no obvious difference exists between the performances of the power plant in the two years of VL2 operation.

5.4 Solar Stirling Power Engine

A more realistic solar thermal engine is considered in this section. It consists of a solar collector - Stirling engine combination. The Stirling engine is indeed very attractive due to its operation at low temperature difference between the two heat reservoirs. This could enable power production during the Martian dust storm period, when both the incoming solar energy flux and collector temperature are small. Preliminary results were briefly reported in Badescu et al. (2001b) where the diurnal solar plant operation is studied. One concluded that the efficiency of a usual (Earth based designed) Stirling solar engine diminishes with about 8 % when operates under Martian weather conditions. However, the recommended volumetric ratio lies between 1 and 2.6 and the best thermal agent is helium, in agreement with conclusions based on Earth made experiments. The model previously proposed in Badescu et al. (2001b) is now presented. A number of improvements are also included (Badescu 2004).

5.4.1 Solar Engine Model

The solar engine consists of a selective flat-plate solar collector coupled to a Stirling engine with partial heat regeneration. No heat storage system is considered in this approach.

5.4.1.1 Solar Collector Model

First, we shall consider the flat-plate solar collector (Fig. 5.16). Its main components are the collector plate and a transparent cover of thickness a . The gap of thickness s between these two components is filled with carbon dioxide at Martian atmospheric pressure. Solar direct and diffuse radiation penetrates the transparent cover and is absorbed by the collector plate. The thermal energy generated in the solar collector plate is transferred to the thermal agent at the hot head of the Stirling engine and later-on part of it is converted into work.

Steady state energy balance for the solar collector yields

$$\dot{Q}_{abs} - \dot{Q}' - \dot{Q}_{lost} = 0. \quad (5.54)$$

Here \dot{Q}_{abs} is the thermal energy flux generated into solar collector plate, \dot{Q}_{lost} is the heat flux lost towards the ambient and \dot{Q}' is the heat flux transferred to the

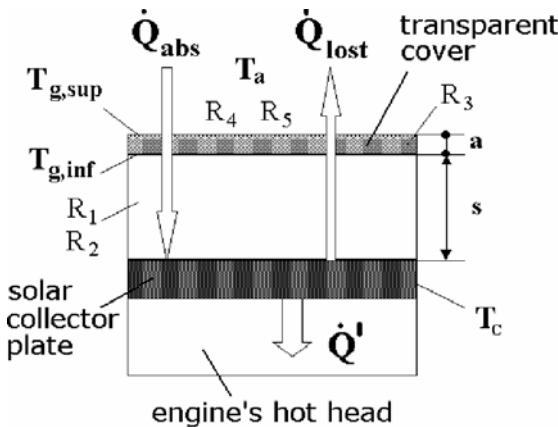


Fig. 5.16 The solar collector of the Stirling engine

thermal fluid inside the Stirling engine. The thermal energy flux generated into the solar collector plate is given by

$$\dot{Q}_{abs} = \varphi_{abs} A_C, \quad (5.55)$$

where A_C is the area of collector plate surface. The heat flux lost towards the ambient is given by the usual Newton relationship:

$$\dot{Q}_{lost} = U_L A_C (T_C - T_a), \quad (5.56)$$

where U_L is the overall heat loss coefficient, while T_C and T_a are collector and ambient temperature, respectively. Note that T_C is an averaged value over collector surface and normally is not constant in time.

Seven temperatures are associated to solar collector operation (Badescu 2004). With a single except (i.e. the average temperature of bottom insulation) they have the meaning explained in Tables 5.2 and 5.3. The collector heat losses could be studied by using the thermal resistance method. Then, the flux of heat losses \dot{q}_{lost} per unit collector surface area is defined by:

$$\dot{q}_{lost} \equiv \frac{\dot{Q}_{lost}}{A_c} = U_L (T_C - T_a) = \frac{T_C - T_a}{R_{tot}}, \quad (5.57)$$

where R_{tot} is the total thermal resistance given by

$$R_{tot} = \frac{R_1 R_2}{R_1 + R_2} + R_3 + \frac{R_4 R_5}{R_4 + R_5} \quad (5.58)$$

Note that R_{tot} here is equivalent to R_{1-5} in the model of solar collector of section 5.1.1. Therefore, Fig. 5.2 and Table 5.3 show the meaning of the thermal resistances R_1 to R_5 . Details concerning the computation of these thermal resistances are similar to those presented in Section 5.1.1 and will be not repeated here.

5.4.1.2 Stirling Engine Model

During the 1980s research was started on low temperature differential Stirling engines. A first Stirling engine running with a temperature difference below 20 degrees Celsius has been developed (Senft 1996). Meanwhile a rather large number of research or commercial Stirling engines using unconcentrated solar radiation were built, with powers from a few to tens of Watt (Badescu 2004).

A simple Stirling engine model is used at this stage of the analysis. It was first developed in Howell and Bannerot (1977) and slightly improved in Badescu (1992). The model is based on two simplifying assumptions. First, one neglects the heat losses during the thermal transfer from the solar collector to the working fluid inside the engine. Second, one neglects the friction of the moving parts. These make of course the results reported below to be upper bounds for the performance of a real solar Stirling system. The Stirling cycle efficiency η_{Stirling} is defined as:

$$\eta_{\text{Stirling}} = \frac{P}{\dot{Q}'} = \frac{\eta_{\max}}{1 + D\eta_{\max}}, \quad (5.59)$$

where P is the output mechanical power while $\eta_{\max} = 1 - T_a/T_c$ is the maximum efficiency of the cycle in case of perfect heat regeneration. In case the thermal agent is an ideal gas the parameter D is given by Howell and Bannerot (1977), Badescu (1992):

$$D = \frac{x}{(k - 1) \ln \varepsilon_v}. \quad (5.60)$$

Here $x (\in [0,1])$ is the heat regeneration factor, k is the adiabatic exponent of the thermal agent and ε_v is the volumic ratio (i.e. the ratio of the extreme values of the thermal agent total volume). $x = 0$ means ideal heat regeneration. In this case the efficiency of the Stirling engine Eq. (5.59) equals the Carnot efficiency.

Use of Eqs. (5.54), (5.55), (5.56) and (5.59) yields the following dependence of the power provided P on the collector plate temperature T_c :

$$P(T_c) = \frac{1 - T_a/T_c}{1 + D(1 - T_a/T_c)} [\varphi_{abs} - U_L(T_c - T_c)] A_C. \quad (5.61)$$

Note that U_L is not a constant but depends on T_c (among other variables) (see Fig. 5.17). Consequently, a simple numerical procedure to maximize directly the power P given by Eq. (5.61) was used here, as follows. One monotonously increases the plate temperature T_c starting from the ambient temperature. For any value of T_c one evaluates first the overall heat loss coefficient U_L (Eq. (5.57)) and then the output power P (Eq. (5.61)). The maximum power P_{\max} and the appropriate optimum plate temperature $T_{c,\text{opt}}$ are finally selected from the series of results.

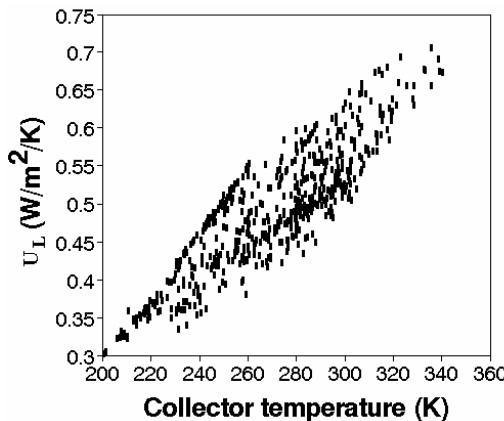


Fig. 5.17 Dependence of overall heat loss coefficient U_L on collector temperature. All available meteorological data were used (see Table 2.2 of Chap. 2).

Figure 5.17 shows that the overall heat loss coefficient U_L increases by increasing the collector temperature. The values of U_L range between 0.3 and $0.7 \text{ W m}^{-2}\text{K}^{-1}$. Broadly, these values are smaller than those of Earth located flat-plate vacuum solar collectors (see e.g. Benz and Beikircher (1999) where the experimentally derived overall heat loss coefficient is about $1.25 \text{ W m}^{-2}\text{K}^{-1}$ during a rather heavy utilization as steam production). They are slightly smaller than the results obtained during simulation of the Carnot solar engine in section 5.3 (U_L values between 0.5 and $1 \text{ W m}^{-2}\text{K}^{-1}$). This is to be expected as an (ideal) Carnot engine normally operates at higher collector temperature than the Stirling engine considered here. The dependence of U_L on collector temperature justifies a posteriori the numerical optimization procedure we used in this section.

The way of using the output power of the Stirling engine is not relevant for this study as the key aspect here is the comparison with the output power of a similarly sized PV-based system. Consequently, electric generators and/or gear boxes are not considered, neither efficiencies for these (possible) components are included in the model.

Use of the output power P allows to define the solar energy conversion efficiency (or solar-to-user efficiency) as $\eta \equiv P / (\varphi_{\text{inc}} A_C)$. The maximum solar efficiency is given by

$$\eta_{\text{sol,max}} = P_{\text{max}} / (\varphi_{\text{inc}} A_C). \quad (5.62)$$

Note that in the case analysed here both η_{sol} and P have their maximum for the same value $T_{C,\text{opt}}$.

5.4.2 Results and Discussions

A horizontal solar collector is considered. This is the easiest-to-use strategy. It gives for many periods of time a performance less than 10 % smaller than the

strategies involving more sophisticated, orientable, solar collectors (Badescu et al. 2000a, 2000b, 2001a). Details about solar collector design are given in Table 5.8 (where the information about the bottom insulation should be omitted of course). The values of the effective transmittance-absorptance products were computed with $(\tau\alpha)_{dir} = (\tau\alpha)_{dif} = \tau\alpha$.

Three Stirling engine thermal agents are usually considered: carbon dioxide (adiabatic exponent $k = 1.33$), air ($k = 1.4$) and helium ($k = 1.66$). Helium gives the best performance and is our choice for this section. A simple analysis proves that the Stirling engine efficiency increases by increasing the compression ratio ε_V . Here the value $\varepsilon_V = 2$ is adopted as a good compromise between better performance and smaller engine size. Increasing the heat regeneration parameter x from the ideal value $x = 0$ diminishes the engine output power. Here the value $x = 0.2$ is accepted, in agreement with current terrestrial practice.

Results of computations performed by using the meteorological data during the autumn and winter of the first year at VL1 site are shown in Fig. 5.18. It is obvious that both the maximum power P_{max} and the maximum solar efficiency $\eta_{sol,max}$ increase around the solar noon. However, $\eta_{sol,max}$ seems to be less dependent on solar time than P_{max} . The maximum power of the solar Stirling engine around the noon is comparable in magnitude with the designed power of Sojourner's PV cells system (i.e. 16 W). Also, the solar efficiency is as large as 18 %, which is close to the design value of Sojourner's PV cells efficiency (about 18 %). Note, however, that the horizontally placed collector is a good solar energy collection strategy especially during autumn and winter. Indeed, results not shown here prove that during spring and summer the solar Stirling engine performance is slightly worse. During the warm season other strategies of collecting solar energy are much more effective (Badescu 1998a).

The large dispersion of the results in Fig. 5.18 is mainly due to the influence of optical depth. The performances of solar Stirling engine obviously decrease by

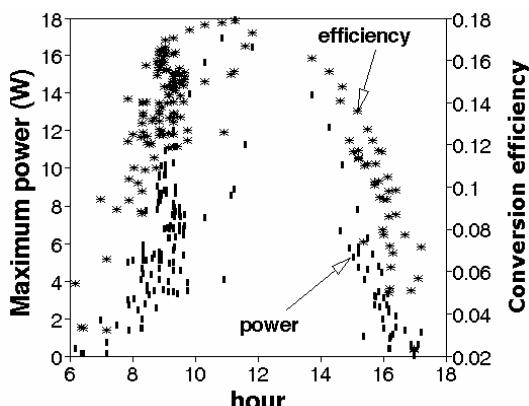


Fig. 5.18 Dependence of maximum power P_{max} and maximum solar efficiency $\eta_{sol,max}$ on local solar time (in Earth hours). Results obtained by using the meteorological data during the autumn and winter in the first year at VL1 site.

increasing the atmospheric optical depth (Badescu 2004). The maximum efficiency $\eta_{sol,max}$ is less dependent on the optical depth than the maximum power P_{max} . Note that under-unitary optical depth values are usually associated with "clear sky" conditions while during the dust storm the optical depth has larger values (Badescu 2001).

The dependence of the maximum output power P_{max} on optimum solar efficiency $\eta_{sol,max}$ is shown in Fig. 5.19. All the meteorological data of Table 2.2 in Chap. 2 were used. The maximum output power increases when the optimum solar efficiency increases. Generally, the dependence of P_{max} on $\eta_{sol,max}$ seems to indicate a quadratic relationship. This is in agreement with previous results obtained in case of solar Carnot engines.

A certain optimum solar efficiency threshold (around 5 %) must be exceeded in order the system provide useful power. Again, this confirms previous results reported in section 5.3. The optimum solar collector temperature $T_{C,opt}$ increases by increasing the solar efficiency, as expected. The larger dispersion of the $T_{C,opt}$ values for the same value of the solar efficiency is mainly due to the difference in the solar time.

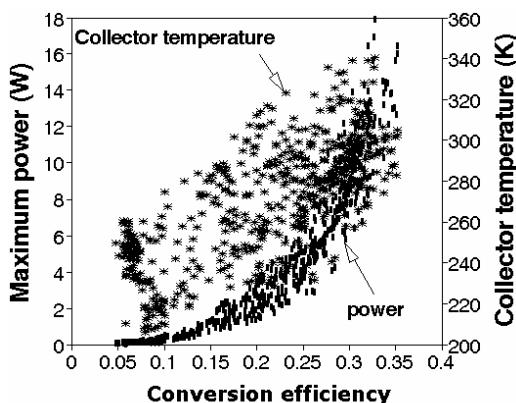


Fig. 5.19 Dependence of the maximum power P_{max} and optimum collector temperature $T_{C,opt}$ on maximum solar efficiency $\eta_{sol,max}$. All available meteorological data were used (see Table 2.2 of Chap. 2).

5.5 Conclusions

The chapter focuses on solar thermal power generation on Mars surface. A "dynamic" solar power plant (which consists in a solar collector - thermal engine combination) is proposed as an alternative for the more usual photovoltaic cells.

Upper bounds for the efficiency of solar thermal power plants operating in the Martian environment are evaluated in section 5.2. The earlier theories predict too high efficiencies for radiation energy conversion into work to be of practical interest. Much more accurate simple upper bounds were already proposed in literature. For reader convenience they are summarized in section 5.2 and applied to power plants

operating on Mars under both direct and diffuse solar illumination. Both black-body and selective absorbers were considered.

A more elaborated but still idealized model uses a Carnot cycle to describe solar engine operation in section 5.3. The solar power plant is designed to operate during the whole Martian year. Consequently, it contains a *selective flat - plate* solar collector. A detailed model of collector heat losses towards the ambient is developed. The optimization procedure developed in section 5.3.1 is based of finite-time thermodynamics methods. The following two strategies of collecting solar energy were analysed: (i) horizontal collector - strategy H; (ii) the collector tilt and orientation are continuously adjusted to keep the receiving surface perpendicular on Sun's rays - strategy P. The first strategy is easier to implement and use in practice while the later strategy provides higher collected solar energy most time of the year. A medium and a high efficiency thermal engine were analysed. Practical situations should normally lie between these two cases. All the computations were performed for a solar collector comparable in size with Pathfinder's Sojourner (solar energy collection area of 0.22 m^2). During the application we used as input data values of atmospheric pressure and temperature, wind speed and atmospheric optical depth measured at Viking Landers sites. The main results are:

(1) The heat losses between the solar collector absorber plate and transparent cover are exclusively by conduction. This differs from the Earth-based solar collector where the main heat loss mechanism is convection.

(2) The convection heat transfer above the transparent cover is laminar. This differs from the Earth-based solar collectors where turbulence is the main convection mechanism.

(3) For the design solution we selected, the values of the overall heat loss coefficient U_L range between about 0.5 and $1 \text{ Wm}^{-2}\text{K}^{-1}$. Roughly, these values are comparable with those of Earth located vacuum solar collectors. The dispersion of the U_L values is slightly smaller during winter and higher during autumn. No obvious dependence of U_L on the optical depth is observed. However, the dispersion of the U_L values is larger for values during the "clear" days. U_L does not depend significantly on the latitude of the solar collector.

(4) Two parameters are used to quantify the heat transfer inside the thermal engine. They are h_a (the heat transfer coefficient between the working fluid passing through the radiator of the heat engine and ambient; see Eq. (5.40)) and h_C (the heat transfer coefficient between solar collector and working fluid (see Eq. (5.37)). The influence of h_a on performance is more significant.

(5) The strategy of collecting solar radiation has a rather weak influence on the dependence of performances on h_a and h_C . As expected, strategy P leads to higher performances.

(6) The dynamic solar power plants equipped with selective flat - plate collectors could provide power comparable to that of similar-size PV cell systems.

(7) The influence of latitude on performance is obvious. Generally, the solar efficiency is smaller at VL1 site as compared to VL2 site for both strategies of collecting solar radiation. In most cases the maximum power provided by a horizontal collector at VL1 site is smaller than at VL2 site. However, in some

situations the meteorological effects compensate the latitudinal effects and the output power is quite similar at both VL1 and VL2 sites.

(8) In case a medium efficiency engine is coupled to a horizontal collector, the solar efficiency does not exceed 0.13 at VL2 site. Generally, it is lower during summer and higher during winter's dust storms. In case the medium efficiency engine is coupled to a P collector, the solar efficiency increases during summer.

(9) The solar efficiency is as high as 0.18 in case of a horizontal collector attached to a high efficiency engine. If the high efficiency thermal engine is connected to a P collector the solar efficiency increases significantly during summer and spring but does not exceed 0.18.

(10) Using a high efficiency engine coupled to a horizontal collector leads to output power up to 13 W during spring, autumn and winter.

(11) The P strategy is recommended mainly during summer and spring in combination with high efficiency engines. In this case the output power could be as high as 25 W.

(12) During an autumn day at VL1 site (sol 301) the optimum engine efficiency is around 33 %. The system efficiency lies around 20 % for the high efficiency engine and around 15 % for the medium efficiency engine. The optimum solar efficiency doesn't exceed 15 % and 10 %, respectively, in the two cases. The maximum output power can exceed 15 W and doesn't exceed 8 W for the high and medium efficiency engine, respectively.

(13) During a winter dust-storm day at VL1 site (sol 328) the maximum output power is much smaller than during autumn. At noon it reaches 10 W in case of the high efficiency thermal engine but doesn't exceed 5 W in case of the medium efficiency engine.

(14) During a spring day at VL2 site (sol 420) the optimum solar efficiency values are quite close to those from autumn and summer. The high efficiency engine connected to a solar collector kept perpendicular on Sun's rays can ensure a maximum output power around 15 W during most part of the day. This is close to the design power of the PV system of Mars Pathfinder's Sojourner.

(15) During an autumn day at VL2 site (sol 872) the optimum solar efficiency ranges from 8 to 16 % in case of a good quality engine. The high efficiency thermal engine should be used in combination with a solar collector kept perpendicular on Sun's rays. A maximum output power of 15 W might be provided a short time period during the day.

(16) When a horizontal solar collector is considered, the dependence of the maximum output power on optimum solar efficiency seems to be quadratic at both VL1 and VL2 sites. When a collector perpendicular on Sun's rays is considered, this dependence is more complicate, but keeps the quadratic feature.

(17) A certain optimum solar efficiency threshold (around 5 %) must be exceeded in order the system provide useful power.

(18) The dependence of the maximum output power on the optimum solar efficiency is stronger in case of a high efficiency thermal engine than in case of a medium efficiency engine, on one hand, and in case of a collector kept perpendicular on Sun's rays than in case of a horizontal collector, on the other one hand.

(19) No obvious difference exists between power plant performances in the two years of VL2 operation.

A solar Stirling engine based on a horizontal selective flat-plate converter was analyzed in section 5.4. A numerical optimization procedure was used to maximize the power provided by the engine. The main results are as follows:

(1) The values of the overall heat loss coefficient U_L range between about 0.3 and $0.7 \text{ Wm}^{-2}\text{K}^{-1}$.

(2) The solar energy conversion efficiency at noon could be as high as 0.18 at VL2 site during autumn and winter. Generally, it is lower during spring and summer.

(3) The output power at noon produced by the Stirling engine at VL2 site was as high as 16 W during autumn and winter. It obviously decreases during the dust storm periods.

(4) Properly designed solar Stirling engines equipped with horizontal flat-plate collectors could provide output power comparable in magnitude to that of PV cell systems. However, this conclusion should be treated with caution as it is based on figures referring to an existing PV system, on one hand, and to a simplified solar Stirling engine model, on the other one hand.

(5) To increase the output power and the efficiency, the design parameters of Stirling engines operating on Mars should be different from those of terrestrial engines. For example, one expects the compression ratio to exceed the usual value $\varepsilon_V=2$ while a smaller regeneration factor than $x = 0.2$ to be implemented.

Acknowledgments

The author thanks Professors Gheorghe Popescu, Monica Costea (Polytechnic University of Bucharest) and Michel Feidt (University of Nancy) for previous collaboration.

References

- Angelino, G., Invernizzi, C.: Cyclic Methylsiloxanes as working fluids for space power cycles. *J. Solar. Energy Engng.* 115, 130–137 (1993)
- Angelo Jr., J.A., Buden, D.: The nuclear power satellite (NPS) - key to a sustainable global energy economy and solar system civilization. In: Proc. of SPS 1991, Power from space, Paris, August 27-30, pp. 117–124 (1991)
- Badescu, V.: Discussion on the unification of three different theories concerning the ideal conversion of enclosed radiation. *J. Sol. Energy Engng.* 110, 349 (1988)
- Badescu, V.: On the thermodynamics of the conversion of diluted radiation. *J. Phys. D* 23, 289–292 (1990)
- Badescu, V.: Maximum conversion efficiency for the utilization of multiply scattered solar radiation. *J. Phys. D* 24, 1882–1885 (1991)
- Badescu, V.: Optimum operation of a solar converter in combination with a Stirling or Ericsson heat engine. *Int. J. Energy* 17, 601–607 (1992)

- Badescu, V.: Dynamic solar space power systems: optimum design and operation. *Space Technol.* 14, 331–337 (1994)
- Badescu, V.: Different strategies for maximum solar radiation collection on Mars surface. *Acta Astronautica* 43, 409–421 (1998a)
- Badescu, V.: Accurate upper bounds for the conversion efficiency of black-body radiation energy into work. *Phys. Lett. A* 244, 31–34 (1998b)
- Badescu, V.: Accurate upper bounds for the efficiency of converting solar energy into work. *J. Phys. D* 31, 820–825 (1998c)
- Badescu, V.: Simple upper bound efficiencies for endoreversible conversion of thermal radiation. *J. Non-Equilib Thermodyn* 24, 196–202 (1999)
- Badescu, V.: Accurate upper bound efficiency for solar thermal power generation. *Int. J. Solar Energy* 20, 149–160 (2000)
- Badescu, V.: Inference of atmospheric optical depth from near-surface meteorological parameters on Mars. *Renewable Energy* 24, 45–57 (2001)
- Badescu, V.: Simulation of a solar Stirling engine operation under various weather conditions on Mars. *J. of Solar Energy Engng.* 126, 812–818 (2004)
- Badescu, V., Popescu, G., Feidt, M.: Model of optimized solar heat engine operating on Mars. *Energy Conv. Mngmnt.* 40, 1713–1721 (1999)
- Badescu, V., Popescu, G., Feidt, M.: Design and optimisation of a combination solar collector - thermal engine operating on Mars. *Renewable Energy* 21, 1–22 (2000a)
- Badescu, V., Popescu, G., Feidt, M.: Simulation of a Martian solar thermal power plant: diurnal operation and power-efficiency correlations. *J. of the British Interplanetary Soc.* 53, 131–144 (2000b)
- Badescu, V., Popescu, G., Feidt, M.: Simulation of a thermal solar power plant operating on Mars under clear sky and dust storm conditions. *Acta Astronautica* 49, 667–679 (2001a)
- Badescu, V., Popescu, G., Feidt, M., Costea, M.: Optimisation du fonctionnement sur Mars d'un moteur de Stirling solaire (in French). *Termotecnica* 5(1), 24–28 (2001b)
- Bejan, A.: Unification of three different theories concerning the ideal conversion of enclosed radiation. *J. Sol. Energy Engng.* 109, 46–51 (1987)
- Bejan, A.: Advanced engineering thermodynamics. Wiley, New York (1988)
- Benz, N., Beikircher, T.: High efficiency evacuated flat-plate solar collectors for process steam production. *Solar Energy* 65, 111–118 (1999)
- Candau, Y.: On the exergy of radiation. *Solar Energy* 75, 241–247 (2003)
- Duffie, J.A., Beckmann, W.A.: Solar Energy Thermal Processes. Wiley, New York (1974)
- Eaton, C.B., Blum, H.A.: The use of moderate vacum environments as a means of increasing the collectors efficiencies and operating temperatures of a flat-plate solar collectors. *Solar Energy* 17, 151–158 (1975)
- Golombek, M.P., Cook, R.A., Economou, T., Folkner, W.M., Haldermann, A.F.C., Kallemeyn, P.H., Knudsen, J.M., Manning, R.M., Moore, H.J., Parker, T.J., Rieder, R., Schofield, J.T., Smith, P.H., Vaughan, R.M.: Overview of the Mars Pathfinder mission and assessment of landing site predictions. *Science* 278, 1743–1748 (1997)
- Hourdin, F., Forget, F., Talagrand, O.: The sensitivity of the Martian surface pressure and atmospheric mass budget to various parameters: A comparison between numerical simulations and Viking observations. *J. Geophys Res.* 100(E3), 5501–5523 (1995)
- Howell, J.R., Bannerot, R.B.: Optimum solar collector operation for maximum cycle work output. *Solar Energy* 19, 149–153 (1977)
- Jeter, S.J.: Maximum conversion efficiency for the utilization of direct solar radiation. *Solar Energy* 26, 231–236 (1981)

- Landsberg, P.T., Tonge, G.: Thermodynamic energy conversion efficiencies. *J. Appl. Phys.* 51, R1–R20 (1979)
- Landsberg, P.T., Mallinson, J.R.: Thermodynamic constraints, effective temperatures and solar cells. In: *Coll. Int. sur l'Electricite Solaire*, Toulouse, CNES, pp. 27–35 (1976)
- Landsberg, P.T., Badescu, V.: The geometrical factor of spherical radiation sources. *Euro-phys Lett.* 50, 816–822 (2000)
- Lide, D.R. (ed.): *Handbook of chemistry and physics*, 71th edn., pp. 15–39. CRC Press, Boca Raton (1991)
- Martin, L.J., Zurek, R.W.: An analysis of the history of dust activity on Mars. *J. Geophys. Res.* 98(E2), 3221–3246 (1993)
- McLallian, K.L., et al.: The solar dynamic radiator with a hystorical perspective. In: *Proceedings of the 23rd International Energy Conversion Engineering Conference*, Denver, CO, ASME, July 31-August 5, 1988, vol. 3, pp. 335–340 (1988)
- Meinel, A.B., Meinel, M.P.: *Applied Solar Energy*. Addison-Wesley Publishing Company, Reading (1976)
- Menetrey, W.R.: Space applications of solar energy. In: Zarem, A.M., Erway, D.D. (eds.) *Introduction to the utilization of solar energy*, p. 326. Mc Graw Hill, New York (1963)
- Mozjorine, Y.A., Senkevich, V.P., Koval, A.D., Narimanov, E.A.: Small - scale space power stations: feasibility and usage prospects. In: *Proc. of SPS 91, Power from space*, Paris, August 27-30, 1991, pp. 381–392 (1991)
- Petela, R.: Exergy of heat radiation. *J. Heat Transfer* 86, 187–192 (1964)
- Petela, R.: Exergy of undiluted thermal radiation. *Solar Energy* 74, 469–488 (2003)
- Pop, M.G., Leca, A., Prisecaru, I., Neaga, C., Zidaru, G., Musatescu, V., Isbasoiu, E.C.: *In-drumar -Tabele, monograme si formule termotehnice* (in Romanian), vol. 1. Editura Tehnica, Bucuresti (1987)
- Press, W.H.: Theoretical maximum for energy from direct and diffuse sunlight. *Nature* 264, 734–735 (1976)
- Prisnjakov, V.F.: SPS interest and studies in USSR. In: *Proc. of SPS 1991, Power from space*, Paris, August, 27-30, p 36 (1991)
- Prisnjakov, V.F., Statsenko, I.N., Kondratjev, A.I., Markov, V.L., Petrov, B.E., Gabrinets, V.A.: Developing space power Brayton systems with solar heat input. Research of working process of high temperature latent heat storage system. In: *Proc. of SPS 1991, Power from space*, Paris, August 27-30, pp. 465–470 (1991)
- Prisnjakov, V.F., Statsenko, I.N., Kondratjev, A.I., Markov, V.L., Petrov, B.E., Gabrinets, V.A.: Developing a space power Brayton system. *Space Power* 13, 135–144 (1994)
- Secunde, R., Labus, T.L., Lovely, R.G.: Solar dynamic power module design. In: *Proc. 24th International Energy Conversion Conf.*, vol. 1, pp. 299–307. IEEE, Piscataway (1989)
- Senft, J.R.: An introduction to low temperature differential Stirling engines. Moriya Press, River Falls (1996)
- Spanner, D.C.: Introduction to thermodynamics, p. 218. Academic Press, London (1964)
- Stefanescu, D., Marinescu, M., Danescu, A.: *Transferul de caldura în tehnica* (in Romanian), vol. 1. Editura Tehnica, Bucuresti (1982)
- Weingartner, S., Blumenberg, J., Ruppe, H.O.: Influence of orbit on solar - dynamic power systems. *Space Power* 13, 103–120 (1994)
- Zurek, R.W., Barnes, J.R., Haberle, R.M., Pollack, J.B., Tillman, J.E., Leovy, C.B.: Dynamics of the atmosphere of Mars. In: Kieffer, H.H., et al. (eds.) *Mars*, ch. 26, pp. 835–933. University of Arizona Press (1992)

Chapter 6

Fuel Cell Power System Options for Mars Surface Mission Elements

Simon D. Fraser

Graz University of Technology, Austria

6.1 Introduction

Fuel cells have played a key role in manned space exploration since the 1960s. Alkaline fuel cells have powered the Apollo spacecraft, safely taking men to the Moon and back, and still provide electrical energy to the Space Shuttle Orbiter today. Recent advances in fuel cell research have provided promising new cell materials and technologies. It will only be a question of funding and time until these scientific advances can be successfully turned into durable and reliable systems designed for even the most challenging applications in space exploration.

Space industry has been one of the few branches where fuel cell systems have already been economically competitive in niche applications in the past (Fraser 2001). This will even more so be the case in the near future, when power sources will have to be developed for very challenging and demanding applications in planetary exploration.

In space industry, the objective is always to design each system as light as possible to keep the launch costs low. In case of power systems for Mars surface applications, a promising approach of how to provide propellants to ascent and Earth-return spacecraft - as well as to various surface-bound mission elements - has been proposed. This so-called In-Situ Resource Utilisation (ISRU) aims at producing propellants and/or consumables directly on location – e.g. on the surface of Moon or Mars - rather than launching a full supply from Earth (Zubrin and Wagner 1996; Hoffman and Kaplan 1997). These ISRU-based split-mission scenarios introduced an elegant way of avoiding the assembly of a huge single-shot spacecraft in Orbit, as suggested by previous concepts (Cohen 1989; National Space Council 1990).

Different ISRU options have been considered in the past. Taking a feedstock of hydrogen, water or hydrocarbons from Earth and producing methane and oxygen on location is one of the most promising options. Directly splitting carbon dioxide taken from the Martin atmosphere into carbon monoxide and oxygen is another very interesting option.

Each ISRU-approach has certain drawbacks and offers certain advantages. Determining the ideal fuel/oxidant combination is therefore no straight-forward decision, but rather depends on various mission parameters and has to be determined on a comprehensive evaluation of mission profile (e.g. mission duration, energy

versus power demand, operating conditions) and local availability of resources (see Chap. 16).

Assuming that an ISRU system is available for propellant production, the fuels and oxidants produced could not only be used as propellants for ascent and Earth-return vehicles, but also for power generation applications with surface-bound mission elements. Mobile pressurised rovers designed to act as mobile habitat and laboratory for scientists, for instance, require substantial quantities of energy when operated on a continuous 14-day surface exploration trip. This energy demand cannot be easily satisfied by photovoltaic arrays and/or batteries, as the external surface area of a rover is normally too small for mounting large solar panels, and mass estimations presented within this chapter will show that the estimated battery mass would also be prohibitively high.

A fuelled system, however, could easily provide the energy storage capacity if an efficient energy conversion technology is utilised. Many different technologies for fuel-to-electrical conversion are available. This includes dynamic conversion technologies such as internal combustion engines, microturbines or Stirling engines as well as static conversion technologies such as thermoelectrical (TE) and thermophotovoltaic (TPV) generators or alkali metal thermal-to-electrical converters (AMTEC) (Fraser 2001; Baker 2002).

One of the most promising options for fuel-to-electric conversion are fuel cells, as they combine the high specific energy and the high stored energy density of a fuel/oxidant combination with an inherently simple conversion system design capable of operating at sufficiently high conversion efficiency levels. Fuel cells therefore seem perfectly suited for some of the most demanding applications in Mars surface exploration.

This chapter will provide the reader with a basic overview of fuel cell technology and discuss the possibilities as well as the limitations of this energy conversion technology in relevant Mars surface applications.

6.2 Fuel Cell Fundamentals

Fuel cells combine features of batteries and fuelled generators utilising principles known since 1839, when Sir William Grove demonstrated the first operational fuel cell.

Today, 170 years after these first demonstrations, fuel cells still have not gained the relevance in real-life applications almost constantly predicted by experts since the 1950s. This is partly due to the fact that fuel cells will have to substitute established and cost-optimised technologies in mass-market applications; and in part this is also due to the fact that fuel cell technology – although promising an inherently simple system design – still requires extensive fundamental research to understand design and material challenges faced on the way towards designing und building systems capable of managing a wider penetration into commercial mass markets.

This section is therefore dedicated to providing the reader with a compressed overview of principles, types and configurations of fuel cells and fuel cell systems. This introduction is followed by a short overview of the history of fuel cells in space flight applications.

6.2.1 Basic Principles of Fuel Cells

Fuel cells are electrochemical conversion devices that directly produce electrical energy from fuels and oxidants without the intermediate conversion step into thermal energy required by heat engines such as internal and external combustion engines.

Fuel cells essentially are batteries that do not have to be recharged, as the electrodes are not consumed or intentionally modified in operation, but only provide the catalytically active sites required to engage the continuous flows of reactants in the electrochemical reactions. As long as an input of fresh reactants is available, fuel cells can therefore constantly produce electrical energy without being subject to intermittent charge and discharge cycles.

The basic elements common among all different types and designs of fuel cells are shown in Fig. 6.1.

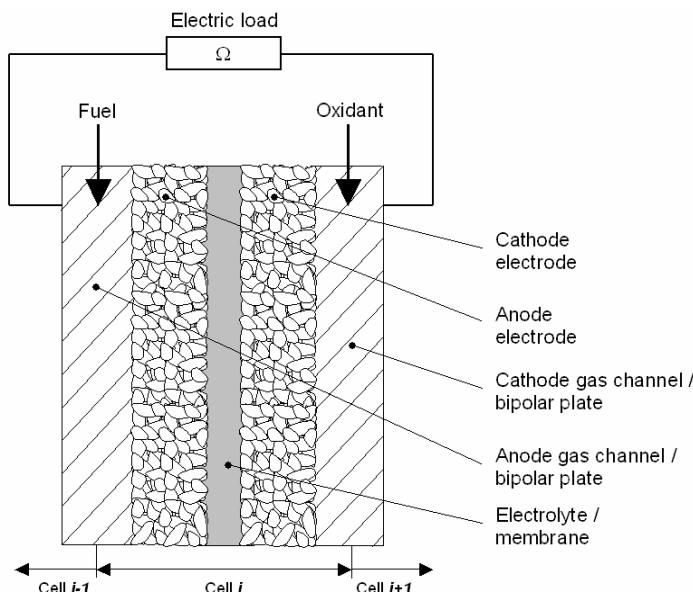


Fig. 6.1 Basic elements of a fuel cell

A constant stream of electrons flows from the anode electrode to the cathode electrode when an active fuel cell is connected to an electric load. Fuel is supplied to the anode electrode utilising a system of gas channels, whereas oxygen or ambient terrestrial air are fed to the cathode electrode utilising a similar gas distribution system.

An electrolyte is sandwiched between the anode and the cathode electrode. This electrolyte can either be a liquid, a liquid bound in a matrix, or a solid polymeric or ceramic membrane. Catalyst materials can be applied at the interface regions between the electronically conductive electrodes and the electronically isolating electrolyte to increase the electrochemical reaction kinetics, and thus to increase the current and power output of the cell.

A single fuel cell run on hydrogen/oxygen will normally produce operational output voltages of less than one Volt. In order to provide a useful output voltage, actual fuel cell systems normally consist of a number of individual cells electrically connected in series. Such a module consisting of a number of individual cells is referred to as a *fuel cell stack*. This is also indicated in Fig. 6.1., where single cell '*i*' of a stack is shown, and the adjacent cells '*i-1*' and '*i+1*' are attached to the right and left hand side by so-called bipolar plates. These bipolar plates do not only electrically connect the individual cells in series, but also have the gas channel systems required for fuel and oxidant distribution machined into the volume of the cell interconnectors.

A full fuel cell stack is shown in Fig. 6.2. The stack consists of a number of individual cells sandwiched between two end plates. Anode and cathode input and output ducts as well as the electrical contacts where the fuel cell stack is connected to the electric load are normally installed into the end plates. Fresh reactants are distributed among the individual cells utilising a system of gas channels that assures that each cell is constantly and sufficiently supplied. Gaseous and sometimes in part liquid (water) products of the electrochemical reactions are removed from the cells either continuously or intermittently (dead-end operation) with the offgas.

Thermal management is often achieved with an air- or water-based cooling system, and cell monitoring systems are often used for inspecting the proper operation of each cell. The thermal management and cell monitoring systems are not shown in Fig. 6.2 for the sake of simplicity.

In fuel cell technology it is common practise to give area-specific specifications (i.e. per unit of geometrical electrode area) in order to provide performance and design figures independent of cell or stack dimensions. A power output of 0.5 W cm^{-2} therefore specifies that half a Watt of net electrical output power is available

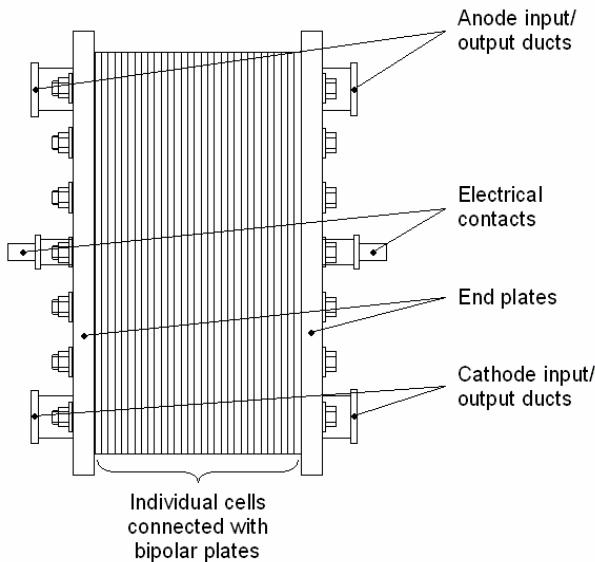


Fig. 6.2 Basic construction of a fuel cell stack

per square centimetre of geometrical electrode area. This notation is also applied throughout this chapter.

6.2.2 Fuel Cell Types

Many different types of fuel cells have been developed and investigated primarily since the 1950s (Kordesch and Simader 1996; Larminie and Dicks 2000).

Fuel cell types are normally distinguished by their electrolyte and – more recently – are further sub-divided into different operating temperature sub-regimes. Thus, the five classical types of fuel cells - i.e. Alkaline Fuel Cells (AFCs), Polymer Electrolyte Membrane Fuel Cells (PEMFCs), Phosphoric Acid Fuel Cells (PAFCs), Molten Carbonate Fuel Cells (MCFCs), Solid Oxide Fuel Cells (SOFCs) – can be subdivided into the range of modern fuel cells types presented in Table 6.1.

Table 6.1 Overview of modern fuel cell types

Type	Abbreviation	Electrolyte
Alkaline Fuel Cell	AFC	Potassium hydroxide solution (KOH), liquid or in matrix
Alkaline Anion Exchange Membrane Fuel Cell	AAEM-AFC	Alkaline anion exchange membrane
Phosphoric Acid Fuel Cell	PAFC	Phosphoric acid (H_3PO_4) in matrix
(Low-Temperature) Polymer Electrolyte Membrane Fuel Cell ^{a,b}	(LT-)PEMFC ^{a,b}	Perfluorosulfonic acid membrane ^c
Direct-Methanol Fuel Cell	DMFC	Perfluorosulfonic acid membrane ^c
High-Temperature Polymer Electrolyte Membrane Fuel Cell	HT-PEMFC	HT membrane ^d
Molten Carbonate Fuel Cell	MCFC	Molten mixture of alkali metal carbonates
Low-Temperature Solid Oxide Fuel Cell	LT-SOFC	Ion-conducting ceramic material
Intermediate-Temperature Solid Oxide Fuel Cell	IT-SOFC	Ion-conducting ceramic material
(High-Temperature) Solid Oxide Fuel Cell ^e	(HT-)SOFC ^e	Ion-conducting ceramic material

^aNormally referred to as PEMFC without the prefix LT

^bPEMFCs are also known as PEFCs and SPFCs (Solid Polymer Fuel Cells)

^ce.g. NAFION membranes (registered trademark of Dupont)

^de.g. membranes based on Polybenzimidazole (PBI)

^eNormally referred to as SOFC without the prefix HT

Each type of fuel cell offers specific possibilities and limitations. Two of the most important fuel cell characteristics are operating temperature and possible fuels as well as poisonous species.

The operating temperature generally has a very strong impact on system design and application, considering that start-up will obviously rather be an issue with a system operating at 800°C than with a system operating at room temperature. An increase in cell temperature – on the other hand – offers considerable advantages with respect to reaction kinetics and catalyst materials.

The possibility of avoiding the utilisation of platinum group metal (PGM) catalysts such as Platinum, Palladium or Ruthenium in high-temperature fuel cells such as SOFCs and MCFCs does not only reduce the system costs, but also simplifies the utilisation of hydrocarbon (HC) fuels such as methane. This is due to the fact that neither carbon mono- nor dioxide are poisonous for high-temperature fuel cells, whereas even small traces of the former in case of acidic low-temperature fuel cells (e.g. PEMFCs) and the latter in case of alkaline low-temperature fuel cells have to be removed by upstream gas processing. Not removing these species may result in deactivation of catalyst materials or clogging of the porous gas diffusion electrodes.

An overview of operational characteristics and technological status of the different types of fuel cells is provided in Table 6.2.

Table 6.2 Overview of operating parameters and technological status

Fuel Cell type	Operating temperature	Fuel options (poisonous species)	Technological status
AFC	60 - 90°C	H ₂ (poison: CO ₂)	available (space)
AAEM-AFC	60 - 90°C	H ₂ (poison: CO ₂)	current R&D
PAFC	160 - 220°C	H ₂ (poison: CO/mildly)	outdated
(LT-)PEMFC	25 - 95°C	H ₂ (poison: CO/strongly)	available
DMFC	25 - 95°C	CH ₃ OH	available
HT-PEMFC	120 - 200°C	H ₂ (poison: CO/mildly)	current R&D
MCFC	620 - 660°C	H ₂ , CO, reformate	available
LT-SOFC	< 650°C	H ₂ , CO, reformate	current R&D
IT-SOFC	650 - 800°C	H ₂ , CO, reformate	current R&D
(HT-)SOFC	800 - 1000°C	H ₂ , CO, reformate	available

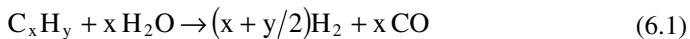
Half-cell reactions and mobile ions of the five major types of fuel cells are presented in Fig. 6.3.

Hydrogen is the only species directly engaged in the electrochemical anode half-cell reactions of the low- and medium-temperature fuel cell types (AFCs, PEMFCs, and PAFCs). High-temperature fuel cells (MCFCs and SOFCs), on the other hand, can directly utilise hydrogen and carbon monoxide as fuel, and are thus particularly well-suited for operation with hydrocarbon fuels sources. Hydrocarbons can normally not be directly utilised in fuel cells, with the only relevant exception being methanol in DMFCs, but have to be converted into a hydrogen

	Anode electrode		Cathode electrode
AFC	$H_2 + 2OH^- \rightarrow 2H_2O + 2e^-$	OH^-	$\frac{1}{2}O_2 + H_2O + 2e^- \rightarrow 2OH^-$
PEMFC	$H_2 \rightarrow 2H^+ + 2e^-$	H^+	$\frac{1}{2}O_2 + 2H^+ + 2e^- \rightarrow H_2O$
PAFC	$H_2 \rightarrow 2H^+ + 2e^-$	H^+	$\frac{1}{2}O_2 + 2H^+ + 2e^- \rightarrow H_2O$
MFC	$H_2 + CO_3^{2-} \rightarrow H_2O + CO_2 + 2e^-$ $CO + CO_3^{2-} \rightarrow 2CO_2 + 2e^-$	CO_3^{2-}	$\frac{1}{2}O_2 + CO_2 + 2e^- \rightarrow CO_3^{2-}$
SOFC	$H_2 + O^{2-} \rightarrow H_2O + 2e^-$ $CO + \frac{1}{2}O^{2-} \rightarrow CO_2 + 2e^-$	O^{2-}	$\frac{1}{2}O_2 + 2e^- \rightarrow O^{2-}$

Fig. 6.3 Half-cell reactions and mobile ions of the five major fuel cell types

and carbon monoxide rich synthesis gas. This can be done with a steam reforming reaction written in Eq. (6.1) in general form and in Eq. (6.2) for methane:



The carbon monoxide content in the reformat will normally be a problem if low-temperature fuel cells are utilised (i.e. AFCs, PEMFCs, and PAFCs). These problems are not only due to the fact that a direct utilisation of carbon monoxide is not possible with these cells types, but also – as mentioned above - due to the fact that carbon monoxide actually poisons many of the noble metal catalysts utilised with these cells chemistries. Complex gas clean-up is therefore required to remove most of the carbon monoxide before feeding the reformat to a low-temperature fuel cell. This is not required with high-temperature fuel cells, where the carbon monoxide can be directly utilised as fuel without the danger of poisoning the catalyst materials.

6.2.3 Electrolysers, Regenerative and Reversible Fuel Cells

Fuel cells and electrolyzers can essentially be based on identical electrochemical cells, utilising the same concepts, designs and materials. The only difference is that the overall and half-cell reactions - as well as mass transport and consumption of species and charges - are reversed when switching from fuel cell to electrolyser operation and vice versa. Two figures determine whether a cell is operated as a fuel cell or as an electrolyser: the operational cell voltage and the reversible cell voltage.

The *operational cell voltage* is the voltage available at the electrical contacts of the cell during operation. The term *operation* indicates that fresh reactants are supplied to the cell and that a net current is either supplied to or drawn from the cell.

The *reversible cell voltage*, on the other hand, is the theoretical voltage available if the conversion process proceeded without any losses. The reversible cell voltage can be derived from the molar change in Gibbs free energy of formation ΔG_f (kJ mol⁻¹) according to Eq. (6.3), and written explicitly for the case of a hydrogen/oxygen reaction in Eq. (6.4):

$$\Delta G_f = G_f(\text{products}) - G_f(\text{reactants}) \quad (6.3)$$

$$\Delta G_f = G_f(\text{H}_2\text{O}) - G_f(\text{H}_2) - 1/2 G_f(\text{O}_2) \quad (6.4)$$

If a fuel cell was operated with the reversible cell voltage, the change in Gibbs free energy would be fully available as electrical energy. In case of an electrolyser, only the Gibbs free energy would have to be supplied as electrical energy.

The reversible cell voltage can then be computed according to Eq. (6.5) where the change in Gibbs free energy (kJ mol⁻¹) is divided by the Faraday constant F (96,485 C mol⁻¹) and the number of electrons transferred in the reaction n_e (-). In case of a hydrogen/oxygen reaction, two electrons are transferred for each mole of hydrogen (anode half-cell reaction: $\text{H}_2 \rightarrow 2 \text{H}^+ + 2 \text{e}^-$; cathode half-cell reaction: $1/2 \text{O}_2 + 2 \text{H}^+ + 2 \text{e}^- \rightarrow \text{H}_2\text{O}$) and n_e therefore has a value of two:

$$E_0 = \frac{-\Delta G_f}{n_e F} \quad (6.5)$$

The Gibbs free energy of formation is not independent of operating conditions, but changes with temperature and pressure of the species. Tabulated values of the Gibbs free energy of formation for the hydrogen/oxygen reaction as a function of temperature can be found in literature (Barin 1989). The reversible cell voltage of a hydrogen/oxygen fuel cell is plotted as a function of temperature in Fig. 6.4 for the case of liquid water and steam as reaction product. All of the species are assumed to be present with an activity of one (i.e. standard pressure of 0.1 MPa or in liquid state), as discussed with the pressure-dependence of the reversible cell voltage below.

Figure 6.4 shows a decline in reversible cell voltage with temperature. This decline does not necessarily imply that the operational cell voltages of high-temperature fuel cells are smaller than the ones available with low-temperature fuel cells. The faster reaction kinetics of high-temperature operation tend to compensate the reduced reversible voltage. The operational voltages of high- and low-temperature fuel cells are therefore often in a similar range.

The change of the Gibbs free energy of formation with the pressure of the gaseous species can be computed utilising the Nernst equation written in Eq. (6.6) for a hydrogen/oxygen reaction. In this, the reversible cell voltage E_0 – computed for each species present at standard pressure (0.1 MPa) – and a pressure-dependent second term are added to give the reversible cell voltage E_{rev} available with the actual pressure levels of the species. The pressure-dependent term is a function of

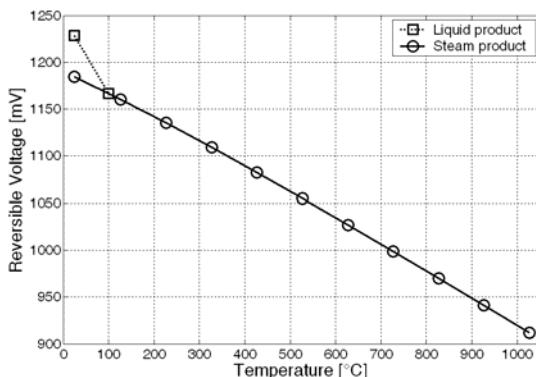


Fig. 6.4 Reversible cell voltage of a hydrogen/oxygen fuel cell as a function of temperature (all species at standard pressure of 0.1 MPa)

the universal gas constant R ($8.314 \text{ J K}^{-1} \text{ mol}^{-1}$), the investigated temperature T (K), the Faraday constant F ($96,485 \text{ C mol}^{-1}$) and the activities of the species:

$$E_{rev} = E_0 + \frac{RT}{2F} \ln\left(\frac{a_{H_2} a_{O_2}^{1/2}}{a_{H_2O}}\right) \quad (6.6)$$

Assuming ideal gas law, the activity of a gaseous species ‘x’ can be computed according to Eq. (6.7a). The activity of liquid water can be approximated according to Eq. (6.7b):

$$\text{Ideal gases: } a_x = p_x / p_x^0 \quad (6.7a)$$

$$\text{Liquid species: } a_x = 1 \quad (6.7b)$$

The reversible cell voltage is therefore independent of the cell being operated as fuel cell or as electrolyser and remains constant during operation as long as no change in operation conditions occurs (e.g. an increase in cell temperature due to ohmic heating would change the reversible cell voltage). The operational voltage, however, is not constant but depends on the electrical current supplied to or drawn from the cell.

Comparing reversible and operational cell voltage, one can determine whether a cell is operated as fuel cell or as electrolyser. If the operational cell voltage is higher than the reversible voltage, the cell is operated as an electrolyser and electrical energy is consumed in the electrochemical reaction. If the operational cell voltage is smaller than the reversible voltage, the cell is operated as a fuel cell and therefore generates an output voltage.

This correlation between cell operation and output voltage is shown in Fig. 6.5, where a typical cell voltage of a low-temperature cell (e.g. a PEMFC) is shown as a function of the normalised current density.

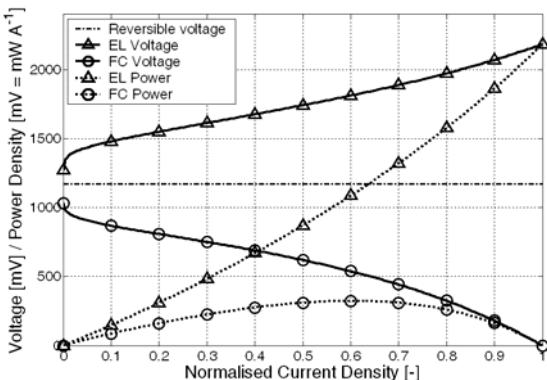


Fig. 6.5 Cell voltage and power density as a function of the normalised current density

The normalised current density gives the dimensionless ratio of investigated and limiting current density. The limiting current density specifies the area-specific electrical current that can be drawn from and/or supplied to cell before either the cell voltage drops to zero or a further increase could potentially damage the cell.

The power density plotted in Fig. 6.5 is given in mV or mW A^{-1} because it is computed as the product of cell voltage and dimensionless normalised current density. The absolute power density in mW cm^{-2} can be computed by multiplying the plotted power density values (mW A^{-1}) with the limiting current density (A cm^{-2}).

In fuel cell operation, the operational cell output voltage is initiated below the dash-dotted line indicating the reversible cell voltage, and then gradually drops down to zero when the investigated current density approaches the limiting current density, as shown in Fig. 6.5. This reduction in operational cell output voltage with current density is due to losses associated with the electrochemical reaction kinetics (the so-called activation losses), ohmic losses as well as mass transport limitations. These losses increase either linearly with current density (e.g. ohmic losses) or non-linearly (e.g. activation losses of the electrochemical reactions and mass transport limitations) (Fisher 1996; Hamann and Vielstich 1998; Fraser 2004).

In electrolyser operation, the voltage gradually increases until again a limiting current density is achieved that should not be exceeded e.g. to avoid damaging the porous electrode structure when the gas generation rate is too high.

The power output density of a fuel cell – and correspondingly also the power input to an electrolyser – can simply be calculated by multiplying the operational cell voltage with the current density. The power input to an electrolyser therefore gradually increase with current densities, whereas the power output of a fuel cell reaches a maximum before falling back to zero when the operational cell voltage also drops down to zero at high current density operation. Power density curves for fuel cell and electrolyser operation are shown as dashed lines in Fig 6.5.

6.2.4 Fuel Cell System Configurations

An electrochemical cell can be utilised for converting electrical into chemical energy of a fuel when operated as electrolyser (EL) or for power generation, converting chemical energy of a fuel into electricity, when operated as fuel cell (FC). The combined operation of electrolyser and fuel cell - providing the possibility of intermittently storing electricity in the chemical energy of a fuel/oxygen combination - is possible. Four basic fuel cell system configurations can be distinguished:

- (a) Direct-fuel fuel cell systems (primary)
- (b) Indirect-fuel fuel cell systems (primary)
- (c) Regenerative fuel cell systems (secondary)
- (d) Unitised regenerative (or reversible) fuel cell systems (secondary)

The primary direct-fuel fuel cell system offers a very simple design. This system has to be refuelled from an external source, as suggested by the word *primary* borrowed from primary (i.e. non-rechargeable) batteries as opposed to secondary (i.e. rechargeable) systems. *Direct-fuel* indicates that the fuel stored in the system can be directly utilised in the electrochemical reactions of the fuel cell without any further processing steps. A typical terrestrial application for such a system would be a direct-hydrogen fuel cell system. Hydrogen taken from the storage vessel just has to be conditioned (e.g. warmed and humidified) before feeding it to the fuel cell module in such a system design.

The primary *indirect-fuel* fuel cell system also has to be refuelled from an external source. The fuel stored in the system, however, cannot be directly utilised in the electrochemical reactions of the fuel cell and has to be pre-processed in a fuel processor (FP) before utilising it in the fuel cell.

A typical application for such a system is if a liquid hydrocarbon fuel is to be utilised in a low-temperature fuel cell. As a direct utilization is not possible, the fuel therefore has to be processed prior to supplying it to the cell module, making the system an *indirect-fuel* system.

The only exception where a direct utilisation of a liquid hydrocarbon fuel in a low-temperature fuel cell is possible are direct-methanol fuel cells (DMFCs). These fuel cells have special Platinum/Ruthenium catalysts that enable a direct utilisation of (liquid) methanol without prior pre-reforming; DMFCs therefore are direct-fuel fuel cell systems.

Liquid hydrocarbon fuels normally have to be evaporated and reformed into a hydrogen-rich synthesis gas before being fed to the fuel cell module. This pre-processing of the stored fuel increases the overall fuel cell system mass, volume and complexity. This additional burden of pre-processing is considered in the nomination as *indirect-fuel* fuel cell system.

Secondary regenerative fuel cell systems can be recharged by supplying electrical energy to a built-in electrolyser, thus providing similar electrical/electrical energy storage characteristics as a secondary battery. Two basic designs are possible: a regenerative fuel cell system where fuel cell and electrolyser are two separate units, and a *unitised* regenerative fuel cell (URFC) system where fuel cell and electrolyser are identical (i.e. the same electrochemical cell is operated as electrolyser and as fuel cell).

The former offers the advantage of separately optimising fuel cell and electrolyser to their specific applications, whereas the latter offers the full functionality in one – potentially – very compact and light-weight unit by avoiding the installation of two separate cell modules.

Choosing or not choosing a unitised approach primarily depends on the available cell technology and the intended application. The current status of fuel cell and electrolyser technology often suggests the use of two separate units; a unitised approach is a very interesting target for future power system research. However, the unitised technology will always suffer from reduced efficiency in fuel cell and/or in electrolyser mode.

Simplified system layouts of these four basic fuel cell system configurations are shown in Fig. 6.6. The abbreviation PMAD indicates the power management and distribution to which the electric terminals of fuel cells and electrolyzers are connected to.

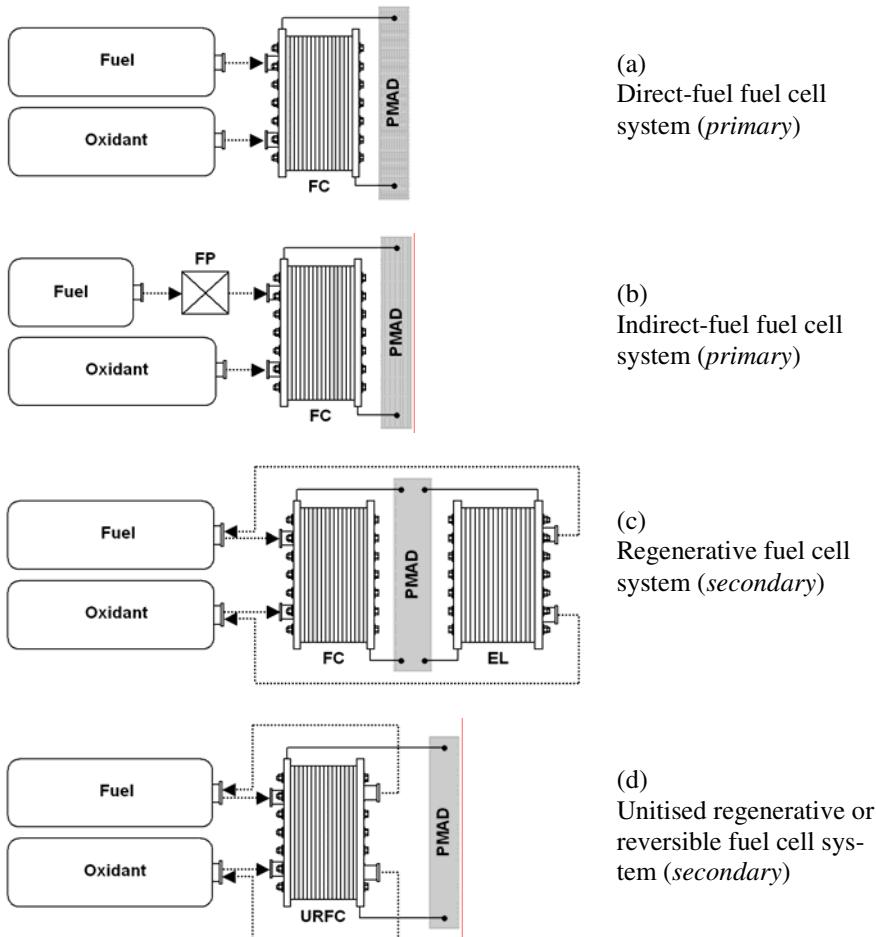


Fig. 6.6 Overview of fuel cell system configurations

6.2.5 Short History of Fuel Cells in Space Applications

Three manned capsules were developed and launched in the manned space flight program of the National Aeronautics and Space Administration (NASA) in the 1960s and 1970s: the Mercury, Gemini and Apollo capsules.

The issue of onboard power supply was not a very pressing issue with the Mercury Program. Due to the short flight durations, a comparably small battery system was sufficient to provide electrical energy to instrumentation and life support systems.

6.2.5.1 *The Gemini Program*

Within the Gemini Program, a battery system did not seem the appropriate choice for onboard power supply anymore due to the fact that longer flight durations were considered and higher output power levels were required. In addition, the Gemini missions were also considered as possibility for testing and training personnel and hardware for the upcoming Apollo Program, pursuing even more ambitious goals with respect to flight duration and onboard power system requirements. Fuel cells were therefore chosen for the Gemini program.

General Electrical (GE) developed an acidic fuel cell system with an electrical output power of 1 kW for the Gemini capsules. The GE modules were based on acidic cell chemistry and applied a polystyrene sulfonate ion-exchange membrane as electrolyte. Although the GE fuel cell systems performed well in the Gemini missions, issues of membrane stability and low power output density finally led to the decision not to consider polymer electrolyte membrane technology with the upcoming Apollo Program, but to switch to alkaline fuel cell technology instead (Warshay and Prokopiou 1990).

General Electrical continued to pursue the development of polymer electrolyte membrane fuel cells after the Gemini Program and eventually installed the new NAFION membranes from Dupont in their cells. NAFION ultimately became the standard membrane for PEMFCs, but GE at the time nevertheless decided to switch to phosphoric acid fuel cells due to the many challenging design issues of polymer electrolyte membrane fuel cell technology. The interest in PEMFCs renewed in the early 1990s, and many of the challenging design issues of the early GE systems were – and still are – again in the focus of intensive research and development (Larminie and Dicks 2000).

PEMFCs are nowadays comparably well-understood and managed to become the most advanced type of fuel cell. Most terrestrial demonstration applications are nowadays based on PEMFCs, and even a replacement of the Space Shuttle Orbiter alkaline fuel cell system with PEMFC technology was discussed (Larminie and Dicks 2000), but could not be realised due to cost impact.

6.2.5.2 *The Apollo Program*

Pratt & Whitney Aircraft Corp. licensed the Bacon patents on alkaline fuel cell technology in the early 1960s. Based on this technology, Pratt & Whitney won the NASA contracts for the Apollo spacecraft power supply (Kordes and Simader 1996).

The fuel cells developed within the Apollo program were based on sintered nickel electrodes with platinum metal catalysts. The catalyst loadings of up to

40 mg cm^{-2} provided comparably high output power levels. Although the operating temperatures were between 200 and 230°C , the operating pressures were only in the order of 0.3 MPa, and therefore fairly low. This could only be achieved by utilising highly concentrated electrolyte solutions (85% KOH) to prevent boiling (Kordesch and Simader 1996).

Round electrodes measuring a diameter of 200 mm were utilised. Each fuel cell module consisted of 31 individual cells. Three of these modules were operated aboard the Apollo spacecraft in parallel. The nominal rating of each module was 1.5 kW in normal and 2.3 kW in overload operation. Each module weighed 109 kg (Kordesch and Simader 1996).

Of the 92 fuel cells modules delivered by Pratt & Whitney, 54 have been utilised. This includes nine flights to the Moon, three Skylab missions and the Apollo-Soyuz mission (Kordesch and Simader 1996).

6.2.5.3 *The NASA Space Shuttle Orbiter Fuel Cell System*

After the Apollo Program, NASA also decided to use alkaline fuel cell technology for their Space Shuttle Orbiters. The Orbiter fuel cell modules are based on the technology previously developed during the Apollo Program and consist of 32 cells each.

Each Orbiter is equipped with three of these modules. Each module has a nominal rating of 12 kW and an overload capability of up to 16 kW. This documents the technological advantages achieved since the Apollo Program, where each module was only rated at 1.5 kW for nominal and 2.3 kW for overload operation. The operating temperature was reduced to 92°C , whereas the operating pressure was slightly increased to a maximum of 0.44 MPa (Kordesch and Simader 1996).

A picture of the Space Shuttle Orbiter fuel cell module is shown in Fig. 6.7.

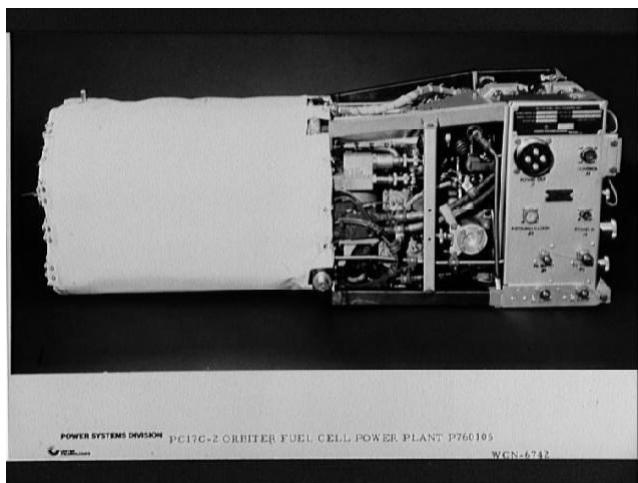


Fig. 6.7 The Space Shuttle Orbiter fuel cell module (reproduced with kind permission of Johnson Space Center/NASA)

6.2.5.4 The Photon Fuel Cell System

Plans for the Russian equivalent of the Space Shuttle Orbiter, the Buran, also included an alkaline fuel cell system. The Photon systems were to be installed in the front section of the payload bay and included cryogenically stored hydrogen and oxygen (Hendrickx and Vis 2007).

The Buran power system was to consist of four fuel cell modules with 32 cells each (the American Space Shuttle Orbiter only had three 32-cell modules). Each module was rated at 10 kW nominal and 25 kW overload operation. Similar to the Space Shuttle Orbiter systems, the cells were also designed for low-temperature operation at 100°C and 0.4 MPa (Hendrickx and Vis 2007).

A picture of the Photon fuel cell module is shown in Fig. 6.8.

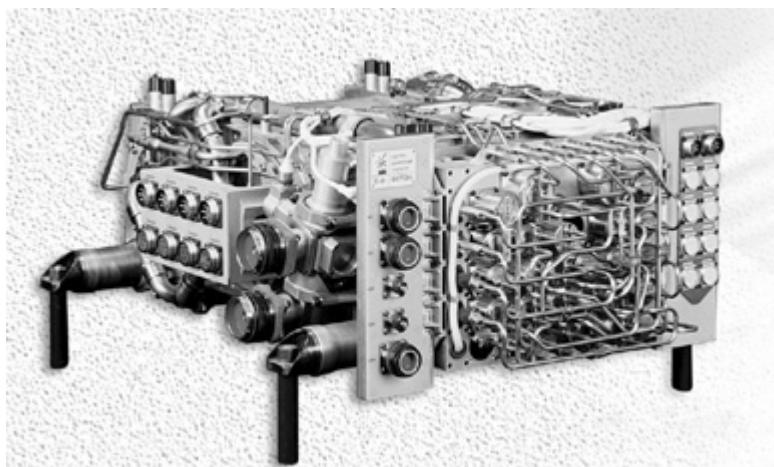


Fig. 6.8 Photon fuel cell module (reproduced with kind permission of Ural Electrochemical Integrated Plant)

The Buran Program was officially cancelled in 1993 after completing a single unmanned flight in 1988, where an unmanned Buran Space Shuttle was lifted into orbit by Energia boosters. Due to the short flight duration of the 1988 mission, the fuel cells modules were not installed.

In the course of European fuel cell power plant development activities for the Hermes spaceplane, a Photon fuel cell generator was tested in a realistic Hermes mission profile at ESA/ESTEC in 1992 (Schautz et al. 1993).

After a total of 80,000 hours of operation during development and maintenance, the Photon fuel cell system is now offered by Ural Electrochemical Integrated Plant (UEIP) for terrestrial and aerospace applications in a fully automated version (UEIP 2008).

Renewed interest for the Photon system is also reported with the presentation of the Kliper spacecraft, where modified Photon fuel cell modules could be used.

6.3 Fuel Cell for In-Situ Resource Utilisation

For any spacecraft or mission element relying on fuelled power and/or propulsion system elements, one initial question in system design suggests itself: should the spacecraft be equipped with sufficient quantities of fuel and oxidant prior to launch to operate throughout the whole mission duration without refuelling/recharging or is either an external supply or an on-board production of fuel and/or oxidant a feasible option.

Up to date, space missions solely relied on a ready cache of fuel and oxidant supplied to the spacecraft prior to launch, but the situation is different with the upcoming missions to the surface of Mars. Today's launch costs still are such a dominant factor that any potential reduction of the power and/or propulsion system mass is not only a technological, but also an economical challenge that has to be addressed. The buzzword for on-location production of fuel and/or oxidants is ISRU.

In-situ Resource Utilization (ISRU) is currently being understood as one of the key elements in advanced surface exploration of Moon and Mars. ISRU aims at reducing the total system mass that has to be transferred from the Earth onto the surface of Moon or Mars by utilizing resources that are easily accessible on location. This could be surface regolith with Moon and Mars, and – in case of Mars – primarily the atmosphere containing more than 95% carbon dioxide (Zubrin and Wagner 1996; Hoffman and Kaplan 1997; Fraser et al. 2005).

6.3.1 Sabatier/RWGS/Water Electrolysis ISRU System

Methane is commonly used as energy carrier in terrestrial applications, as it is the major constituent of natural gas. A considerable share of terrestrial fuel cell systems is operated with natural gas, as methane can be easily converted into a hydrogen-rich synthesis gas utilizing a catalytic fuel processor. In addition, methane is a particularly interesting fuel option for high-temperature fuel cells, as neither carbon mono- nor dioxide – both by-products of the reforming reaction – are poisonous for MCFCs and SOFCs.

Methane is also considered as the fuel of choice with many state-of-the-art Mars surface ISRU scenarios. Hence, synergies in fuel production and storage could be used if stationary and/or mobile mission elements could also be powered with the same fuel/oxidant combination as the ascent vehicle's rocket motors.

A simplified overview of the Sabatier/RWGS/water electrolysis ISRU system is shown in Fig. 6.9.

The key element of this ISRU approach is the Sabatier reaction. This reaction produces methane and steam from hydrogen and carbon dioxide. Thus, carbon dioxide taken from the ambient atmosphere could be utilized as in-situ resource. At ESA, this concept has been developed over 20 years for human space flight. ESA's Air Revitalisation System (ARES), currently in preparation for a flight demonstration experiment aboard the Columbus module of the International Space Station, is essentially based on a very similar plant layout and uses a Sabatier reactor and a water electrolyser for carbon dioxide removal and oxygen production in an air revitalisation system (Bockstahler et al. 2008; Witt et al. 2008). The Sabatier/RWGS/water

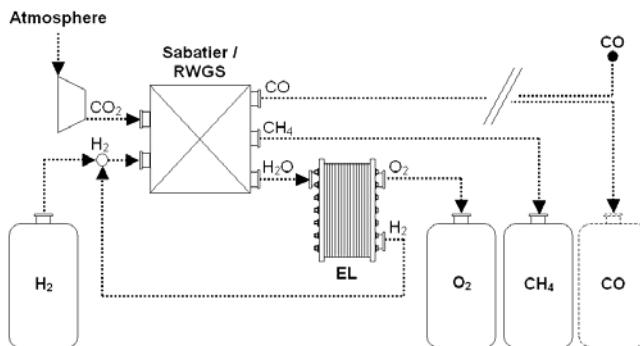


Fig. 6.9 Basic design of an ISRU system with Sabatier/RWGS/water electrolysis

electrolysis ISRU system essentially applies the same technology, just on a bigger scale and primarily for propellant production.

The Sabatier reaction, Eq. (6.8), produces two kilograms of methane from each kilogram of hydrogen. Steam, the second reaction product, can be electrolysed, producing hydrogen and oxygen. The hydrogen can be re-supplied into the process and the oxygen can be stored in compressed or liquefied form.

The Sabatier reaction with subsequent electrolysis of product water provides a mass leverage of twelve (i.e. 12 kg of methane and oxygen are produced from each kilogram of the hydrogen) according to Eq. (6.9), where the mass leverage is derived from the molar masses M_x of the species involved. Values of mass leverage utilised within this chapter generally refer to the leverage in species mass landed onto the surface of Mars excluding tanks and processing equipment. The effective mass leverage will be obviously reduced by the masses of production plant and storage systems:



$$\text{Mass leverage } (\text{H}_2\text{O electrolysed}) = \frac{M_{\text{CH}_4} + M_{\text{O}_2}}{2 \times M_{\text{H}_2}} \approx 12 \quad (6.9)$$

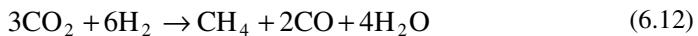
The reverse water gas shift (RWGS) reaction, Eq. (6.10), is well-known from terrestrial methane reforming processes. In an ISRU plant, the RWGS reaction could be used to produce additional quantities of oxygen (and carbon monoxide, either vented or utilised as fuel) by electrolysing the product water or in a combined process working in parallel to the Sabatier reaction as discussed below.

The mass leverage of the RWGS reaction, Eq. (6.11), is theoretically infinite, as no hydrogen is consumed. The mass leverage of a practical system will again be limited due to the mass of hydrogen required in the system as well as the processing and storage equipment:



$$\text{Mass leverage (CO utilised or vented)} = \frac{M_{\text{CO}} + M_{\text{O}_2}}{0} \rightarrow \infty \quad (6.11)$$

Combining Sabatier and RWGS reaction in a single reactor will result in one of the most promising ISRU cycles presented in Eq. (6.12). This combined reaction gives a mass leverage of 20 if CO is vented, and 34 if CO is utilized as fuel, as shown in Eqs. (6.13) and (6.14):



$$\text{Mass leverage (CO vented)} = \frac{M_{\text{CH}_4} + 2 \times M_{\text{O}_2}}{2 \times M_{\text{H}_2}} \approx 20 \quad (6.13)$$

$$\text{Mass leverage (CO utilised)} = \frac{M_{\text{CH}_4} + 2 \times M_{\text{CO}} + 2 \times M_{\text{O}_2}}{2 \times M_{\text{H}_2}} \approx 34 \quad (6.14)$$

6.3.2 Sabatier/Water Electrolysis/Methane Pyrolysis ISRU System

A slightly different approach - particularly suitable for smaller (robotic) missions - is to combine a Sabatier reactor and a water electrolyser with a methane pyrolysis reactor.

Hydrogen and carbon are produced in the pyrolysis reactor, following Eq. (6.15). Hydrogen can be re-fed into the Sabatier reactor and the (solid) carbon can be discarded:



By combining methane pyrolysis with a Sabatier reactor and a water electrolyser, carbon dioxide can be decomposed into carbon (produced in the pyrolysis reactor and subsequently discarded) and oxygen (produced in the water electrolyser). The oxygen can be stored and utilized as oxidant in rocket motors or fuel cells or fed into the manned habitat. Fuel is not produced by this cycle, though. A full cache of fuel therefore has to be taken from Earth.

Combining a fuel with good storage characteristics - e.g. a hydrocarbon such as propane (Kaplan 1996) - with ISRU oxygen production can result in a comparably simple but yet mass-optimised system. A significant fraction of the propellant mass does not have to be landed onto the surface of Mars as the oxygen tanks of the ascent and/or Earth-return vehicle are empty, but the ISRU system does not have the complexity of a system that is intended for production of oxygen and fuel on the surface of Mars.

A simplified overview of the Sabatier/water electrolysis/methane pyrolysis ISRU system is shown in Fig. 6.10.

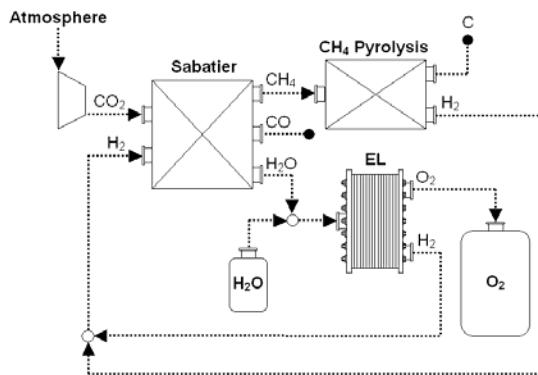


Fig. 6.10 Basic design of an ISRU system with Sabatier/water electrolysis/methane pyrolysis

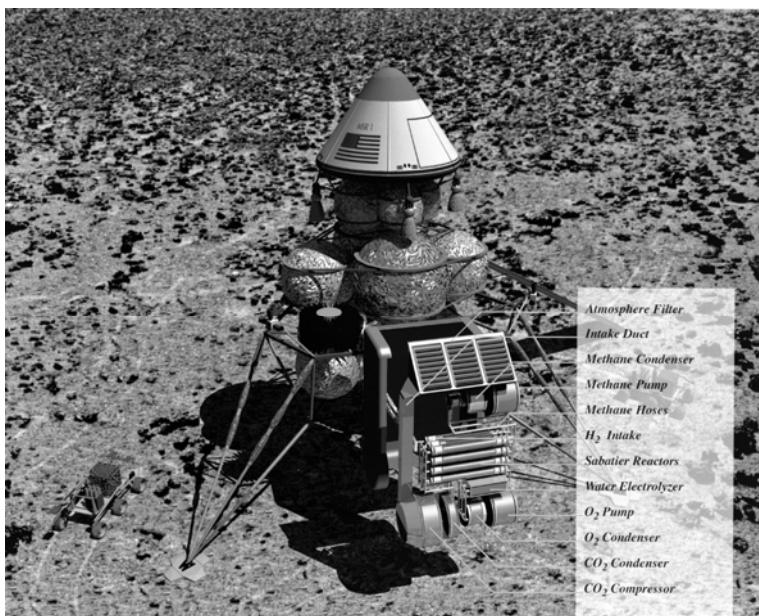


Fig. 6.11 Artist's impression of a Mars In-Situ Resource Utilization Sample Return mission (reproduced with kind permission of Johnson Space Center/NASA)

Fig. 6.11 shows a small sample return mission producing propellants for the return capsule utilizing ISRU technology.

This so-called Mars In-Situ Resource Utilization Sample Return (MISR) mission would consist of a robotic lander designed to collect soil and rock samples and send them on a return trip back to Earth. The lander would set down with the propellant tanks of the return vehicle empty. These propellants would then be

produced in the following roughly 300 days utilizing ISRU technology installed into the lower section of the lander.

After completing the propellant production cycle, the return vehicle could make the return trip and thus take the precious soil and rock samples back to Earth.

The ISRU system of such a small robotic lander could basically consist of all of the methane-based approaches presented above. Utilising a combination of a full cache of fuel with simple transport and storage characteristics and ISRU-produced oxygen is certainly an interesting option for such an application. Propane is suggested as fuel option with a MISR mission description presented in literature, for instance (Kaplan 1996).

6.3.3 Carbon Dioxide Decomposition ISRU System

All of the ISRU approaches presented above were based on a feedstock of hydrogen, water and/or a hydrocarbon fuel taken from Earth. An ISRU of water taken is theoretically feasible, as the Martin atmosphere contains about 210 ppm of water (NASA online 2009). An extraction utilising adsorption of water vapour into molecular sieves such as Zeolites is discussed in literature (Williams et al. 1995, Coons et al. 1997, Grover and Bruckner 1998). Water ice, recently confirmed in soil samples analysed by the NASA Phoenix Mars Lander, remains a long-term option.

In a short and medium timeframe, however, the utilization of a hydrogen- or hydrocarbon-based ISRU approach will most likely imply that a hydrogen-containing feedstock will have to be transported from Earth and landed on the surface of Mars. This is certainly a feasible approach, but any loss of the feedstock species during transport and surface storage could have catastrophic effects as long as no (technically and economically) feasible way of producing hydrogen from ambient resources can be found on Mars. These problems could be avoided if fuel and oxygen are directly produced without requiring any resources other than a supply of carbon-dioxide-rich Martian ambient air.

Carbon monoxide and oxygen have been previously considered as propellants for Mars ascent vehicles. Theoretical and experimental investigations of these propellants were made at NASA in the early 1990s (Linne et al. 1990; Linne 1991). A small rocket engine was developed (Linne 1997) and used to experimentally verify that carbon monoxide/oxygen can be successfully ignited and steadily combusted (Linne 2005).

Carbon monoxide can also be used as fuel in high-temperature fuel cells. MCFCs and SOFCs are regularly operated with high molar fractions of carbon monoxide in the anode electrode input synthesis gas. Even the operation of SOFCs with pure carbon monoxide is possible.

Producing carbon monoxide and oxygen on the surface of Mars by utilizing atmospheric carbon dioxide is probably the most elegant way of implementing ISRU. Thus, the power system would not rely on a feedstock of species brought from Earth, but only on a sufficient on-site supply with energy to drive the decomposition process and to store the reaction products.

Carbon dioxide decomposition according to Eq. (6.16) will result in an infinite theoretical mass leverage according to Eq. (6.17) neglecting the processing and storage equipment:



$$\text{Mass leverage} = \frac{M_{\text{CO}} + 1/2 M_{\text{O}_2}}{0} \rightarrow \infty \quad (6.17)$$

Two possible ways of decomposing carbon dioxide are considered: electrochemical reduction with a high-temperature solid oxide electrolyser and photocatalytic carbon dioxide decomposition.

Electrolytic reduction of carbon dioxide is based on a high-temperature solid oxide electrolyser, which is essentially a reversed solid oxide fuel cell (Förstner 1998). Yttria Stabilized Zirconia (YSZ) and other ceramic materials (e.g. BiO_2 , CeO_2 , Ta_2O_5) can be used as ceramic electrolyte, conducting negatively-charged oxygen ions from the cathode to the anode electrode. The conductivity of these ceramic electrolytes strongly depends on the operating temperature. Many of the ceramic materials require operating temperatures as high as 1000°C and above for efficient operation with acceptable output power densities. Recent advances in SOFC technology have led to the development of cell materials capable of operating at considerably lower temperatures of approximately 750°C .

A simplified overview of the electrolytic carbon dioxide decomposition ISRU system is shown in Fig. 6.12.

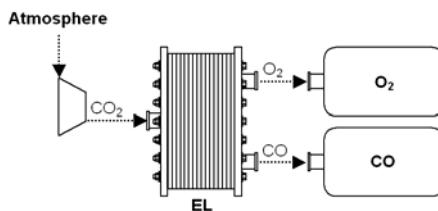


Fig. 6.12 Basic design of an ISRU system with carbon dioxide electrolysis

The problems linked to elevated operating temperatures and power consumption of solid oxide electrolyser technology has led to investigations of alternative processes that may be particularly well suited for early Mars surface applications. One promising approach is the photocatalytic decomposition of carbon dioxide (Pipoli et al. 2000, 2002). This process is based on dissociating carbon dioxide with solar irradiation utilizing TiO_2 and Ni nanoparticles. Such a process could be operated in a static device and would not require electrical energy as driving force.

Carbon monoxide and oxygen produced in either of these approaches could be stored in compressed or cryogenic state and utilised for power generation applications utilising high-temperature fuel cell technology.

6.3.4 The Interrelation between Fuel Cells, ISRU and a Surface Power Plant

A crewed surface mission inevitably requires a sufficient, reliable and safe supply with electricity over months and years prior, during and inbetween crewed surface

operations to ensure the availability of life support systems, technical and scientific equipment. The required electrical power levels are even further increased if an ISRU approach is (at least partly) driven by electricity (e.g. electrolysis of carbon dioxide and/or water). Propellant storage in compressed or cryogenic state also requires a considerable power input for compression, liquefaction and storage.

State-of-the-art crewed surface mission scenarios therefore often include a small nuclear power plant with static or dynamic conversion systems for surface power generation. Such a reactor could provide a continuous electrical power output in the order of 200 kW in a very compact and robust package. Alternatively, non-nuclear technology could also be utilised for power generation. Photovoltaics and wind turbines are both generally feasible options, but very challenging for such an application because their power output literally depends on wind and weather, and a continuous power production cannot be guaranteed with either of these technologies without the burden of including a strong energy storage subsystem; not to mention erection and maintenance of the huge array of photovoltaic panels or wind turbines required to provide an average of 200 kW (Fraser, 2001; Baker 2002).

Assuming that the base power plant is able to provide sufficient output power to run all of the stationary systems associated with a crewed mission, some challenges in power system design still remain: how to power all the mobile and disseminated surface systems? How to store the 1000+ kWh single-trip energy required by a mobile pressurized laboratory? And how to provide backup power to vital systems in case the power supply of the stationary plant is disrupted?

For the case of fuel cell operation, two basic approaches are possible: firstly, the classical refuelling approach, where mobile mission elements are refuelled with (landed and/or ISRU-derived) fuels and oxidants at a stationary (ISRU production and storage) plant. Depending on the fuel/oxidant combination, the reaction product of the mobile fuel cell system could either be vented (e.g. in case of a carbon monoxide/oxygen system) or returned to the propellant plant (e.g. water directly produced in a hydrogen/oxygen fuel cell or condensed out of the product gas stream of a high-temperature fuel cell).

The second option would be to install a closed-loop regenerative fuel cell system aboard the mobile mission elements, and to simply utilise the electricity generated by the base power plant for recharging, utilising the onboard electrolyser of the mobile mission element. Thus, no exchange of species would have to be made between the refuelling/ISRU plant and the mobile or disseminated mission elements. Conversion losses due to the limited efficiency of regenerative fuel cell systems would then have to be compensated by a corresponding surplus production of the stationary power plant.

6.4 Possibilities and Limitations of Fuel Cell Power Systems in Mars Surface Applications

Robotic missions have been - and still are - the key element in the early science-driven stage of Mars surface exploration. Stationary (e.g. Viking 1 & 2 in 1975,

Pathfinder 1996) as well as mobile robotic missions (Sojourner with Pathfinder in 1996, Mars Exploration Rovers Spirit and Opportunity since 2004) have, for instance, provided valuable information about surface geology and atmosphere.

As the preparations for crewed missions proceed, stationary and mobile robotic systems will also be used in advanced human pre-cursor missions. One of these challenging missions will be the sample return mission, where soil and rock samples will be collected on the surface of Mars and returned to Earth in a small Earth-return capsule.

Upon arrival of the first elements of a crewed surface base on Mars, robotic systems will assist in base erection and thus provide the technical and scientific environment required by the first crew.

The crewed surface base will consist of a habitat, an Earth return vehicle, a stationary power plant for power generation (nuclear plant; non-nuclear technology as option), most likely an ISRU facility for propellant production as well as many different systems for surface transport and a wide range of technical and scientific equipment (Hoffman and Kaplan 1997; Baker 2002; Cougnet 2002; von Richter 2002, Fraser et al. 2004).

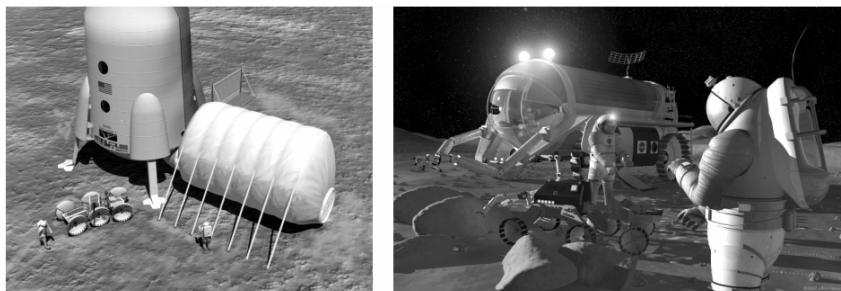


Fig. 6.13 Artist's impressions of two-storey lander habitat, inflatable laboratory and unpressurised rover (left) and remote operations with a pressurized rover (right) (reproduced with kind permission of Johnson Space Center/NASA)

Stationary mission elements (e.g. habitat, greenhouse module, ISRU plant) will be directly connected to one or more stationary power plant(s). A fuel cell system could nevertheless be included for intermittent - e.g. diurnal - energy storage or for peak and backup power applications in stationary systems.

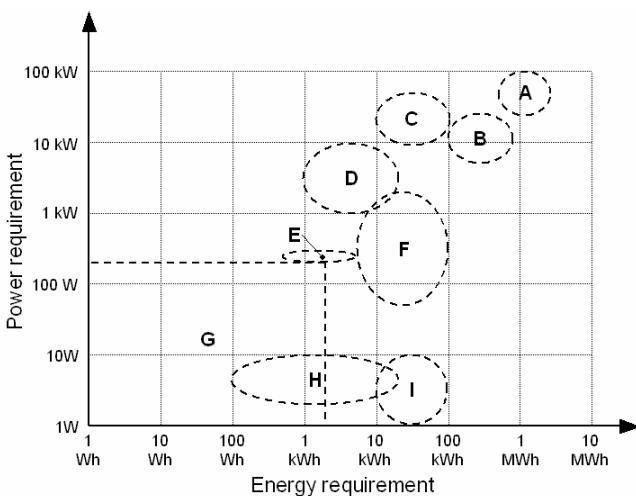
Mobile and disseminated mission elements, on the other hand, will not be directly powered by the base power plant; these mission elements are the primary candidates for fuel cell power systems.

6.4.1 Overview of Potential Fuel Cell Applications

Estimates of the power and energy requirements of mission elements considered as primary candidates for fuel cell power systems are shown in Fig. 6.14. Indices A-I of Fig. 6.14 are explained in Table 6.3.

Table 6.3 Abbreviations applied in Fig. 6.14

Index	Mission element
A	Pressurised mobile laboratory
B	Unpressurised utility truck
C	Habitat auxiliary power system
D	Small APU system
E	EVA suit
F	Robotic exploration rover
G	Power tools
H	Nano-rovers
I	Remote sensors

**Fig. 6.14** Power and energy map of selected Mars surface mission elements

6.4.2 Possibilities and Limitations of Fuel Cells

Engineers involved in (terrestrial as well as in space) power system design normally have a selection of different technologies they can choose from. In case of the upcoming Mars surface missions this will e.g. most likely include different primary and secondary batteries, photovoltaic panels, radioisotope units with different conversion systems and different fuelled power systems options (Fraser 2001). Primary and secondary fuel cell systems with different fuels and oxidants and with different storage technologies will be among the fuelled power system options.

Some of these power system technologies will not be suitable for specific applications because they require large surface areas (e.g. the geometrical area available for mounting photovoltaic arrays on a surface rover is limited) or because they are not compatible with the harsh environmental conditions expected during operation.

It is beyond the scope of this chapter to provide a comprehensive evaluation of the possibilities and limitations of fuel cell systems compared to all rival power

generation and energy storage technologies. Four key aspects are therefore discussed in the following, which are absolutely essential to understand the possibilities and the limitations of fuel cell power systems in Mars surface applications: conversion efficiency of primary and secondary fuel cell systems and specific energy as well as energy storage density of relevant fuel/oxygen combinations in selected storage systems.

Knowing the efficiency limits and the mass and volume requirements of relevant fuel and oxygen storage options, the initial question whether or not a fuel cell system could perform favourably in a certain application can be answered quite easily.

6.4.3 Fuel-to-Electric Conversion Efficiency of Fuel Cells

The primary fuel cell system efficiency defines how efficient the electrochemical conversion of fuel and oxidant into electrical energy is. In general, the efficiency of a fuel cell is directly proportional to the operational output voltage E (Volt). The higher the operational output voltage, the more of the chemical energy content of a fuel is available as electrical energy after the conversion process.

Three different efficiency values are normally considered with fuel cells: the efficiency versus lower and higher heating value (LHV and HHV) and the electrochemical efficiency:

$$\eta_{\text{LHV}} = E / E_{\text{LHV}} \quad (6.18)$$

$$\eta_{\text{HHV}} = E / E_{\text{HHV}} \quad (6.19)$$

$$\eta_{\text{electrochem}} = E / E_{\text{rev}} \quad (6.20)$$

The fuel cell efficiencies versus LHV and HHV (η_{LHV} and η_{HHV}) give the percentages of the calorific value of the fuel transformed into electrical energy according to Eqs. (6.18) and (6.19). For the case of hydrogen, the theoretical fuel cell voltages available if the heating values could be completely converted into electrical energy are 1.25V for the LHV (E_{LHV}) and 1.48V for the HHV (E_{HHV}). Dividing the operational fuel cell output voltage by these two theoretical voltages provides a measure for the conversion efficiency. This efficiency value is particularly helpful when comparing fuel cells to rival fuel-to-electric conversion technologies such as heat engine powered generators.

The electrochemical efficiency ($\eta_{\text{electrochem}}$) is another way of defining the efficiency of the conversion process. This efficiency is derived from reversible cell voltage (E_{rev}) and gives the fraction of the theoretical electrical energy output effectively available in operation according to Eq. (6.20).

All three efficiency values are shown for the example of a typical terrestrial PEMFC in Fig. 6.15 again as a function of the normalised current density (current/voltage data is derived from a laboratory-scale PEMFC run on hydrogen/oxygen at 70°C and terrestrial ambient pressure operation).

The greyed area indicates the fuel cell output voltage labelled on the Y-axis located on the right hand side, lines indicate the three efficiency values labelled on the y-axis located on the left hand side. In this example, a normalised current

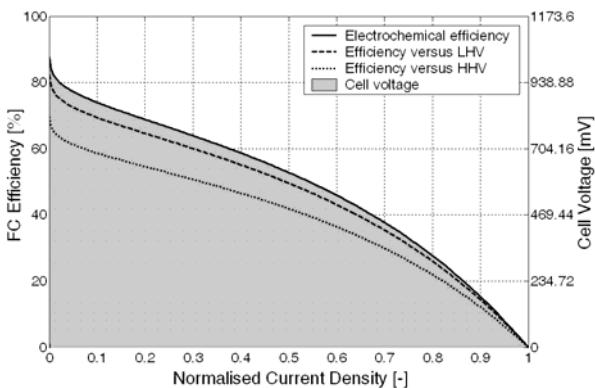


Fig. 6.15 Fuel cell output voltage versus conversion efficiency

density of 0.4 corresponds to a fuel cell output voltage of 690 mV, which is a good average for the output voltage achieved in state-of-the-art terrestrial applications in medium output power operation. In this specific case, the value 690 mV correspond to the following values of efficiency: 59% electrochemical efficiency, 55%/47% thermodynamic efficiency versus LHV/HHV.

A fuel cell operated in space applications will definitely be optimised with respect to materials and design and therefore achieve higher operational output voltages and efficiencies than a cost-optimised system for terrestrial applications. Conversion efficiencies considerably higher than 50% can therefore be achieved even in high output power operation with optimised fuel cell systems; even higher efficiencies can be achieved in part-load operation where both, fuel cell output voltage and efficiency, tend to increase if efficient auxiliary components are utilised.

The efficiency values given above correspond to the fuel-the-electric efficiency of the electrochemical reaction. The actual operational efficiency of the full system will be reduced by auxiliary components (e.g. blowers, control systems, PMAD). These parasitic loads reduce the available net output power of the fuel cell system and therefore also have to be considered in a full comparison of system efficiencies.

In general, a primary fuel cell is a very efficient means of converting the chemical energy of a fuel into electrical energy, keeping in mind that the figures are for fuel-to-electric conversion, and that these figures have to be compared to the combined efficiency of heat engine plus generator.

6.4.4 Electric/ Electric Efficiency of Regenerative Fuel Cells

The issue of conversion efficiency is of even greater importance with a regenerative fuel cell system, where two conversion steps are required to convert electrical energy into chemical energy of a fuel in an electrolyser, and to re-convert this chemical energy back into electrical energy in a fuel cell.

Modern secondary battery systems based on lithium ion technology provide a very energy-efficient means of storing energy. In general, these high electrical/electrical efficiencies can normally not be achieved with a regenerative fuel cell system.

When considering the electrical/electrical efficiency of a regenerative fuel cell system, only one efficiency value is considered. This is due to the fact that the efficiency of a regenerative fuel cell system gives the fraction of the electrical energy supplied to the electrolyser that is available as electrical energy from the fuel cell after the intermittent storage in fuel and oxidant. We refer to this efficiency as *electrical/electrical efficiency*.

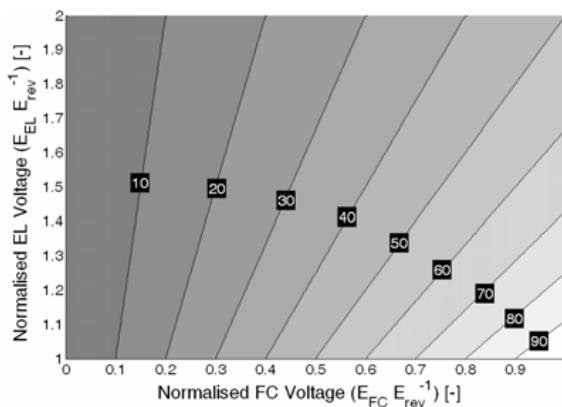


Fig. 6.16 Electrical/electrical efficiency of a regenerative fuel cell system

Figure 6.16 shows the variation of the electrical/electrical efficiency as a function of the operational voltages of electrolyser and fuel cell. The cell voltages are normalised to the one common reversible cell voltage. Assuming that the fuel cell voltage is 30% smaller and the electrolyser is 30% higher than the reversible voltage, the electrical/electrical efficiency is in the order of 54% excluding losses for auxiliary equipment, intermittent fuel and oxygen storage and input-/output-PMAD. This efficiency is reduced down to 33% if the operational cell voltages deviate 50% from the reversible voltage.

A deviation in the order of 30% from the reversible cell voltage is quite a reasonable estimate for fuel cell and electrolyser operation. A rough estimate for the electrical/electrical efficiency will therefore be in the order of 50-65%, considering an optimised system with efficient auxiliary components and energy-efficient energy storage systems for fuel and oxidant.

A regenerative fuel cell can therefore normally not match the high electrical/electrical efficiencies offered by other energy storage technologies such as secondary batteries or high-performance flywheels. Regenerative fuel cells are nevertheless an interesting option for energy storage applications, considering that they offer a very mass-efficient way of storing large quantities of energy. The possibility of storing large quantities of energy in a very mass- and volume-efficient way is the most important argument for the use of regenerative fuel cell systems despite of their comparably low electrical/electrical efficiencies.

6.4.5 Specific Energy and Energy Density of Relevant Fuels

Specific energy and energy density are two of the most important figures when it comes to storing energy in applications where mass and volume are strongly limited. Any reductions in mass and volume are always welcome in space applications, as this cuts down launch costs and/or increases the performance of the investigated system. This is not only true for mission elements launched from Earth, but also for mission elements refuelled and/or recharged on the surface of Mars.

Storing the 1000+ kWh of energy required by a manned pressurised laboratory, for instance, is not an easy task if the storage space is limited and each additional kilogram reduces range and mobility of the rover. Primary as well as secondary fuel cell systems can offer very high specific energies and energy densities, particularly if the ratio of stored energy and output power is high. This makes them attractive options for energy storage applications, considering that fuel cells are a very efficient means of converting a fuel/oxidant-combination into electricity.

The overall mass of a fuel cell system can be derived by summing up the masses of fuel and oxidant including storage vessels, the mass of the fuel cell and – if available – the electrolyser plus any auxiliary subsystems required e.g. for thermal control or PMAD. The first issue to be discussed in the following is the energy content of relevant fuels.

Table 6.4 presents an overview of key data for the three most important fuels for Mars surface applications considered in state-of-the-art ISRU scenarios: hydrogen, carbon monoxide and methane. Figures are given for the liquid (cryogenic) state as well as for gaseous state at pressure levels of 300 and 700 bars.

Cryogenic and compressed storage are comparably well-developed and proven in terrestrial and aerospace applications and can therefore be referred to as baseline against which innovative storage technologies can be evaluated. The figures given in Table 6.4 are for pure species excluding storage systems and electrochemical conversion system(s).

A selection of hydrocarbons is also included in this overview. This includes the alcohols methanol and ethanol, the two Liquid Petroleum Gases (LPGs) butane and propane and the two higher hydrocarbons heptane and octane. All of these hydrocarbons can be stored in light-weight vessels and can be easily converted into fuel cell grade synthesis gas by a catalytic fuel processor. All of these hydrocarbons could theoretically be supplied from Earth and/or produced in an ISRU system if the energy capacity and/or the stored energy density was strongly in favour of using hydrocarbons rather than hydrogen, carbon monoxide or methane.

Storage in liquid (cryogenic) state is indicated with (l) and compressed gaseous storage is indicated with (g) and the respective storage pressure given in bars.

A realistic comparison of these fuel options is only possible when the mass of oxygen required for stoichiometric combustion of the fuel as well as mass and volume of the storage vessels for fuel and oxygen are included in this compilation.

The following assumptions are applied: a 15wt% (weight-percent) storage vessel for hydrogen – a figure achievable by advanced small-scale cryogenic and pressure storage systems currently under development for terrestrial automotive applications – is used as reference with all of the species considered in this study.

Table 6.4 Properties of relevant fuels and oxygen

Species		Density [kg m ⁻³]	Boiling temp. [°C] ^b	Specific energy [kWh kg ⁻¹]	Energy density [kWh l ⁻¹]
H ₂	Hydrogen (l)	71	-252.8	33.3	2.4
H ₂	Hydrogen (g) 300 bar	24 ^a	-252.8	33.3	0.8
H ₂	Hydrogen (g) 700 bar	57 ^a	-252.8	33.3	1.9
CO	Carbon mono. (l)	789	-191.6	2.8	2.2
CO	Carbon mono. (g) 300 bar	339 ^a	-191.6	2.8	0.9
CO	Carbon mono. (g) 700 bar	791 ^a	-191.6	2.8	2.2
CH ₄	Methane (l)	423	-161.6	13.9	5.9
CH ₄	Methane (g) 300 bar	194 ^a	-161.6	13.9	2.7
CH ₄	Methane (g) 700 bar	453 ^a	-161.6	13.9	6.3
CH ₃ OH	Methanol (l)	790	64.5	5.6	4.4
C ₂ H ₅ OH	Ethanol (l)	790	78.0	7.5	5.9
C ₃ H ₈	Propane (l)	582	-42.1	12.8	7.4
C ₄ H ₁₀	Butane (l)	601	-0.5	12.8	7.7
C ₇ H ₁₆	Heptane (l)	680	98.0	12.2	8.3
C ₈ H ₁₈	Octane (l)	700	126.0	12.5	8.8
O ₂	Oxygen (l)	1130	-183.0	n/a	n/a
O ₂	Oxygen (g) 300 bar	387 ^a	-183.0	n/a	n/a
O ₂	Oxygen (g) 700 bar	904 ^a	-183.0	n/a	n/a

^aComputed for 298.15 K and ideal gas law^bAt ambient terrestrial pressure (1.01325 bar)

The weight-percentage of fuel stored in a vessel is scaled linearly with the molar mass of the stored species. This is a simplified but nevertheless feasible approach, considering that the mass of a pressure storage vessel, for instance, will be primarily determined by the storage pressure and not by the density of the stored species.

A tank for storing 30 kg of hydrogen has a computed mass of 200 kg (15 wt% of a 200 kg vessel is 30 kg), whereas a tank capable of storing the energy-equivalent 72 kg of methane only has a mass of 60 kg utilising the estimation technique presented above. Although this only provides rough estimates of the storage vessel mass, it nevertheless catches the trend very well.

The oxygen storage systems are chosen according to the fuel storage systems; hence, cryogenic storage of oxygen for liquid/cryogenic fuels, and 300 or 700 bar storage of oxygen for pressurized fuel storage, respectively.

Results of the mass and volume estimations required for storing 1000 kWh (versus LHV and excluding the conversion system efficiency) are compiled in Table 6.5. Specific energy and energy density are also computed for each fuel/oxygen

combination to provide a straight-forward means of comparing the energy storage properties with secondary battery systems and other energy storage techniques.

The figures presented in Table 6.5 do not give the useable energy content – which would have to include the efficiency of the fuel-to-electrical conversion system – but only the stored chemical energy of the fuel, more specifically the lower heating value (LHV). Mass and volume of fuel cell, electrolyser and auxiliary equipment as well as the efficiency of the conversion process are also neglected in these figures. Assuming a fuel cell operating at 67% fuel-to-electrical conversion efficiency, mass and volume of the storage system would have to be increased by 50% and specific energy and energy density would have to be correspondingly reduced because only a certain fraction of the chemical energy content of the fuel is effectively available as electrical energy.

Comparing the estimated figures with the two data sets for secondary batteries included in Table 6.5 clearly shows the potential of fuelled power systems despite of the fact that the fuel-to-electric conversion efficiency is neglected. In addition, the figures taken for the secondary batteries are long-term performance targets defined by the United States Advanced Battery Consortium - USABC (USABC 2008; Chalk and Miller 2008) and do not reflect the real status of battery technology.

Thus, specific energy and stored energy density of fuelled systems are far beyond the figures available or expected for rival energy storage techniques such as

Table 6.5 Comparison of storage system mass and volume

Species	Storage system mass [kg]	Storage system volume [l]	Specific energy [Wh kg ⁻¹]	Energy density ^a [Wh l ⁻¹]
Hydrogen (l)	568	697	1760	1435
Hydrogen (g) 300 bar	568	2030	1760	493
Hydrogen (g) 700 bar	568	870	1760	1150
Carbon monoxide (l)	825	702	1213	1424
Carbon monoxide (g) 300 bar	825	1752	1213	571
Carbon monoxide (g) 700 bar	825	701	1213	1427
Methane (l)	540	467	1851	2142
Methane (g) 300 bar	540	1224	1851	817
Methane (g) 700 bar	540	525	1851	1906
Methanol (l)	638	513	1567	1949
Ethanol (l)	567	456	1765	2192
Propane (l)	505	424	1979	2357
Butane (l)	494	416	2024	2405
Heptane (l)	501	412	1997	2426
Octane (l)	487	398	2053	2510
Secondary battery ^b	6667	4348	150	230
Secondary battery ^c	5000	3333	200	300

^aAssuming the storage vessel consumes an additional 10% of the storage volume

^bFigures correspond to the USABC minimum goal for EV batteries

^cFigures correspond to the USABC long term goal for EV batteries

batteries and flywheels now or in the foreseeable future. The specific energies of the investigated hydrocarbon fuels in combination with oxygen and storage vessels is in the same order as the values computed for hydrogen and methane; only carbon monoxide offers a significantly lower specific energy. This mass-penalty in operation has to be weighed against the potential reduction of the landed mass due to the omission of landing a feedstock species as required with hydrogen-based fuels and ISRU production systems.

The stored energy density clearly shows that high pressure storage requires significantly more space than liquid/cryogenic storage even if storage pressures of up to 700 bar are applied. The stored volume of power systems based on liquid hydrocarbons is considerably smaller than the volume required by a hydrogen, carbon monoxide or methane storage system.

The choice of the best fuel and oxidant option has to be made on the basis of a comprehensive analysis of mission scenario, ISRU options and mass and volume constraints in the actual application. Choosing a prudent and well-integrated solution for power and propellant production, a combination of a fuelled power system and a highly-efficient primary or secondary fuel cell system offers very interesting possibilities for many Mars surface mission elements.

6.4.6 Fuel Cell Power System Mass and Volume Estimates

Mass and volume estimates of secondary battery and fuel cell power systems are compared in the following. The comparison is based on specific energy and energy density estimates presented in Table 6.5, assuming a low-temperature PEMFC operated with pressurised hydrogen/oxygen and a secondary battery system complying to the long term goals of the USABC for electric vehicle applications (USABC 2008; Chalk and Miller 2008).

Mass and volume estimates are computed for a 100 kW peak output power system as it could e.g. be used in a mobile pressurised laboratory. Mass and volume of fuel cell and battery power system are computed as a function of the available net electrical energy output.

The limited fuel-to-electric conversion efficiency is considered with the fuel cell system; limitations in the depth-of-discharge and the energetic discharge efficiency are considered with the secondary battery system. Efficiency values of fuel cell and secondary battery include losses and auxiliary loads required by control and diagnostic systems as well as fuel cell balance-of-plant-components (e.g. blowers and humidification subsystem).

Mass, volume and efficiency of power management and distribution (PMAD) systems are neglected. This assumption is reasonable as the PMAD components of the fuel cell and battery modules would both be designed for the identical peak output power of 100 kW.

Mass and volume of the thermal control systems (TCSs) can hardly be estimated by simple mass and energy balances. The useable window of operating temperatures and the issue of waste heat utilisation, respectively, have to be specifically evaluated for the investigated cell technologies and applications. Mass and volume of the TCS components are therefore also neglected in the estimates presented below.

Table 6.6 Parameters applied with the mass and volume estimations

Property	Value	Unit
Specific energy (700 bar H ₂ /O ₂ including storage vessels)	1760	Wh kg ⁻¹
Energy density (700 bar H ₂ /O ₂ including storage vessels)	1150	Wh l ⁻¹
Specific power H ₂ /O ₂ fuel cell ^a	1000	W kg ⁻¹
Power density H ₂ /O ₂ fuel cell ^a	1000	W l ⁻¹
Fuel-to-electric conversion efficiency fuel cell ^{b,c}	65	%
Specific energy secondary battery (USABC long term goal)	200	Wh kg ⁻¹
Energy density secondary battery (USABC long term goal)	300	Wh l ⁻¹
Depth-of-Discharge secondary battery	80	%
Energetic discharge efficiency secondary battery ^d	90	%
Electrical power output (fuel cell and battery) ^e	100	kW

^aIncluding auxiliary components, excluding PMAD and TCS

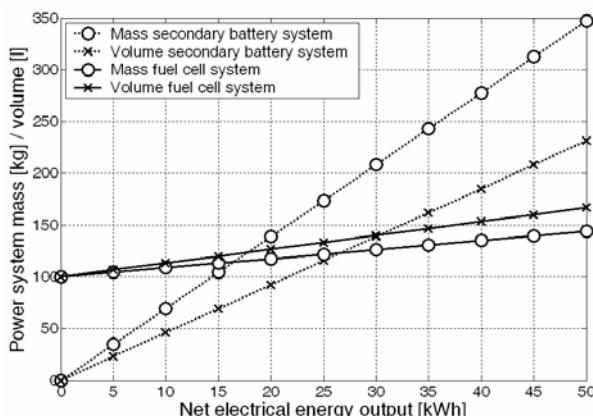
^bNet electrical energy output versus LHV of the fuel

^cCorresponds to an operational cell output voltage (PEMFC) of about 810 mVs

^dFraction of theoretical energy content available as electrical energy

^eUsed for sizing the fuel cell module

Simulation parameters applied with the computation of the mass and volume estimates are compiled in Table 6.6. Mass and volume of the fuel cell power system are computed as the sum of the fuel cell module (i.e. fuel cell stack including auxiliary components but without hydrogen/oxygen storage system) and the hydrogen/oxygen storage system. An average operational fuel-to-electric conversion efficiency is used to scale the stored energy content of fuel/oxygen with respect to the net electrical energy output. Mass and volume of the battery system are estimated by considering depth-of-discharge (DoD) and discharge efficiency with the long-term battery performance goals of the USABC for electric vehicle battery systems (USABC 2008; Chalk and Miller 2008).

**Fig. 6.17a** Comparison of mass and volume estimates for secondary battery and fuel cell power systems (range: 0-50 kWh net electrical energy output)

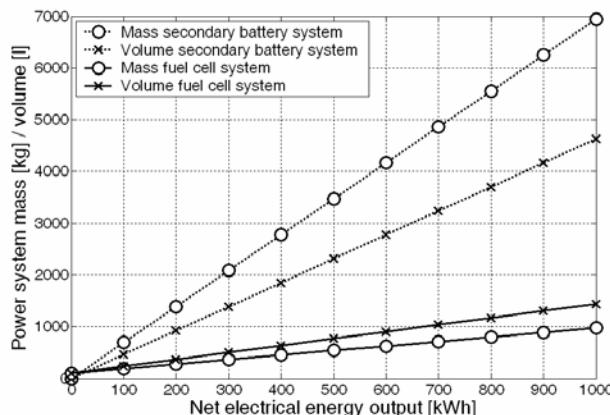


Fig. 6.17b Comparison of mass and volume estimates for secondary battery and fuel cell power systems (range: 0–1000 kWh net electrical energy output)

Mass and volume estimates of the secondary battery and fuel cell power systems are plotted in Figs. 6.17a and 6.17b for two different electrical energy output ranges.

Two characteristic features can be derived from these mass and volume estimates: due to the fact that mass and volume of the fuel cell (i.e. fuel cell stack plus auxiliary components excluding hydrogen/oxygen species plus storage vessels) are independent of the net electrical energy output, but only a function of the output power of 100 kW, the fuel cell system is both heavier and larger if only a small electrical energy output is required.

Moving towards higher values of electrical energy output, mass and volume of the fuel cell system increase far less than the values computed for the secondary battery system. This is due to the high specific energy and the high energy density of the fuelled system. This energy content is very high despite of the fact that the chemical energy content of the stored species is almost twice as large as the net electrical output due to the limited efficiency of the fuel-to-electric conversion.

If more than 17 kWh of electrical energy are drawn from the power systems, the fuel cell system becomes lighter than the secondary battery equivalent despite of having to carry the additional burden of the fuel cell stack plus auxiliary components on top of the stored mass of hydrogen and oxygen. Pressurised storage of gases is commonly linked to comparably large storage volumes. This is the reason why the fuel cell power system only becomes more compact than the secondary battery system if the net electrical energy output exceeds 30 kWh.

Approaching a net electrical energy output of 1000 kWh, significant reductions in power system mass and volume can be achieved with a fuel cell system. The fuel cell system reduces the overall power system mass in the order of 85% and the power system volume is reduced by almost two-thirds compared to the secondary battery system.

These mass and volume estimates – although computed with strongly simplified energy balances – clearly demonstrate the possibilities of fuel cell power systems in high energy output applications.

6.5 Summary

There are many operational characteristics and design features inherent to fuel cell systems that make them an interesting option for a wide range of applications in space exploration.

Fuel cells can be built from simple, repeating elements with few moving parts; this will make them highly reliable and long lasting even under harsh operating conditions. Fuel cells are very efficient; their superior fuel-to-electrical conversion efficiency will reduce mass and volume of power systems. Specific energy and energy density of fuel cell systems can be significantly higher than the design targets defined for future battery and flywheel energy storage systems. Mobile mission elements could thus be equipped with a full energy supply prior to launch from a Mars surface base instead of relying on battery recharging procedures during mission time, as it would be the case with a combined photovoltaic and secondary battery power system.

The vision of producing fuels and oxidants directly on the surface of Mars is one of the key aspects in favour of the use of fuelled power system options. On-site fuel and/or oxidant production capabilities could either be achieved by producing oxygen from regolith or ambient atmosphere and methane from a feedstock of hydrogen brought from the Earth. Thus, a significant mass leverage with respect to the Earth launch mass could be achieved. Alternatively, decomposing carbon dioxide taken from the Martian atmosphere into carbon monoxide and oxygen could also provide a very robust in-situ resource utilisation approach without relying on a feedstock of species brought from Earth.

Provided that fuels and oxidants are produced on location for propelling ascent and/or Earth-return vehicles, the utilisation of these readily-available fuels and oxidants in power generation applications suggests itself. This requires an efficient and reliable fuel-to-electrical energy conversion technology; future fuel cell technology holds the promise of providing these features.

Fuel cells are even attractive in a scenario where a small nuclear power plant is installed for surface power generation. In this case, a regenerative fuel cell system could be used to store electrical energy in a very mass- and volume-efficient way, and thus help to reduce power system mass and volume in mobile and stationary backup systems.

The question to be asked is therefore not *if* fuel cells are going to be part of future Mars exploration missions, but rather *when* and *how* they are going to be utilized.

Acknowledgements

I would like to thank Mr. Martin Lang and Dr. Max Schautz (ESA/ESTEC) for previous collaborations, NASAImages.org for their assistance in obtaining the

NASA images and the Joint-Stock Company Ural Electrochemical Integrated Plant (UEIP) for the kind permission to reproduce the picture of the Photon fuel cell system.

References

- Baker, A.M.: Future power systems for space exploration: executive summary. ESA Contract #14565/00/NL/WL (2002)
- Barin, I.: Thermochemical data of pure substances. VCH, Weinheim (1989)
- Bockstahler, K., Funke, H., Lucas, J., Witt, J., Hovland, S.: Design status of the closed-loop air revitalisation system ARES for accommodation on the ISS. In: SAE 48th International Conference on Environmental Systems, San Francisco, USA, SAE document number 2008-01-2189 (2008)
- Chalk, S.G., Miller, J.F.: Key challenges and recent progress in batteries, fuel cells, and hydrogen storage for clean energy systems. *J. Power Sources* 159, 73–80 (2006)
- Cohen, A. (ed.): Report of the 90-day study on human exploration of the Moon and Mars. NASA-TM-102999 (1989)
- Coons, S.C., Williams, J.D., Bruckner, A.P.: Feasibility study of water vapor adsorption on Mars for In Situ Resource Utilization. In: 33rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Seattle, USA, July 6-9, 1997, AIAA paper number 97-2765 (1997)
- Cougnet, C.: European Mars system architectures study: executive summary. ESA Contract #14566/00/NL/WK (2002)
- Fisher, A.C.: Electrode dynamics. Oxford University Press, Oxford (1996)
- Förstner, R.: Characterisation of a regenerative solid oxide fuel cell for Mars applications. Diploma thesis, Universität Stuttgart (1998)
- Fraser, S.D.: Non-nuclear power system options for a mission to mars and derived terrestrial applications. Diploma thesis, Graz University of Technology (2001)
- Fraser, S.D.: Mathematical modelling of polymer electrolyte membrane fuel cells for stack design and power system analysis. Ph.D. thesis, Graz University of Technology (2004)
- Fraser, S.D., Monsberger, M., Hacker, V.: Fuel cell power system options for Mars rovers. In: Proc. 2nd International Conference on Green Propellants for Space Propulsion, Sardinia, Italy, June 7–8 (2004)
- Fraser, S.D., Hacker, V., Besenhard, J.O.: Sponge iron process for manned space exploration: final report. ESA Contract #18461/04/NL/MV (2005)
- Grover, M.R., Bruckner, A.P.: Water vapor extraction from the Martian atmosphere by adsorption in molecular sieves. In: 34th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Cleveland, USA, AIAA paper number 98-3301, July 13-15 (1998)
- Hamann, C.H., Vielstich, W.: Elektrochemie, 3rd edn. Wiley-VCH, Weinheim (1998)
- Hendrickx, B., Vis, B.: Energiya-Buran: The Soviet Space Shuttle. Springer, Heidelberg (2007)
- Hoffman, S.J., Kaplan, D.I. (eds.): The reference mission of the NASA Mars exploration study team. NASA Special Publication 6107 (1997)
- Kaplan, D.I.: Mars ISRU sample return (MISR) – a mission description. *Lunar and Planetary Science* 27, 641–642 (1996)
- Kordesch, K., Simader, G.: Fuel cells and their applications. Wiley-VCH, Weinheim (1996)
- Larminie, J., Dicks, A.: Fuel Cell Systems Explained, 2nd edn. John Wiley & Sons, Chichester (2000)

- Linne, D.L.: Carbon monoxide and oxygen combustion experiments: a demonstration of Mars in situ propellants. NASA-TM-104473 and AIAA PAPER 91-2443 (1991)
- Linne, D.L.: A rocket engine for Mars sample return using in situ propellants. NASA-TM-107396 and AIAA-97-0893 (1997)
- Linne, D.L.: Demonstration of oxygen and carbon monoxide propellants for Mars missions. NASA Document ID 20050179358 (2005)
- Linne, D.L., Roncace, G., Groth, M.F.: Mars in situ propellants: carbon monoxide and oxygen ignition experiments. NASA-TM-103202 and AIAA-90-1894 (1990)
- National Space Council, Human exploration of space: a review of NASA's 1990-day study and alternatives. National Academy Press, Washington (1990)
- NASA online, Mars fact sheet (2009) <http://nssdc.gsfc.nasa.gov/planetary/factsheet/marsfact.html>
- Pipoli, T., Besenhard, J.O., Schautz, M.: Feasibility of a CO/O₂ fuel cell to be used on Mars. In: Proc. 2nd Conf. Academic and Industrial Cooperation in Space Research, Graz, Austria, November 15-17 (2000)
- Pipoli, T., Besenhard, J.O., Schautz, M.: In situ production of fuel and oxidant for a small solid oxide fuel cell on Mars. In: Proc. 6th European Space Power Conference (ESPC), Porto, Portugal, May 6-10, 2002, pp. 699–704 (2002)
- Schautz, M., et al.: Results of tests on the Photon fuel cell generator. In: Proceedings of the European Space Power Conference, Graz, Austria, August 23-27, 1993, pp. 653–658 (1993)
- UEIP, Homepage of the Joint-Stock Company Ural Electrochemical Integrated Plant (UEIP) (2008), <http://www.ueip.ru>
- USABC, Goals for advanced batteries for EVs. Downloaded from the USABC energy storage systems goals (2008), <http://www.uscar.org>
- von Richter, A.: Automation and robotics for human Mars exploration (AROMA): executive summary. ESTEC/Contract 14563/oo/NL/WK (2002)
- Warshay, M., Prokopiou, P.R.: The fuel cell in space: yesterday, today and tomorrow. *J. Power Sources* 29, 193–200 (1990)
- Williams, J.D., Coons, S.C., Bruckner, A.P.: Design of a Water Vapor Adsorption reactor for Martian in situ resource utilization. *Journal of British Interplanetary Society* 48, 347–354 (1995)
- Witt, J., Hovland, S., Bockstahler, K.: ARES – ESA's regenerative air revitalisation experiment on the International Space Station. SAE 48th International Conference on Environmental Systems, San Francisco, USA, SAE document number 2008-01-2093, June 29 - July 2 (2008)
- Zubrin, R., Wagner, R.: The case for Mars: the plan to settle the red planet and why we must. Touchstone, New York (1996)

Chapter 7

The Evolution of Nuclear Power Generation for Mars

Liviu Popa-Simil

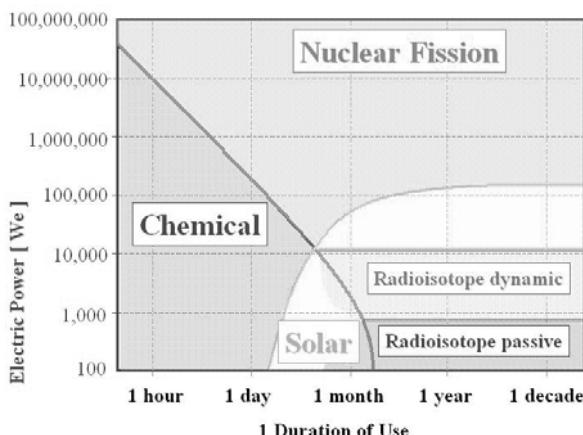
LAVM Co., Los Alamos, USA

7.1 Introduction

Among the available energy alternatives nuclear power offers important advantages and in many cases is the only viable alternative given actual operation conditions on Mars. We know that nuclear is the most compact form of energy available. Nuclear power is required at every step of space exploration as a backup form of energy ready to be delivered when all other energy sources cease to deliver. Unfortunately, it also has a number of drawbacks.

Figure 7.1 shows one-selection criteria for the available power sources on Mars, i.e. power as a function of duration of use for various sources of energy. The use of chemical sources of energy becomes prohibitive in the case the chemicals are not produced locally. The only remaining alternatives are solar and nuclear. Fission reactions are used when high power is needed; or radioisotope sources, mainly alpha emitters, when stable moderate power is required, with low complexity over a large period of time. Missions are usually longer than a few months, which make chemical sources mostly prohibitive on Mars, especially where both reactants have to be stored and used together. For solar powered applications, one also has to know that the sun irradiance on Mars is about $\frac{1}{2}$ of that on the Earth at same incidence angle (see Chap. 2).

Fig. 7.1 Domains of selection of energy sources



7.2 Goals of Mars Exploration

7.2.1 Reconnaissance Missions to Mars

The first US explorations were performed by the Mariner 9, followed by the Viking landers in the 1970's, continued in by the Mars Exploration Rovers (MER) Spirit and Opportunity as part of initial unmanned missions searching for water (Jet Propulsion Laboratory 2004). NASA plans to send human precursor missions by 2011 to prepare for a man landing by 2030 (for details see Chap. 1). There are many issues to fix such as: landing, outpost habitat, resources exploitation, communication and recovery (NASA 2004). Challenging requirements have been set for both space transportation and outpost systems that have to minimize the transit time and the astronaut's radiation exposure to assure longer duration stays.

The main enabling technology for Mars missions is nuclear power. Nuclear power for space applications is not a new idea, as there are presently 1 US and 33 Russian non-operational fission reactors in space. There are two types of Nuclear power: Radioisotope Thermoelectric Generators (RTGs) and fission sources (Fig. 7.2).

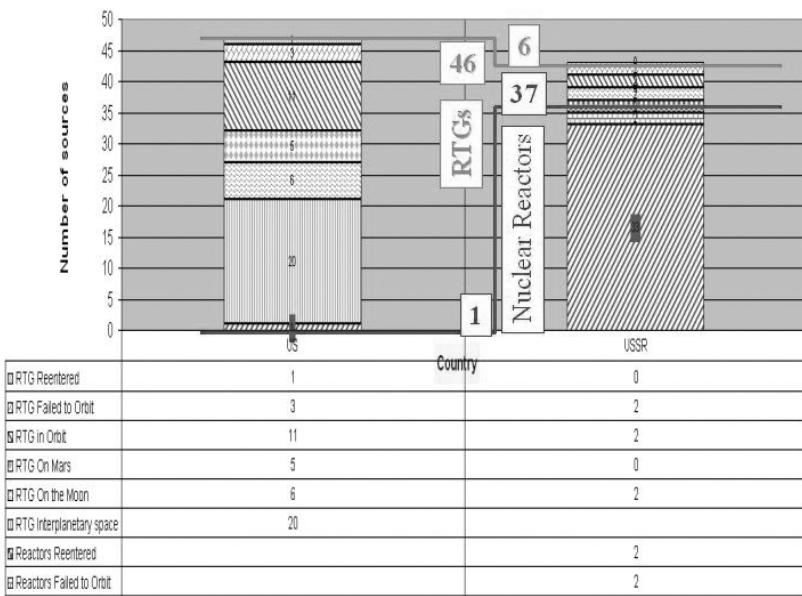


Fig. 7.2 Nuclear power sources in space

RTG is powered by alpha radioisotopes such as ^{238}Pu with a half-life ($T_{1/2}$) of 87.7 years and a specific power of 0.5 W/g, or ^{210}Po with $T_{1/2}=138$ days giving a specific power of 140 W/g. As the radioisotope decays, it auto-absorbs its radiation thereby heating up. Up to 7% of that heat is converted to electrical power by thermo-electrics. The rest of the heat flows through the structure and is radiated to space by the RTG's radiator fins. RTGs were the actual nuclear power source used

on the Moon in the Apollo program, along with the Galileo, Cassini, and Voyager missions delivering a specific power in the range of 4.5-10 W_e/kg (El-Genk 2002). Exploration vehicles need between several kW up to few MW given by “the road curve” that depends on mass, terrain conditions, cross sections, and gasodynamic coefficients making fission power a realistic alternative. This is also the case for vehicles on Mars, where the irradiance is rather low (see Chap. 2).

Nuclear Thermal Propulsion (NTP) and Nuclear Electric Propulsion (NEP) may deliver both thrust and electric power on board.

The reactor coolant, typically hydrogen, flows from the storage tank to the nozzle, cooling the nozzle, and then flows through the reactor, where is heated to approximately 2700 K and expelled through the nozzle as rocket propellant generating thrust. This is a 1960’s concept developed in the KIWI and NERVA programs with the advantage of having a high specific impulse (up to 900 second long), high thrust, and a simple design with minimal moving parts. Its main disadvantage is high fuel temperatures, leaving limited abort scenarios in a Mars mission.

More advanced designs separate the nuclear reactor and thrust generator, by using the nuclear power entirely to produce electricity, and then using the electricity in specialized electric propulsion devices (Dudinski 2004). This concept has been developed in the NASA SP-100 programs (NASA 2004) following the success obtained on the SNAP-10A test platform, the equivalent of the Russian TOPAZ (Angelo and Buden 1985).

A review of the main features of nuclear power sources showing overall specific mass versus electric power output favors the development of nuclear fission reactors (McGinnis 2004).

7.2.2 Rovers and Robots

One of the most significant successes of the Mars missions are NASA’s twin rovers, Spirit and Opportunity. They seem to be the robots that never quit (Malik 2009).

The two probes were each built for a 90-day romp across the Martian surface, a mission that began when Spirit bounced to a stop on the planet’s expansive Gusev Crater five years ago today. Opportunity touched down on the other side of Mars a few weeks later and now, after half a decade, both rovers are still exploring the red planet after surviving more than 20 times their planned lifetime (other details are given in Chap. 1).

7.2.3 Outposts on Mars

The location of any outpost is of exceptional importance, and one actual strong choice is the North Pole because of the interesting temporal and spatial variations it exhibits compared with the rest of Mars:

- Steeper temperature and pressure than the rest of Mars;
- Reduced winds, wind gradients, and dust.
- More seasonal frost growth.
- Possible ancient sea-bed.

- Ideal region for underground water.
- Ideal region to search for fossilized/existing life.
- Low elevation which is much better for aircraft.
- Smoother terrain than the southern hemisphere that assures safer rover traverses and aircraft landings.
- The terrain seems better for base construction and is very smooth by terrestrial standards.
- Is known to contain water, which makes it potential future-fuel production area.
- It offers a constant summer site with continuous 220 sols, acceptable temperature, more tranquil weather, more imagery, and more time to erect the base and achieve operability.

7.3 Power Options on Mars Surface

NASA has considered several options to power a Mars outpost (Marshall 2004):

- Solar array by modifying the solar power module for mid-latitude operations
- Radioisotope thermoelectric generator (RTG)
- Nuclear reactor.

The data for this study comes from a previous robotic polar outpost study that utilized 3 kW of solar power, in a near polar location provided with 220 sols of continuous sunlight in the Martian year. During the winter eclipse period, a majority of the polar base was shut down. The use of a nuclear power source will enable continuous operation of the polar base station. In this approach, NASA considered the following:

- Assumed 3 kW power requirements and a near polar base location
- Assess the impacts of moving the outpost to the mid-latitudes
- Investigate the operational feasibility of solar, RTG, and nuclear reactors
- Use “HOMER” (Heat Pipe Operated Mars Exploratory Reactor) nuclear reactor design as the initial basis for a outpost reactor. A RTG and a reactor have roughly same mass for 3 kW_e, but RTGs provide the lowest technical risks if provided plutonium challenges (cost/production/ perception) are overcome.

Nuclear power has a significantly higher power weight efficiency (kWhr /Martian yr/kg) than solar power, while solar power is optimized for summer daylight only operations (Table 7.1).

7.3.1 Radioisotope Based Power Sources

Other potential units are:

- RHU Radioisotope Heating Units
- RTG Radioisotope Thermoelectric Generator
- GPHS General Purpose Heating Systems
- SRG Stirling Radioisotope Generator
- MMRTG Multi-Mission RTG

Table 7.1 Comparative Performance of Solar and Nuclear Power for Mars

System	Power output (kW _e)	System Mass (kgS)	Specific Power (W _e /kgS)	Power weight efficiency over Martian Year (kW-hr/kg)	Benefits	Constrains
SNAP-10	0.65	435	1.49	24.62		
HOMER	3	775	3.87	63.79	Better HEDS support	Large development
Topaz-2	10	1060	9.43	155.46		
SP-100	100	5000	20	329.58		
MMRTG	3	728	4.12	67.89		
SRTG	3	711	4.22	69.54	Low risk, Modular availability	²³⁸ Pu availability
Polar Array	3	363	8.26	8.26	Available technology	Dust management
Mid Latitude sleep mode	3	1177	2.55	22.36		
Mid Latitude continuous	3	2095	1.43	27.24		Oversized for summer

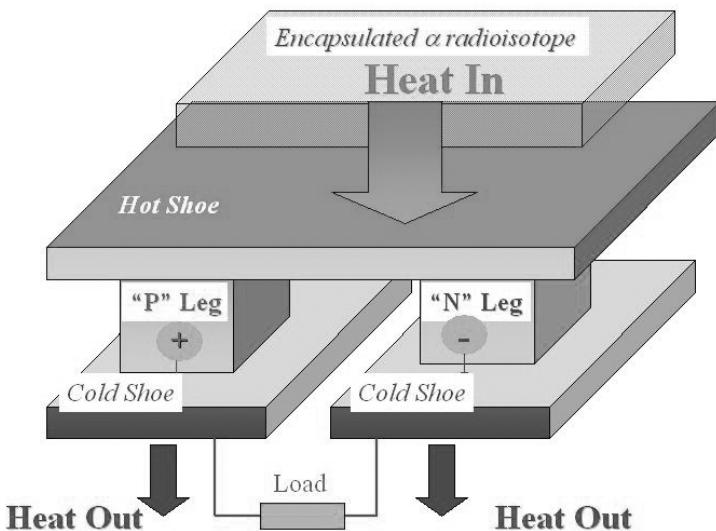
**Fig. 7.3** RTG – block diagram

Figure 7.3 shows the essential elements of a RTG. The encapsulated radioisotope, usually an alpha emitter (i.e. ²³⁸Pu, ²⁴¹Am, ²¹⁰Po, etc.), warms up auto absorbing its own radiation up to temperatures of 700-900K applied on the hot shoe.

The heat flows through the legs towards the cold shoes, creating a potential voltage between them. The efficiency of such devices is between 5-10%, forcing the rest of the heat to flow through the legs to the cold shoes only to be discharged into a vacuum compatible heat sink.

One example is a RTG with the following characteristics:

- Fuel: Plutonium 238.
- Specific Power: 25 W/kg Plutonium.
- Specific RTG system Power: 3-5 W/kg.

RTG was proven in space missions as Cassini that used 3 GPHS for 870 Watt. The 44 RTGs that have powered 24 spacecrafts were very expensive and scarce. Needs a continuation of improvements in:

- Thermoelectric conversion
- Dynamic conversion (Stirling engine)
- Lighter refractory materials
- Lighter structural materials
- Production lines for ^{238}Pu .

7.3.2 HOMER (Heat Pipe Operated Mars Exploratic Reactor)

HOMER has the following characteristics: Power: $15 \text{ kW}_t - 3 \text{ kW}_e$; Dimensions: 2.5 m diam x 2.4 m length (12 m^3); Mass of Fissile Material: 73 kg (5 kg/kW_t, 24 kg/kW_e); Total Power System Volume: 48 m^3 ; Reactor Mass: 214 kg; Total Power System Mass: 775 kg; Sodium heat-pipe cooled reactor; Stirling engine energy conversion: 30%; Specific Reactor thermal Power: 70 W/kg; Specific System electric Power: 3.8 W_e/kg; Specific Power-Volume: $0.25 \text{ kW}_e/\text{m}^3$.

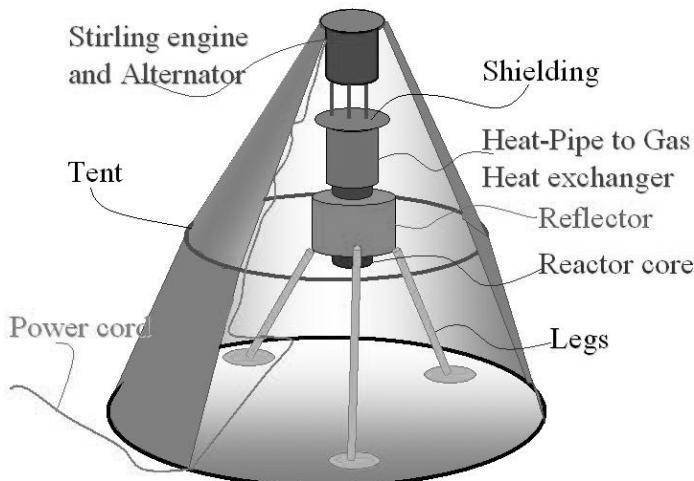


Fig. 7.4 HOMER nuclear reactor

The following future improvements are expected:

- Reduce weight (but not shielding effectiveness)
- Address re-entry shield for launch abort
- Prometheus type is a focal point for space power reactors
 - Jettison when under the way
 - Advocate for surface reactor development

Any mission should exploit space nuclear reactors in Mars orbit to support power intensive science, such as HEDS and Prometheus. Figure 7.4 shows the placement of HOMER on Mars. It shows the electric power cord going to an outpost several hundred meter away from the reactor. The protection tent is intended to insulate the reactor against sand storms, being aerodynamically profiled and maintaining a good IR radiative surface.

7.3.3 Comparative Evaluation of the Power Sources

Figure 7.5 shows the Power-Mass efficiency for various power structures already designed to be used on Mars. Nuclear Reactors are the most likely power sources to be developed since they provide an increase in power without a commensurate increase in mass. The solar power mass efficiency increase shows a limitation with increasing power output, while the isotope powered batteries maintain a constant ratio.

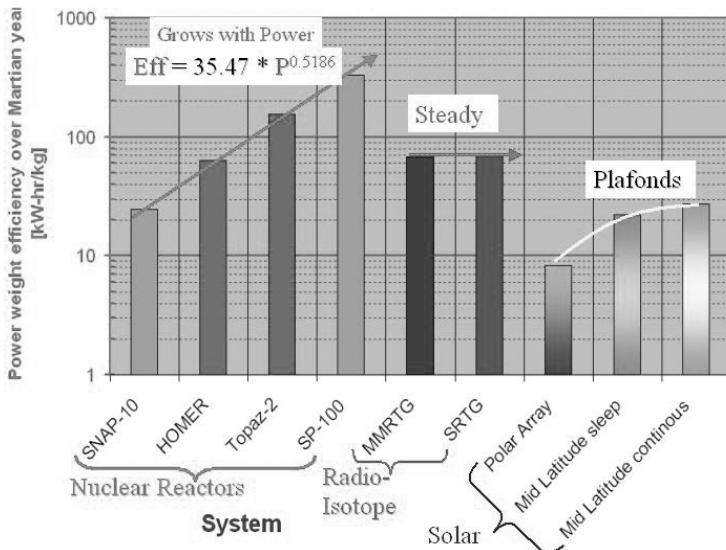


Fig. 7.5 Power-mass efficiency for various power structures

Figure 7.6 shows the relationship between the electric power output and specific power. Photovoltaic structures are at a disadvantage because they have to be over dimensioned to produce and store the required power depending on the outpost position on the planet. Nuclear power is advantageous because its delivery is

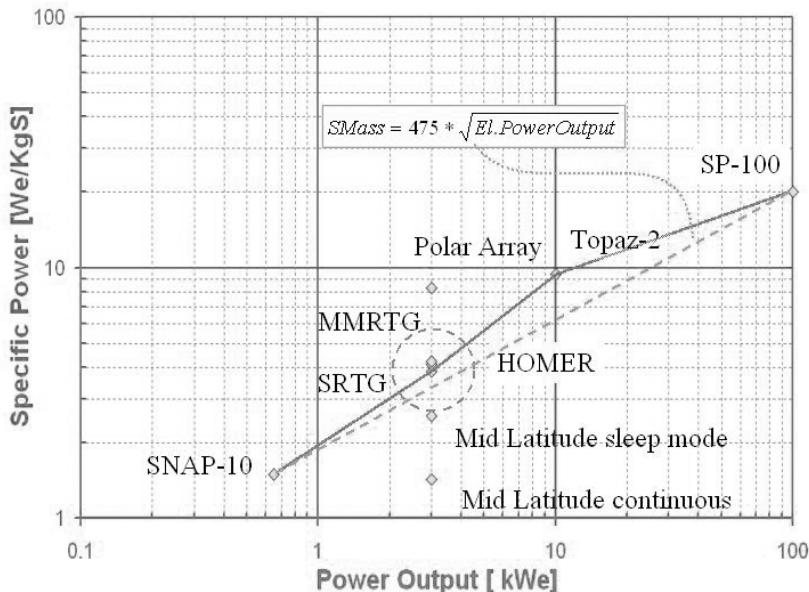


Fig. 7.6 Specific power versus power output evaluation

determined by the power request and the maximum electric power it is able to deliver. A comparison of potential power systems has been done by NASA and JPL (Marshall 2004) which considers the diversity of units needed to be powered in future missions such as: Landers, Rovers, power modules, Flyers, and Shelters. Each has specific power requirements as part of the Mars polar outpost planned to be built by NASA between 2016 and 2030.

NASA's six-wheeled Mars Science Laboratory (MSL) is designed to gather specimens of Martian soil and rock and is scheduled to be launched in 2009 (Golombek 2006). It will then analyze those samples for organic compounds and gauge environmental conditions that may have supported microbial life on that planet, now or in the past.

Promising MSL landing sites will receive extra-special attention, not only by using data accumulated by NASA's MRO, Mars Odyssey, and the Mars Global Surveyor, but also the European Space Agency's Mars Express. Given all that research, the touchdown point of MSL in 2010 will be the most studied locale on Mars in space exploration history. Nuclear-powered MSL is built to live longer than the smaller, solar-energized robotic twins now trekking about Mars—Spirit and Opportunity.

However a crash landing could lead to MSL's radioisotope power source polluting water ice in the region. In addition, "care has to be taken in regards to putting Earth microbes in the places where they can grow and thrive," said John Rummel, Planetary Protection Officer at NASA Headquarters (Washington, D.C.). MSL's nuclear power source will stay hot for a long time, he added, and if it

comes in contact with Martian ice, would produce a warm little pond and might make a comfy home for microbes.

The Mars Science Laboratory (MSL), depending on the pathway, would investigate potential habitats, e.g., water-laden sedimentary rocks or a hydrothermal deposit (active or fossil), and study the geochemistry and mineralogy of the site in detail. If previous missions suggest that Mars was never globally wet, MSL would focus on exploring departures in the evolution of the terrestrial planets on the basis of in-situ analytical tools (Mars Science Program Synthesis Group 2003).

In 30 years Russia plans to establish a base on Mars. The deputy chief engineer of Red Star (a state scientific company closely affiliated with Russia's Nuclear Energy Ministry) says the station will be constructed in the mountainous areas of Mars, possibly in one of the canyons, and it will be nuclear powered (Liss 2003).

7.3.4 Nuclear Power Option for Space

Nuclear systems are well suited for use at planetary/space bases because:

- They are not affected by solar proximity or orientation;
- Do not require power storage to operate through nights;
- Are compact;
- Have a high specific power;
- Are virtually non-radioactive at launch;
- Can remain sub-critical during all credible launch accidents;
- Can scale to very high power levels;
- Have the potential for a very long lifetime;
- Can operate in dusty environments, and
- Can be used as a source of high-quality heat, in addition to electricity.

Nuclear power sources (Bennett 2006) have enabled or enhanced the most challenging space missions. In addition, radioisotope heater units have enhanced or enabled the Mars exploration rover missions Sojourner, Spirit, and Opportunity. Since 1961, the United States has successfully flown 41 radioisotope thermoelectric generators (RTGs) and one reactor to provide power for 24 space systems. The former Soviet Union has reportedly flown at least 35 nuclear reactors and at least two RTGs to power 37 space systems.

The former Soviet Union also had an active space nuclear program, launching about 31 (and perhaps 33) BUK space nuclear reactors to power spacecraft used for marine radar observations. The BUK liquid-metal-cooled fast reactor produced about 3 kWe from 100 kW_t of thermal power using two-stage thermoelectric elements. Launches of BUK may have begun as early as 1967, with the last publicly known launch occurring in 1988 (Bennett 1989; Ponomarev-Stepnoi 2000).

In 1987, the former Soviet Union launched two experimental thermal reactors designated as TOPAZ. These experimental reactors, which used an in-core thermionic conversion system, reportedly produced about 5 kW_e from a reactor thermal power of 150 kW_t (Bennett 1989). In the late 1980s and early 1990s, another Russian thermionic reactor (ENISEY) was tested, but had not been flown

Table 7.2 History of Mars missions (NASA 2008)

Date	Name	Country	Result	Reason
1960	Korabl 4	USSR	Failure	Didn't reach Earth orbit
1960	Korabl 5	USSR	Failure	Didn't reach Earth orbit (flyby)
1962	Korabl 11	USSR	Failure	Earth orbit only; spacecraft (flyby) broke apart
1962	Mars 1	USSR	Failure	Radio Failed
1962	Korabl 13	USSR	Failure	Earth orbit only; spacecraft (flyby) broke apart
1964	Mariner 3	US	Failure	Shroud failed to jettison
1964	Mariner 4	US	Success	Returned 21 images
1964	Zond 2	USSR	Failure	Radio failed
1969	Mars 1969A	USSR	Failure	Launch vehicle failure
1969	Mars 1969B	USSR	Failure	Launch vehicle failure
1969	Mariner 6	US	Success	Returned 75 images
1969	Mariner 7	US	Success	Returned 126 images
1971	Mariner 8	US	Failure	Launch failure
1971	Kosmos 419	USSR	Failure	Achieved Earth orbit only
1971	Mars 2 Orbiter/Lander	USSR	Failure	Orbiter arrived, but no useful data and Lander destroyed
1971	Mars 3 Orbiter/Lander	USSR	Success	Orbiter obtained approximately 8 months of data and lander landed safely, but only 20 seconds of data
1971	Mariner 9	US	Success	Returned 7,329 images
1973	Mars 4	USSR	Failure	Flew past Mars
1973	Mars 5	USSR	Success	Returned 60 images; only lasted 9 days
1973	Mars 6 Orbiter/Lander	USSR	Success/Failure	Occultation experiment produced data and Lander failure on descent
1973	Mars 7 Lander	USSR	Failure	Missed planet; now in solar orbit.
1975	Viking 1 Orbiter/Lander	US	Success	Located landing site for Lander and first successful landing on Mars
1975	Viking 2 Orbiter/Lander	US	Success	Returned 16,000 images and extensive atmospheric data and soil experiments
1988	Phobos 1 Orbiter	USSR	Failure	Lost en route to Mars
1988	Phobos 2 Orbiter/Lander	USSR	Failure	Lost near Phobos
1992	Mars Observer	US	Failure	Lost prior to Mars arrival

Table 7.2 (continued)

Date	Name	Country	Result	Reason
1996	Mars Global Surveyor	US	Success	More images than all Mars Missions
1996	Mars 96	USSR	Failure	Launch vehicle failure
1996	Mars Pathfinder	US	Success	Technology experiment lasting 5 times longer than warranty
1998	Nozomi	Japan	Failure	No orbit insertion; fuel problems
1998	Mars Climate Orbiter	US	Failure	Lost on arrival
1999	Mars Polar Lander	US	Failure	Lost on arrival
1999	Deep Space 2 Probes (2)	US	Failure	Lost on arrival (carried on Mars Polar Lander)
2001	Mars Odyssey	US	Success	High resolution images of Mars
2003	Mars Express Orbiter/Beagle 2 Lander	ESA	Success/ Failure	Orbiter imaging Mars in detail but lander lost on arrival
2003	Mars Exploration Rover - Spirit	US	Success	Operating lifetime of more than 15 times original warranty
2003	Mars Exploration Rover- Opportunity	US	Success	Operating lifetime of more than 15 times original warranty
2005	Mars Reconnaissance Orbiter	US	Success	Returned more than 26 terabits of data (more than all other Mars missions combined)
2007	Phoenix Mars Lander	US	TBD	

(Ponomarev-Stepnoi 2000). In addition to nuclear reactors, the former Soviet Union has used radioisotope heater units (RHUs) and small RTGs (Bennett 1989).

Twenty-three from thirty-eight missions to Mars (see Table 7.2 and Chap. 1) failed mainly for launching and navigation 80%, 10% for communication failures and 10% for Mars Lander failures. The main problem of success stories is the lifetime on Mars mainly limited by the power sources availability both to the missions orbiting Mars and especially for those landing on Mars.

7.3.5 Space Propulsion

7.3.5.1 Electric Propulsion

The application of Nuclear Electric Propulsion (NEP) to space missions has been a topic of increasing interest (Elliott 2008). NEP systems appear ideally suited for a range of deep space missions where high delta-V and high power at the target bodies are enabled by the use of nuclear power and electric propulsion systems. Somewhat less obvious, however, are the benefits of NEP for inner planet

missions to Mars or Venus, or other near-Earth objects (NEOs) where chemical propulsion and solar power have proven adequate in the past. However the utility of NEP vehicles in the inner solar system is greatly enhanced when the versatility, longevity, and reusability of such a system is considered.

NASA, JPO, and DoE started to develop a NEP Interplanetary Transfer Vehicle; a “Space Truck” designed for delivery of payloads from Earth to a variety of destinations, including Mars and Venus, dependent on mission needs. NASA proposes using electrical ion propulsion powered by a nuclear reactor for its Jupiter Icy Moons Orbiter, an element of Project Prometheus, which is scheduled for launch after 2011 (Danneskiold 2008). Shortening the travel to Mars duration requires a better engine with less mass penalty, same as those for interstellar precursor missions and those which stretch our technical (Lipinski 1999) capabilities in the directions needed for later interstellar travel (Malhotra 1999, Jewitt 1999).

Nuclear electric propulsion shown in Fig. 7.7 has the advantages of flexibility in design and flight-proven hardware, but the disadvantages of complexity of design and a big heat rejection subsystem putting value in the amount of money, time, and research already done on various designs (McGinnis 2004).

One possible power system is a SNAP-10A derivative (using UZrH fuel) with a steam Rankine cycle conversion system. Another is gas-cooled reactor (using SP-100 developed UN fuel) with a Brayton conversion system, and the third is a heat-pipe cooled system (with UN or UO₂- fuel) and a Brayton conversion system. None of these options requires any new fuels or materials development.

7.3.5.2 Plasma Propulsion

Magnetic fusion energy research suggests the use of some magneto-plasma configurations to address certain critical issues in the gas-core fission approach to nuclear-thermal propulsion (Gerwin 1995) in a spheromak configuration. The neutronics, criticality properties, and heat transfer properties of gas core fission plasmas for the purpose of nuclear thermal propulsion were studied over parameter regimes to achieve large Lundquist numbers that would open the possibility to utilize certain magneto-plasma configurations known in magnetic fusion energy research. Also, it was pointed out that the moderately larger dimensions considered here compared to those in Poston’s allow thicker propellant and neutron-moderator layers along with

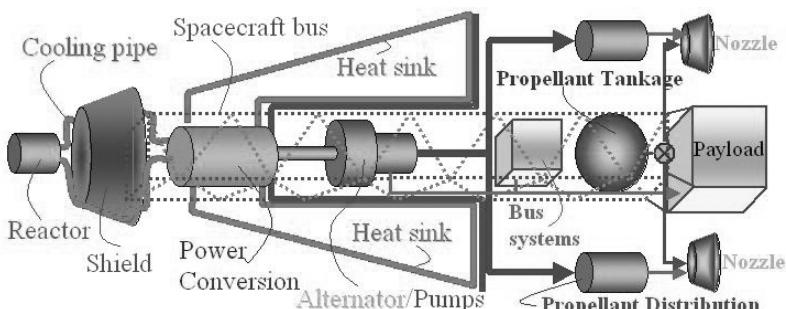


Fig. 7.7 Nuclear Electric Propulsion System

higher propellant temperatures (Poston 1994). Higher specific impulses may be attained; and concomitantly, that the increased energy upscattering of the thermal neutrons by the thicker hotter hydrogen propellant layer may be utilized with ^{233}U fuel to improve the criticality parameter k_{eff} over what it would be with ^{235}U . These enhanced thrust rocket engines are tremendously important to reduce the travel time to Mars.

7.4 Main Issues of Mars Nuclear Reactors

7.4.1 Conversion Cycles

Several procedures to convert the thermal energy released by a nuclear reactor into electrical power for space applications have been studied. These are dynamic conversion cycles such as: Rankine, Brayton, and Stirling cycles, and static conversion systems such as: Alkali-Metal Thermoelectric Converter (AMTEC), Thermoelectric, Thermionic, and Magneto-hydrodynamic. The static thermoelectric and thermionic conversion cycles are inherently less efficient than the dynamic cycles. The magneto-hydrodynamic cycle is highly efficient in theory, but in practice it is typically operated in an open mode that produces exhaust effluents that may interfere with payload operations.

7.4.1.1 *The Rankine Cycle*

The Rankine cycle is based on a working fluid that changes phase during the operation of the conversion cycle. The boiling and condensation phases allow absorption and extraction of heat at constant temperature, which allows this cycle to approach the ideal Carnot efficiency. The circulation of the working fluids is a fairly efficient process that does relatively little to detract from the efficiency of the cycle's operation.

There are two variants of this cycle:

Liquid Metal systems use metals with low melting and boiling temperatures such as mercury, cesium, potassium, sodium, or lithium (in increasing order of operating temperatures) as working fluids.

Organic Rankine systems use organic liquids such as terphenyl, pyridine (CP-32), monoisopropyl biphenyl (MIPB), or a eutectic of biphenyl and biphenyl ether (DowthermA). The potential for pyrolytic degradation of these fluids restricts their use to operating temperatures considerably below those of liquid metal systems.

There are two alternative implementations of this cycle:

Direct Rankine systems consist of a liquid (typically potassium in liquid metal reactors) that is vaporized by the reactor or other power source, passes through turbines generators to produce electric power, and is then condensed back to the liquid phase for recirculation through the system.

Indirect Rankine systems use a liquid (typically lithium in liquid metal reactors) that is heated by the reactor or other power source, then passes through a heat exchanger to vaporize a working fluid (potassium in liquid metal reactors) in a secondary loop,

which in turn is passed through turbine and a condenser. The inclusion of a heat exchanger in implementation of the indirect Rankine cycle adds additional weight and complexity compared to the direct cycle, but avoids the difficulty of assuring in-reactor liquid-vapor separation in the microgravity environment of space.

The primary advantage of the Rankine Cycle is high efficiency and isothermal heat rejection, which minimizes the area of the radiator for a given operating temperature. This system can provide useful levels of power output at relatively low operating temperatures, with radiators operating at relatively high temperatures, compared to the Brayton Cycle.

The primary drawback of the Rankine Cycle is the corrosive and erosive effects of the liquid metal working fluids. In particular, the use of lithium poses a risk of fire or explosion should leaks in the coolant loop expose the lithium to ambient water contaminants.

In addition, the simplest Rankine cycle, which would use a single loop circulating the working fluid from the reactor through the power turbines and the radiators, could expose the entire conversion system to the radioactive potassium coolant. Rankine cycle systems use two coolant loops connected by a heat exchanger, with the secondary loop circulating non-activated fluid through the power turbines and radiators. A third loop may also be added to drive hydraulic systems. These additional loops add to the mass and complexity of Rankine systems. The fact that the Rankine cycle is based on the use of working fluids that are present in both liquid and vapor phases adds a further element of complexity. In-core boiling raises concerns about critical heat flux and reactivity effects, as well as about erosion and corrosion. Particular attention is required to address issues such as control of liquid-vapor interfaces and achieving adequate liquid vapor separation, as well as management of fluid transport and fluid inventory location. These matters are resolved with relative ease in terrestrial systems, but become difficult in the microgravity environment of space.

7.4.1.2 Brayton Cycle

The Brayton Cycle uses a single-phase gaseous working fluid that absorbs energy at a constant pressure through a rise in temperature operated in either open or closed mode. In the open Brayton Cycle, gas is exhausted into the environment after passing through the power generating turbines. Closed Brayton Cycle passes the gaseous working fluid through power generating turbine, after which it is circulated through a radiator for cooling and reuse. The liquid fluids used in the Rankine cycle have heat-transfer coefficients approximately 50 times higher than the gaseous fluids used in the Brayton cycle which gives it a lower thermal efficiency than the Rankine Cycle. The efficiency is further compromised by the energy required to compress the large volume of high-pressure gaseous working fluid and more than one fluid circulation loop.

The attractive features of Brayton Cycle are:

- A large experience base and technological maturity.
- Simplicity of the single loop circulation of the single-phase gaseous working fluid.

- Open-cycle cooling can be more readily implemented.
- The non-corrosive properties of inert gases that may be used as working fluids.
- The cycle can accommodate higher source temperatures than Rankine.
- Can readily respond to rapid changes in power load requirements by adjusting operating pressure, while Rankine systems require longer response times due to the coolant's thermal inertia.

Brayton Cycle drawbacks, resulting from its low thermal efficiency are:

- Gas-cooled reactors require larger heat transfer surface areas than do liquid-metal cooled Rankine reactors, and thus typically require larger and more massive reactors.
- The cycle requires very high efficiencies in power and compressor turbine components.
- Inert gas working fluids may become contaminated with corrosive impurities, compromising system operations.
- Very high operating temperatures and pressures may be required to achieve satisfactory power outputs.
- High rotational speeds of primary generation turbines may require the use of either specially designed high-speed electrical generators, which may add to cost and technical risk, or complex reduction gearing assemblies with added mass and subsequent reduction in efficiency.
- For a given heat source temperature, a large radiator area is required.
- Are typically more massive than other cycles.

7.4.1.3 Stirling Cycle

The Stirling cycle is a closed-cycle reciprocating engine that uses a high-pressure single-phase gaseous working fluid, such as hydrogen or helium. The primary advantage of this cycle is its high thermodynamic efficiency, approaching theoretical Carnot efficiencies, with low operating temperatures. Most of the research was performed on kilowatt-class applications that may not be scaleable to megawatt-class operations. The use of large numbers of small engines is unpractical for high-power applications.

7.4.1.4 Alkali-Metal Thermoelectric Converter (AMTEC)

The Alkali-Metal Thermoelectric Converter (AMTEC) cycle, which is thermodynamically similar to the Rankine cycle, uses high-pressure sodium vapor supplied to one side of a beta-alumina solid electrolyte. The transport of sodium ions across this electrolytic membrane produces a voltage differential, with low-pressure sodium vapor emerging from the other side of the membrane that drives electrons through a useful load.

7.4.1.5 Thermoelectric Conversion

Thermoelectric systems rely on the direct conversion of thermal energy to electricity by means of the Seebeck effect that uses a temperature differential between two adjacent, dissimilar materials generating a voltage between them. The development

of semiconductors in the 1950s led to the fabrication of a range of intermetallic compounds and alloys that produced devices with conversion efficiencies of several percent being used for relatively low power levels. Thus thermoelectric converters are frequently used in Radio-isotope Thermoelectric Generators (RTGs), which provide up to a kilowatt of power for deep space probes.

Thermoelectric converter materials include PbTe, Pb-Sn-Te, and various Ge-Si alloys that have conversion efficiencies of 10% to 15%, but most limit operating temperatures to ranges that result in lower efficiencies. The mass of the converter system is small, but thermoelectric power systems require massive radiators and radiation shielding components. The practical thermoelectric systems net efficiency is about 3-5%.

7.4.1.6 Thermionic Conversion

Thermionic conversion systems are similar to an "electron boiler" in which electrons are thermally boiled off a heated emitter cathode and collected on an anode surface, thereby delivering electrical power to an external load.

There are three configurations of thermionic nuclear power systems:

- A cylindrical reactor core with thermionic converters on its outer surface, each converter having its own radiator that is limited to low power levels by its emitter area.
- Out-of-core configurations can achieve higher power densities and larger emitter areas by circulating a liquid metal coolant through the reactor core, which transfers the heat to an external panel of converters.
- In-core thermionic systems are composed of converters which are directly attached to individual fuel elements within the reactor core. Anode cooling is by a loop which circulates a coolant (typically using a liquid metal) to an external radiator.

Conversion efficiencies increase with emitter temperature, ranging from 5% at 900 K to over 18% at 1,750 K. In practice, efficiencies may range from less than 4% at 900 K to over 9% at 1,750 K. Urania fast reactors are attractive for maintaining the high temperatures and high power densities required for in-core thermionic systems.

Attractive features of thermionic systems include:

- Static operation with reduced vibration and improved reliability;
- High power density;
- Potentially high thermal conversion efficiency compared to thermoelectric systems;
- Very high heat-sink temperatures reduce radiator mass penalties;
- Liquid-metal (NaK) cooled thermionic reactors operate at lower temperatures than other metal-cooled concepts, and thus face reduced corrosion concerns.

Fuel element swelling, leading to structural deformation or radiation leakage may be a particular problem for thermionic reactors using lithium or sodium-potassium coolants that may also run risks of accidental fires.

Fuel fabrication for thermionic reactors is more difficult than for other reactor types since the fuel element includes an emitter, collector, insulators, spacers, sheath, cladding, and fission product purge lines. All these components must be manufactured to tight tolerances. For advanced thermionic concepts, this problem will be more severe because thousands of very small diodes will be required. The thermionic system's ability to maintain satisfactory operations over extended periods is unproven. Experience to date has demonstrated operating lifetimes of a few years.

7.4.1.7 Magnetohydrodynamic (MHD) Conversion

MHD conversion systems produce electrical power by using electrically conductive ionized plasma, which passes through a channel surrounded by a magnetic field perpendicular to the motion of the plasma that produces an electrical current. The power generated is proportionate to the channel volume, plasma velocity, and the magnetic field strength.

Advantages of a MHD conversion system include:

- High enthalpy extraction, which at high temperatures may surpass 50% conversion of thermal energy to electric power in open cycle systems, but practical efficiencies are in the 10% to 15% range.
- No moving parts that improves reliability and reduces vibration.
- Simplicity of operation and long-term storage in space.
- Favorable scaling to larger sizes.
- Almost instant turn on, important for burst-mode applications.
- High power densities, in the range of 5 - 7 kW/kg, though net power densities may range below 2 kW/kg.

The disadvantage of MHD systems is the large volume of high temperature effluents created during operations. The hot gas passing through the generator, hydrogen in the case of nuclear systems, or water vapor or other molecules in the case of chemical rocket systems, must be seeded with conducting materials such as lithium, potassium, or cesium. Cesium is more expensive than potassium but offers higher power densities. These materials, and effluents from chemical power systems, can interfere with payload operations if directly exhausted into the space environment by an open cycle MHD generator. The closed cycle systems avoid this problem but require large and massive effluent absorption and treatment systems.

MHD technology poses several practical problems in addition to its extremely high operating temperatures and the need to obtain adequate electrical conductivity for conversion process (National Research Council 1989). However the management of system effluents emerging from the channel is a problem that requires creating a highly ionized, high-velocity gas stream with adequate uniformity. Small non-uniformities in gas density and/or ionization concentration (conductivity) can result in major flow instabilities, and the excess heating in these regions causes acoustic disturbances and flow disruptions.

Given the disparate range of figures of merit on which systems may be evaluated, proponents of each system are able to point to at least some metrics that support the superiority of their favored approach.

Some general observations are possible:

- None of the proposed conversion cycles has demonstrated reliable long-life operations at multimegawatt power levels -- all require significant development efforts to reach these goals.
- Some conversion cycles, notably the Stirling (for lower power applications), and AMTEC and MHD (for multi-megawatt applications), require major technology improvements, given relatively limited terrestrial experience base.
- In the near-term, the Brayton cycle is the preferred approach for burst-mode open-cycle systems, and appears competitive with the Rankine cycle for closed-cycle steady-state systems.
- In the long run, the Brayton cycle will remain the system of choice, absent any major improvements in MHD conversion, for open cycle systems. Technology improvements may render the Rankine cycle preferable to the Brayton cycle for closed systems.

7.4.2 Radiator Concepts

The choice between open and closed-cycle power systems remained one of the central issues of the Multi-megawatt (MMW) program. For open-cycle systems, the dominant scaling factor is the duration of operation, with extended operation times requiring increasing amounts of expendable coolants. For closed-cycle systems, the primary consideration is the increasing radiator mass needed to accommodate higher power levels.

The selection of either open or closed cycles has direct implications for the choice between chemical or nuclear power sources. Although nuclear systems can be operated in either an open or closed mode, chemical power sources are typically more compatible with open-mode operation, except for long-duration operation as operating on Mars requires.

The effects of effluents from open-cycle operation on spacecraft and payload operation or on a surface outpost are the most significant issue facing such systems (Britt et al. 1989).

In order to prevent effluent from the power system from interfering with the operation, it is important to employ high velocity directional nozzles that produce highly directed jets of the effluent into areas of the space environment out of the field of action. The flow from such nozzles must have a high Mach number, $M = 5$ or greater, in order to achieve a relatively clear local environment, but it remains possible that a significant amount of effluent could back stream into critical areas.

MMW power requirements would entail significant advancement in radiator systems (Department of Energy 1986) to significantly higher power production and rejection rates, and demand the development of radiators and power conversion systems with very low specific weights.

Advanced radiator concepts exhibit relative technical immaturity. The choices among radiator concepts are complicated by a range of potential figures of merit (Begg 1987) that makes the competition between radiator systems concentrate not only on specific weight, but also on factors like deployability, survivability, geometry, maneuverability.

7.4.2.1 *Fin/Tube Radiators*

The advantage of these conventional radiator designs is their simplicity and the extensive existing technology base. Multi-megawatt applications would require massive radiator panels. The mass of radiator panels is increased by the need to protect the coolant loop from punctures by micrometeoroids and space debris or Mars' surface sand winds. A variant of this class is the Heat Pipe radiator that operates on the basis of an evaporation-condensation cycle, in which the gaseous circulating coolant is returned to the liquid phase as it passes through a heat exchanger connected to the radiator. The high effective thermal conductivity associated with this process results in a relatively low mass radiator system. The thickness and mass of the radiator surface is determined only by puncture hazard considerations, rather than thermal conduction requirements. Since each heat pipe element of the radiator is an independent unit that is not connected with the primary thermal management coolant loop, damage to a single element would not result in overall system failure.

Heat pipes have been used in low-power thermal management on spacecraft, but there is no flight experience base with systems for prime power thermal management. A long-standing technical issue is the efficiency of the thermal linkage between the primary coolant loop and the heat pipe evaporators (Mattick 1987).

7.4.2.2 *Membrane Radiators*

Membrane radiators consist of a thin flexible membrane inflated by low gas pressure, with coolant circulating on the inner surface of the rotating membrane. These radiators may use either spherical or disk membranes. In this concept (Koenig 1985) the fluid, a liquid or a condensable vapor, is introduced inside the sphere and impinges on the top and bottom surfaces. The fluid wets the inner surface of the sphere and in microgravity is driven in the form of a liquid film by centrifugal force to the equatorial periphery of the sphere; liquid metal pumps located there return the liquid out of the sphere through rotating shaft seals to its source. On Mars surface it will come towards the bottom under the gravity forces. As the liquid flows along the inner surface of the envelope it loses heat by thermal radiation from the outer surface of the balloon-like radiator.

The use of thin membrane materials in this radiator configuration offers significant reductions in system mass compared to conventional panel technology but are vulnerable to single-point failures and wind forces; therefore, neither the Spherical or Disk Rotating Membrane Radiator are easy usable on Mars' surface.

Another weak point of these assemblies is the rotating machinery elements, which could pose reliability concerns for long-life systems along with the rotating vapor seal between the radiator and the body that may prove a particularly problematic element.

7.4.2.3 *Belt and Filament Radiators*

These radiators are based on mechanical heat transport away from a heat exchanger by moving a thin (1mm) flexible material that radiates waste heat prior to return to the exchanger. The thermal load may be imparted to the moving radiator

either by direct contact or by convection. The advantage of these radiators is the reduced vulnerability to damage from micro-particles, since the exposed radiator belt or filaments are mechanically robust, while the more fragile heat exchanger is protected from exposure to these hazards. The belt may be folded prior to use, reducing storage concerns. These radiators are potentially less massive than conventional designs, but employ complex operational concepts that could pose significant testing and reliability problems. In addition, the large flexible belts would complicate maneuvering and slewing of platforms and would have difficult operation on Mars' surface and require special protection against dust storms.

7.4.2.4 Dust/Droplet Radiators

These radiators are operated by recirculating free streams of sub-millimeter particles to radiate waste heat directly into space (Mattick 1986). Interest in these concepts originated in the late 1970s in response to concerns about the vulnerability of conventional radiator designs to damage by micrometeoroids. The initial concept used solid dust particles, but problems in management of the particle inventory, as well as the low efficiency (except at high temperatures) of heating the radiator particles (via radiation rather than conduction) limited interest in this approach. This may be an interesting approach on Mars where dust is abundant, if advances in particle handling are made.

The liquid droplet concepts have the advantage of a very large surface to mass ratio of the radiating surface compared with more conventional radiators, which could potentially result in an order-of-magnitude reduction in thermal-management system mass. The total mass of the radiator system may be 4 - 8 times that of the droplet array itself. This concept can still produce significant mass savings relative to conventional systems. Its attractive features (Mattick 1982) include:

- Reduced vulnerability to dust, micro-meteoroid, space debris, compared to that of conventional systems;
- Low transport volume, if the dust on surface is used;
- Ease of deployment and stowage.

Major developmental challenges facing this class of radiators involve the complex hydraulics of the systems, including:

- Improvement in generation and precision direction of particle streams is required to reduce losses from misdirection and collision, which could produce contamination interfering with local operations; Surface winds and dust storms may also be an issue.
- Droplets contact electrification may result in significant stream deflection and droplet losses.
- Improved efficiency in particle stream collectors, to reduce losses and contamination, may pose serious challenges in the microgravity environment of space, and real challenges on Mars during the windy season.
- One of the most serious problems results from surface evaporation losses from the droplets, which may constrain maximum operating temperatures to below

500 K. Evaporation could require unacceptably massive coolant reserves for replenishment, and could produce contamination.

- Radiative efficiency of radiating particle streams may be relatively low. While silicone and lithium have satisfactory thermal properties, their use is constrained by evaporation losses. The emissivity of liquid metals (such as tin) that might be used in higher temperature systems is typically of the order of 0.1, resulting in an effective radiator emissivity of 0.2, in contrast to the 0.85 emissivity typical of aluminum used in fin radiators.
- Liquid droplet radiators using liquid metals such as tin could face very complicated shut-down, dormancy, and restart problems with freezing and thawing the metal coolant.

Concerns about droplet losses were the basis for proposals to use electrostatically charged droplets or particles, that would be confined by field lines established by charged generator and collector surfaces.

A new development on this concept is the Curie Point (Carelli 1989) radiator where, solid ferromagnetic metal particles are heated above their Curie Point and ejected in a stream into space. As these paramagnetic particles radiate heat, their temperature eventually drops below their Curie point, whereupon a magnetic field guides the now-magnetic particles into a collector. The Curie point radiator may potentially offer a lower mass solution to thermal management, but employs unproven concepts, with complex and potentially massive support systems. The large magnets required by this concept may result in systems with specific masses comparable to traditional radiator concepts. On Mars' surface this may have the sensibility of mixing with contaminant dust, but may offer more stable yet unproven operation in hard winds.

It is possible on Mars to use sand or regolith to produce dust particles as a cooling agent, but the contact with radiation areas (where transmutation to radioactive isotopes is possible), and abrasion processes have to be carefully considered.

7.4.2.5 Heat Removal Issues

Even for high efficiency energy conversion systems, the heating problem remains, because any other form of produced energy ends in thermal energy warming the shuttle or the outpost in undesired locations (Popa-Simil 2007). This thermal energy has to be managed and directed while the surplus must be eliminated in space, by radiation mainly obeying the Stephan-Boltzman formula.

If the surface is surrounded by other surfaces, the heat exchange becomes more complicated. For two finite surfaces due to the heat exchange between the surfaces depending on their emisivities, the transport factor among surfaces is equivalent to effective solid angles among the objects (the dimensionless Radiation Shape Factor). For the Mars environment with temperatures between 150-300 K, it will not be a problem to provide cooling, but good insulation is needed to keep the inside of an outpost warm to 290 K. The cogeneration solution seems to provide a good opportunity to remove the low-grade heat from the reactor, using IR heat sinks only when needed.

7.4.3 Materials

Material issues beginning with the fuel and ending with the truss structure determine the design and performance. Even the most common materials used in space nuclear power systems have to meet rigorous quality standards.

7.4.3.1 Alloys

The materials used in a nuclear reactor power system are required to meet the following properties (Zinkle 2004):

- Manufacturing capability, availability, and cost
- Mechanical and thermophysical properties
- Irradiation effects
- Chemical compatibility and corrosion properties
- Nuclear properties (such as cross section for neutron absorption)

The operation temperature range for high power space nuclear power system structures is in the 1300K-1400K range. Refractory metals based on Niobium, Tungsten, Molybdenum, Rhenium, and Tantalum are the most used. Super alloys do not have the enough strength at these temperatures and materials based on carbon/carbon and ceramics can be fabricated with difficulty into these complex systems (Ring 2004).

Niobium alloys

The main alloy, PWC-11 is made of Niobium with 1% Zirconium (Nb-1Zr) with 0.1 percent carbon (Buckman 2004). Creep testing has been performed on Nb-1Zr in 1960 for 173,000 hours and 34 other tests have been performed in 1986 for a total of 320,650 hours, a period equal with 37 years. PWC-11 was tested less than Nb-1Zr, but it increases the creep strength of Nb-1Zr, having less strength than Nb-1Zr.

In irradiation testing, Nb-1Zr becomes brittle at less than 800K. At temperatures above 1100K the elongation of Nb-1Zr has experimentally been shown to become significant. Nb-1Zr is compatible with many liquid metal coolants and has a small cross section for fast neutron absorption.

Tungsten alloys

Tungsten is used as a shielding material having a high density for gamma rays, being at the same time unsuited for any other space nuclear power system applications. Tungsten has low fracture toughness even above its ductile to brittle transition temperature (DBTT). Low temperature irradiation tests of tungsten shows it has a severe radiation hardening embitterment below 1200K (Zinkle 2004). Tungsten is difficult to fabricate and weld.

Molybdenum Alloys

In the 1960s, Mo-TZM (Mo-0.5% Ti-0.1% Zr-0.03% C), Mo-TZC (Mo-1%Ti-0.3%Zr-0.15%C), and Mo-13%Re were tested for use as cladding, but molybdenum

is extremely difficult to weld and has poor properties after welding, and it cannot be used in most of the power system components. Therefore, Nb-1Zr was selected for SP-100's cladding because it has the most favorable characteristics of all the refractory metals, but is not fully compatible with the fuel. To mitigate this, a sleeve liner must be in place between the cladding and the fuel.

Rhenium Alloys

The cladding-fuel intermediary sleeve may be made of rhenium because it interacts well with both cladding and fuel, may be easily fabricated, and has a low cross section for fast neutron absorption, and a high thermal-neutron absorption cross-section. Therefore, rhenium gives the added benefit of reducing the chance of criticality in a water submersion accident (Ring 2004).

Tantalum Alloys

The Ti main alloys are ASTAR-811C (Ta-8%W-1%Re-1%Hf-0.025%C) and T-111(Ta-8%W-2%Hf). ASTAR-811C has superior creep properties over T-111, yet it still retains the fabrication and welding characteristics of T-111. Both alloys were tested satisfactorily with lithium coolant. Tantalum alloys are significantly stronger than niobium, but they have higher fast neutron cross sections of absorption, are twice as heavy, and are more difficult to weld than niobium. In its ductile to brittle transition temperature tests, (Zinkle 2004) irradiated tantalum alloys showed embitterment in temperatures less than 1000K.

7.4.3.2 Material Tests for Space Nuclear Power

The following tests have to be performed to increase the database for material studies of space nuclear power systems:

- Irradiated tests on tungsten and tungsten-rhenium alloys in the 950-1300K range.
- Reactor fast fluence testing.
- Bonded rhenium barrier cladding long duration fission product build up tests.
- Impact and fracture toughness tests on tantalum irradiated samples at high temperature.
- PWC-11 irradiation and fracture toughness testing.

7.4.3.3 Coolants and Shielding

Lithium, NaK, Mercury materials

In nuclear power systems, Lithium may be used as primary cooling fluid, and as Lithium Hydride may be used as neutron shield since it has high hydrogen density (5.9×10^{22} hydrogen atoms/cm³), low mass density (0.775 g/cm³), a melting point of 960 K, and produces a minimal amount of secondary radiation. It has the property of expanding up to 25% when it reaches its melting temperature, solved by cooling the shield to prevent melting. It has a high thermal-neutron absorption cross-section. To avoid He generation in the shield, 99.99% ⁷Li enrichment is done. The liquid lithium coolant must be kept above 454K, its melting temperature. In

space applications (i.e. SNAP-10A operating at 816K), Li was preferred to NaK whose operating range is 800-980K. At higher power levels, in the MW range, higher temperatures are required; therefore, lithium or mercury are the best choices. Mercury has the disadvantages of having high density and having a high fast-neutrons absorption cross-section (Angelo and Buden 1985).

Lithium also has the advantages of low vapor pressure, low density, and high specific heat that lead to a lower operating pressure that is less stressful on the reactor components and allows for thinner piping walls. Lithium's high specific heat and low density lower the pumping requirements making the pump more reliable. Additionally, if oxygen levels are kept low in the coolant and in the cladding, there will be minimal corrosion due to the coolant.

7.5 Mars Reactors Derived from Space Propulsion Systems

The choice of technologies for the reactor system is fairly wide, although there is no off-the-shelf space reactor system presently available. The SNAP program flew one space reactor in 1965 and built six other working reactors during the program. All of these used UZrH fuel with a NaK coolant(Anderson 1983). The subsequent SP-100 program designed a reactor that used UN fuel with a Li coolant (Mondt 1994; El-Genk 1995; Mondt 1995); The Russian Rorsat reactors (about 30 flown in space) used UMo fuel (Angelo and Buden 1985). The Russian Topaz II reactor used UO₂ fuel and UZrH moderator.

Given the need for a small reactor, low shield mass, and high conversion efficiency, it was proposed to consider the following three options for the Kuiper Belt Objects (KBO) mission (Lipinski 1999):

(1) a UZrH-fueled, NaK-cooled reactor with a steam Rankine conversion system, or

(2) a UN-fueled gas-cooled reactor with a recuperated Brayton conversion system, or

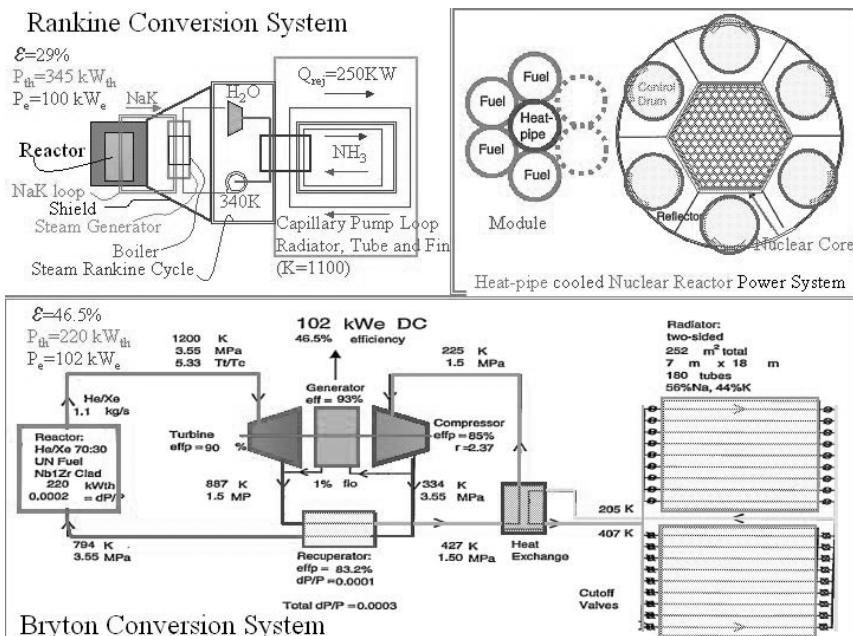
(3) a UN-fueled heat pipe-cooled reactor with a recuperated Brayton conversion system. Table 7.3 summarizes the key features for 100-kWe systems. All three systems assume using the same low-mass deployable thermal radiator to maximize the conversion efficiency (Lipinski 1999).

This kind of design is very important as it includes the most advanced technology available by 2010, and is capable of being used on Mars too, due to the similarity of the external operation conditions. The fact that it was designed for high efficiency in deep space operation is a plus for any Mars mission application.

Figure 7.8 shows the structure of the highest performing reactor structures available for space applications (Lipinski 1999). It is true that better efficiencies may be obtained in Earth tests, but the compliance with both space and Mars' conditions requires reliable continuous operation in very harsh environments that has its price included in efficiency. These configurations appeared as being more promising for deep space applications for electric propulsion to KBO and are reliable enough to be among the first to power medium size Mars' applications.

Table 7.3 Comparative look over the main nuclear reactor approaches

Component	UzrH / NaK-Cooled/Rankine	UN / gas-Cooled/Brayton	UN / Heat-pipe Cooled/Brayton
Electric power [kW]	100	100	100
Thermal power [kW]	345	220	333
Thermal efficiency [%]	29	46	30
Nuclear fuel	UZrH	UN	UN
Primary coolant	NaK	He/Xe	Na
Fuel clad material	Hastelloy	Nb1%Zr/Re	Nb1%Zr/Re
Vessel material	316 SS	Super Alloy	Mo
Conversion cycle	Rankine	Brayton	Brayton
Energy conversion working fluid	Water	He/Xe	He/Xe
Thermal radiator type	Heat-pipe/fin	Heat-pipe/fin	Heat-pipe/fin
Radiator working fluid	Ammonia	Ammonia	Ammonia
Reactor coolant exit temp [K]	723	1200	1200

**Fig. 7.8** Block diagrams and performances of the main space reactors

The heat pipe turns out to be a new reliable and safe heat removal solution making the contamination between the core and fluids even lower than the 2-3 heat exchanger solutions. For all the space systems, water freezing and phase

modifications represent an important problem to be solved, and it seems that heat pipes offer the best trade-off.

The shielding mass is among the main drawbacks of space nuclear power. The SNAP reactors that produced 600 kW of thermal power (kWt): SNAP-8ER and SNAP-8DR had a peak fuel temperature of only 800 K to extend the lifetime of the fuel. An enrichment of 93% was preferred. SNAP-10A space flight reactor and shield weighed about 268 and 98 kg respectively, summing to nearly 366 kg. To allow for a larger power and total burnup capability, was estimated (Lipinski 1999) that the reactor mass would be 500 kg and the shield would be 400 kg, heavy enough for being put on orbit, but not enough shielding to allow the presence living beings near by.

The main issues considered in the selection of a heat pipe reactor are given in Tables 7.4 and 7.5.

Table 7.4 Features of the Heat-Pipe Systems

No	Feature	Description
1	Safety	The HPS is designed to be subcritical for all credible launch accidents.
2	Reliability	The HPS has no single-point failures and is capable of delivering rated power, even if several modules and/or heat-pipes fail.
3	Long life	The design lifetime is in excess of 10 years. The low-power density in the HPS core and the modular design give the potential for long lifetime. At 100 kWt, fuel burnup limits are not reached for several decades.
4	Modularity	The HPS consists of independent fuel heat-pipe modules that can be tested individually.
5	Testing	The HPS system launch hardware can be tested at full power using electrical heaters in place of fuel rods. Unirradiated fuel rods are inserted before launch. Full-power nuclear tests might not be required. In addition, the Heatpipe Bimodal System, a system that uses the same approach as HPS, can be tested in the thermal propulsion mode using electrical heaters.
6	Versatility	The HPS can use a variety of fuel forms, structural materials, coolants, and conversion systems
7	Scalability	The HPS design scales well to beyond 1000 kW thermal power.
8	Simplicity	There are few system integration issues since there are no in-core shutdown rods, no hermetically sealed refractory metal vessel or flowing loops, no electromagnetic pumps, no coolant thaw systems, no gas separators, and no auxiliary coolant loop for decay heat removal.
9	Fabrication	Most of the fabricated parts are small modules with similar metals; there is no pressure vessel because there is no pumped coolant loops. There are no significant bonds between dissimilar metals, and thermal stresses are low.

Table 7.4 (continued)

No	Feature	Description
10	Near Term	The system needs no development of advanced materials or components. It can be developed ‘quickly and inexpensively with few nuclear tests.
11	Low Mass	The HPS system has a high fuel fraction in the core since it uses no in-core shutdown rods. The potential for in-space fueling (because it has no pressure vessel) allows a more compact form while still meeting launch safety requirements.
12	Storability	The HPS is designed such that the fuel can be stored and transported separately from the system until shortly before launch. This capability will reduce storage and transportation costs significantly
13	Milestones	There are several milestones early in the development of the HPS that will prove the viability of the concept. The most significant early milestone is the development and testing of an HPS module.
14	Bimodal	The HPS approach readily evolves into a bimodal system.
15	Dual use	Technology utilized has military, commercial, and civilian uses in both aerospace and terrestrial applications.
16	Mass	The HPS has a high fuel fraction in the core, thus reducing core, reflector, and shield mass for criticality-limited systems. The HI% has no pumped coolant loops and few system integration issues, thus further reducing mass.
17	Schedule	The attributes should allow for quick (<5-yr) development.
18	Cost	The attributes should allow for inexpensive (<\$100M) development. The production unit cost should be <\$20 M.

Table 7.5 Few heat pipe reactors parameters list

No.	Parameter\ Type	HPS7N	HPS30	HPS70	HPS100	HPS120
1	Power (kWt)	100	100	200	330	1000
2	Number of Modules	30	19	30	57	121
3	Fuel Material	UN	UO ₂	UO ₂	UO ₂	UO ₂
4	Fuel Enrichment	97%	97%	97%	93%	93%
5	Fuel Theoretical Density	96%	92%	92%	85%	85%
6	Clad Material	Nb/Zr	Mo3Nb	Mo3Nb	Mo3Nb	Mo3Nb
7	Pin Diameter (cm)	1.50	2.00	1.80	1.40	1.40
8	Core flat-to-flat (cm)	19.7	22.8	23.6	25.7	30.5
9	Core Active Height (cm)	32	36	36	36	42
10	Fuel Burnup per Year	0.05%	0.04%	0.04%	0.13%	0.30%
11	Reactor Mass (kg)	240	305	325	370	480

The HPS primary heat-pipes in uranium-nitride-fueled systems operate at a temperature of 1300 K and transfer heat to secondary heat-pipes operating at 1275 K. Then the heat is transferred from the secondary heat-pipes to the thermal-to-electric

power converters, and waste heat is rejected to space (Houts 1996). The 1275 K converter hot-side temperature is adequate for thermoelectric, AMTEC, or Brayton, although higher or lower temperatures could be used. Unicouple thermoelectric converters that are well suited for use with HPS have been designed (Raag 1995) with a hot-shoe temperature of 1275 K and reject waste heat at 775 K. A HPS using close-spaced thermionics could use lithium heat-pipes and uranium dioxide fuel to operate at a converter hot-side temperature of 1500 K with efficiencies >10%. A version of HPS uses 12 modules, each containing 4 rhenium-lined, Nb-1 Zr-clad uranium nitride fuel pins bonded to a central heat-pipe. The fuel pin's outer diameter is 2.54 cm, which allows existing electrical heaters to be used for testing (Izhvanov 1995).

If Hydrogen propellant flows through the interstitial space around the heat pipe shown in Fig. 7.8, with four/five-pin module and out through a nozzle gives thrust levels up to 400 N at exhaust velocities >8000 m/s. A vacuum gap isolates the heat-pipe from the Hydrogen flow, allowing electric power to be generated during propulsion mode. The vacuum gap also prevents heat-pipe dry-out at the hot end of the core (Poston 1996).

7.6 Safety Issues

Safety is an important design requirement applied since the very first systems (SNAP-3B), where many safety tests and analyses have been conducted. Formal safety analysis reports and independent safety reviews are required before each launch (Bennett 1991). With the enactment of the National Environmental Policy Act in 1969, an independent environmental assessment is also conducted.

For radioisotope power sources, the primary safety requirement is to contain or immobilize the radioisotope fuel so that there is no interaction with humans or the environment. Designs are assessed through a combination of safety tests (often involving explosions, projectiles, fires, and impacts) and analyses (Bennett 1991).

The primary safety requirement of space nuclear reactors is to become subcritical under postulated accident conditions (such as explosions, projectiles, fires, water immersion, and impacts). For the U.S. space nuclear reactor flown (SNAP-10A in 1965), extensive safety tests and analyses were conducted (Staub 1967; Bennett 1991).

All three U.S. missions, Transit 5BN-3 (21 April 1964); Nimbus-B1 (18 May 1968); and Apollo 13 (17 April 1970), were RTG-powered and the Russian's nuclear reactors (Cosmos 954 and Cosmos 1402) and three reentries of radioisotope systems (Cosmos 300, Cosmos 305, and Mars-96), involved in accidents met their safety design requirements and posed no threat to people. The accidents were not caused by the RTGs (Bennett 1991).

The nuclear reactor power systems (Houts 1996) have to remain subcritical during all credible launch accidents by design assured by:

- keeping the system radius small,
- keeping the reflector worth high, and
- strategically placing neutron absorbers in the core.

The negative reactivity worth of the control drums in the reflector, or the negative reactivity effect of losing the reflector and surrounding the reactor with wet sand or water, offsets the positive reactivity effect of core flooding or compaction, thus eliminating the need for in-core safety rods. For deep space or planetary surface missions, where reentry after reactor startup is impossible, fueling the reactor in space or using retractable boron wires to provide shutdown can obtain passive launch safety. This allows the removal of resonance absorbers from the core and reduces system mass and volume.

7.6.1 Control Strategies for Space Nuclear Reactors

Terrestrial reactors internal control systems seem inappropriate for space reactors due to the low temperature environment. The need for large external vessel and shield penetration, requires long life without maintenance and higher overhaul system mass. The favorable aspects of fast-spectrum reactors (Poston 1995) is given by the potential to control the system externally and eliminate the need for in-core safety elements.

The factors that determine if internal control is needed in a specific design are: criticality uncertainty margins, temperature defect reactivity, burnup reactivity loss, core diameter, and core void fraction where some of the factors are strongly dependent on thermal power level of the core; therefore, the number of internal control/safety elements is increasing with power. One effective power control system for fast space reactors is the Radial reflector usually made of Be or BeO with good neutronic performance due to high reflection and low absorption and thermal resilience. Be alloys also have low density and thin reflectors are required. Be is preferred over BeO due to better availability, fabrication experience, and performance under irradiation. A movable reflector drum that adjusts the neutron leakage and operates by sliding or pivoting the reflectors achieves another level of control. Sliders are lighter, but more flight experience has been accumulated with drum systems.

There are various cooling mechanisms such as liquid-metal –Li-cooled, gas-(HeXe)-cooled, heat- pipe-Na-cooled available. The calculation results (Poston D.I. 1995) show that BeO presents more reactivity sensitivity than Be resulting in a smaller radius, lighter control drum in spite of the higher density of BeO. For heat-pipe-cooled reactor these calculations show that the control drums are able to induce a variation in K_{eff} from 1.035 down to 0.935 when are turned by 180 deg. For the worst foreseen scenario when this reactor contains water in the internal voids (heat pipes) and wet sand in the external voids (equivalent with a crash in water and implanted in muddy-sandy bottom) k_{eff} at launch configuration being 0.925 with the drums rotated at 180 deg. It remains sub-critical after this worse case scenario with k_{eff} well under 0.985. as Fig. 7.9 shows. The Heat-Pipe Cooled Fast-Spectrum Reactor uses Be/B₄C rotating Control Drums made of 2 cm thick Be cylinders and boomerang/aerofoil shaped B₄C on one side in Nb1Zr clad being furthest from the core (at zero degree) having the maximum solid angle with the core. These safety measures are even enhanced by other passive measures making the probability of higher criticality after an accident even smaller.

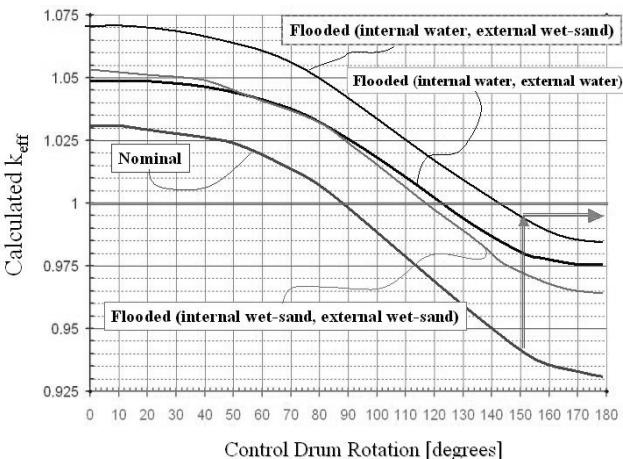


Fig. 7.9 Control Drum Worth HPCR

7.6.2 Reactor and Crew Shielding

Mars has a very light atmosphere compared to Earth (see Chap. 2); therefore, the radiation level on the surface is significantly higher than on Earth (International Commission on Radiological Protection 2003), compared with what High Altitude fighter plane pilots encounter, and not being too far from the conditions on Moon's surface (Cucinotta 2005).

One important requirement for an outpost on Mars is crew shielding from space radioactivity as well as from the own reactor radiation. As we showed in the first chapter related to Earth-Mars-Earth trips, that except for deep space trajectories (DST), most of the in-orbit plane ballistic flights pass near the sun, where the radiation exposure grows and adds a very long time to travel, in the range of ½-1 year.

Cosmic ray power deposition is pretty small due to low flux, at the level of $\mu\text{W/m}^2$, being mainly dominated by the electromagnetic radiation power at about 1.4kw/m^2 at 1 AU from the sun (Popa-Simil 2006).

For highly energetic particles such as protons and muons, the Earth's atmosphere ionization stopping power is about $2 \text{ MeV}/(\text{g/cm}^2)$ (CLBk. 2005). The C and O ions energy distribution at 1 AU from the Sun it is given in Fig.7.10.

At the border of the solar system the energy from galactic particles equals the total energy of the particles coming from the sun, at a value of about $10 \text{ [nW/m}^2 \text{ sr]}$. This is of the same order of magnitude as the total power obtained from light particles like electrons, neutrons, protons, muons and pions at 1AU giving a total dose rate equivalent of 0.5 mSv/h at 25 Km altitude on Earth (Giffin 1996) with a few orders of magnitude increase due to fluctuation of the flux. Cosmic particles spectral composition is showing the predominance of the light particles and the large fluctuation measured by spectrometers (Thayler 1997).

The radiation in space during travel is a matter of special concern because during the trip the exposure becomes higher than the admissible dose (Students for

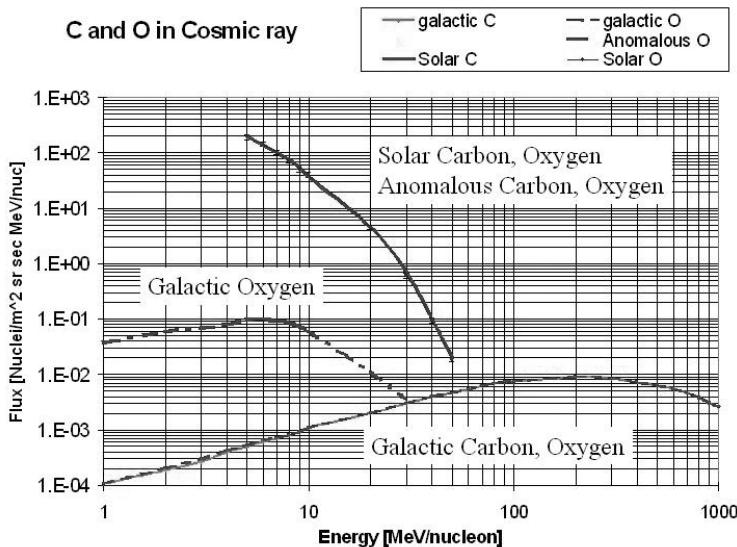


Fig. 7.10 Solar Energetic Particles and Anomalous Cosmic Rays at 1 AU

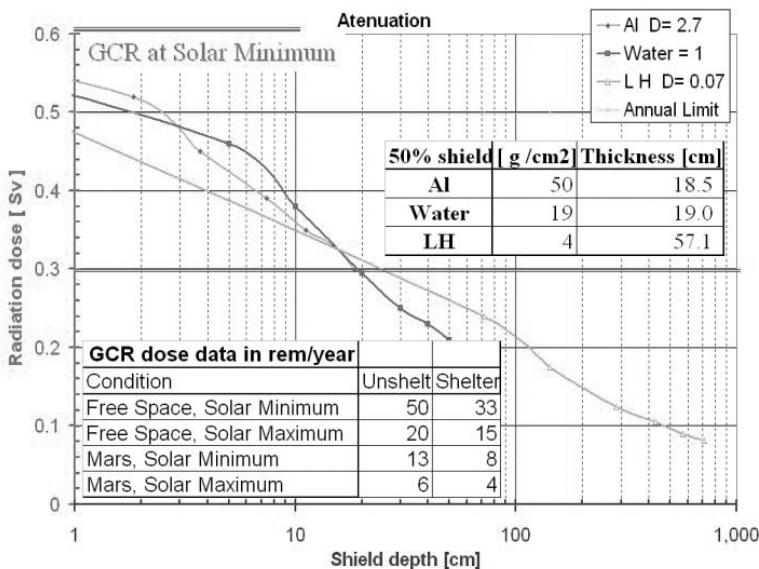


Fig. 7.11 Possible shuttle shielding

the Exploration and Development of Space 2004). Figure 7.11 shows the effectiveness of various shielding materials in space (Johnson Space Center 2005).

In the GCR dose data table for the sheltered condition 35 g/cm^2 of materials have been considered, while for unsheltered condition only 5 g/cm^2 . This gives a

dose estimate for astronauts going in sheltered conditions to Mars between 0.12 and 0.3 Sv/year.

A parametric study was conducted to afford mission designers first order estimates for the amount of lunar regolith required to protect the crew on a lunar surface mission from exposure to GCR, SEP, and neutron fields associated with a small fission reactor (Barghouty 2007).

Shielding is required for both reactor and crew, and a self-consistent approach was taken to estimate, at a given distance, the optimal (total) amount of regolith separating the crew from the reactor (Houts 1999). The additive nature of the solution in this treatment should allow for some flexibility in allocating material resource and/or construction effort between reactor and habitat.

Simple simulation-based dose-depth relations were used for all three-radiation sources in a one-dimensional optimization scheme. The objective was to estimate the optimal regolith mass between crew and reactor, as a function of their separation distance. The optimization scheme was based on the Pontryagin maximal principle. The scheme was applied to both solar maximum and minimum conditions. Depending on the mission's time profile, a savings of up to 30% in mass can be realized between optimized and un-optimized regolith mass estimates. However, it is argued that variation and uncertainty, mainly in lunar regolith attenuation properties and in the radiation quality factor, can easily defeat this and any other similar optimization scheme for 50 cSv/year irradiation dose, including externality cost (Cucinotta 2005).

As a safety precaution, the shielding has to be designed thicker than this optimum, in order to reduce the astronaut radiation load that is already significant due to the Earth-Mars travel time, especially in a very difficult to shield shuttle. The table in the upper right corner of Fig. 7.12 shows the dose for traveling in space, covering some unpredictable non-uniform particle distributions due to specific cosmic or solar activities.

Finally these absorbed doses have to be compliant with ICRP norms (International Commission on Radiological Protection 2003). This easily shows the potential overdose or incompatibility between actual space travel and astronauts radiation resistance providing a strong argument for faster deep space orbits. Finally, an astronaut will not be able to make more than two missions to Mars. Figure 7.12 also shows that a nuclear reactor has to be shielded under 2-3 m of ice where available or 1 m of mud and placed far from the outpost by about 200-1000 m, creating power transport and distribution issues also. Finally, with the proper shielding, a manned Mars mission can be accomplished safely.

7.6.3 The Radiolysis Effect

Radiolysis is an important effect in all materials exposed to radiation (Rajesh 2007). For water, 100 eV of radiation absorbed is enough to split about 5 water molecules (Huerta 2008). For a 250 kW reactor, it becomes possible to obtain 1 mol/h of oxygen and hydrogen, if collected and separated through the appropriate diffusion membranes. Secondary heat may also be used in hydrogen-oxygen production cycles as source of fuel for the O-H fuel cells. Another source of power comes from outer space radiation as cosmic and solar, but at smaller levels.

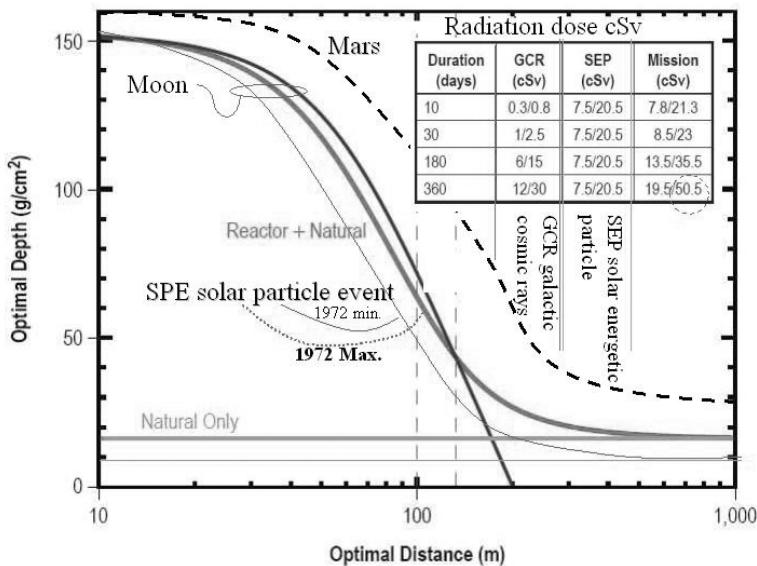


Fig. 7.12 The optimal distance and shielding material weight

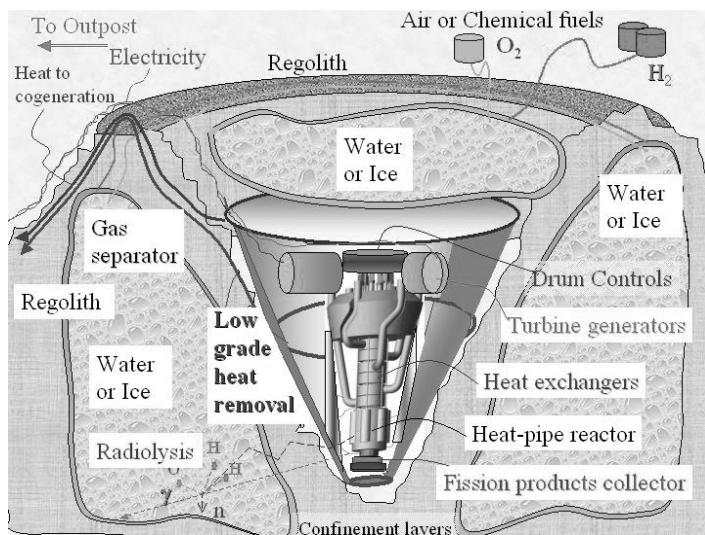


Fig. 7.13 Enhanced Reactor usage suggestion

If one would coat the polar iceberg with large double plastic foils act as H-O separation membranes, it might be possible to harvest 1 kg/Km²/day. Figure 7.13 shows a possible concept for an enhanced HOMER reactor (see Fig. 7.4) by using better Ti based materials, triple redundancy Brayton turbine-generators and a

enhance fuel design, allowing about 100 years of operation. Water/ice tanks using radiation for radiolytic production of Oxygen and Hydrogen make up the shielding, low-grade heat is used for cogeneration, and about 35% of the power is electricity. Finally, the shielding is compliant with the requirements for accepting the presence of the living beings near by the reactor's shielding.

References

- Anderson, R.V.: Space-Reactor Electric Systems: Subsystem Technology Assessment, ESG-DOE-13398, ch. IV. Rockwell International, Canoga Park (1983)
- Angelo Jr., J.A., Buden, D.: Space Nuclear Power, pp. 159–176. Orbit Book Co., Inc., Malabar (1985)
- Barghouty, A.F.: Optimization of Crew Shielding Requirement in Reactor-Powered Lunar Surface Missions, p. 25. Marshall Space Flight Center, Alabama (2007), <https://www2.sti.nasa.gov>
- Barghouty, A.F.: Optimization of Crew Shielding Requirement in Reactor-Powered Lunar Surface Missions, p. 25. Marshall Space Flight Center, Alabama (NASA/TP—2007-215133) (2007), <https://www2.sti.nasa.gov>
- Begg, L., Wetch, J.: Comparison of High Temperature Heat Rejection Concepts to System Related Requirements. In: 22nd Intersociety Energy Conversion Engineering Conference, Philadelphia, PA, vol. 1, pp. 227–234 (1987)
- Bennett, G.L.: A Look at the Soviet Space Nuclear Power Program. In: Proceedings of the 24th Intersociety Energy Conversion Engineering Conference, Crystal City, Virginia, August 1, pp. 7–11 (1989)
- Bennett, G.L.: The Safety Review and Approval Process for Space Nuclear Power Sources. Nuclear Safety 32(1), 1–18 (1991)
- Bennett, G.L.: Space Nuclear Power: Opening the Final Frontier in 4th International Energy Conversion Engineering Conference and Exhibit (IECEC), San Diego, California, June 26–29, 2006. AIAA 2006-4191. Metaspace Enterprises, Emmett (2006)
- Britt, F.G.: Discussion of Open Versus Closed Cycle Space Power Burst Energy Systems. In: Transactions of the Sixth Symposium on Space Nuclear Power Systems, Albuquerque, NM, January 8–12, 1989, pp. 357–362 (1989)
- Buckman, R.W.: Nuclear Space Power Systems Materials Requirements. In: STAIF, Albuquerque, NM, February 2004, p. 815 (2004)
- Carelli, M.: The Currie Point Radiator in Space Nuclear Power Systems 1988, ch. 45, pp. 367–375. Orbit Book Company, Malabar (1989)
- CLBk Cosmic Lawrence Berkeley group, Muons -The Cosmic Connection (2005), <http://cosmic.lbl.gov/>
- Cucinotta, F.A., Kim, M.-H.Y., Ren, L.: Managing Lunar and Mars Mission Radiation Risks Part I: Cancer Risks, Uncertainties, and Shielding Effectiveness, NASA—TP-2005-213164 NASA Johnson Space Center, Houston, TX (May 2005)
- Danneskiold, J.: Los Alamos leading fast-paced reactor research to power planned journey to Jupiter's icy moons in LANL News (December 2008), <http://www.news.lanl.gov>
- Department of Energy Office of Defense, Strategic Defense Initiative Multimegawatt Space Nuclear Power Program - Summary (1986)

- Dudinski, L.: Space Nuclear Power short course in STAIF. Albuquerque, Albuquerque, NM (2004)
- El-Genk, M.S.: In Ed. Univ. of New Mexico, Institute for Space Nuclear Power Studies, pp. 143–155. Albuquerque, NM and AIP Press, New York, CONF 940101 (1995)
- El-Genk, M.: Energy Conversion Technologies for Advance Radioisotope and Nuclear Reactor Power Systems for Future Planetary Exploration. In: 21st International Conference on Thermoelectronics 1 (2002)
- Elliott, J.O., Nakagawa, R., Spilker, T., Lipinski, R., Poston, D.: Design of a Shuttle-Tended Interplanetary Transfer Vehicle Using Nuclear Electric Propulsion, NASA Jet Propulsion Laboratory NEP (2008)
- Gerwin, R.A., Poston, D.I., Nebel, R.A.: Possibilities for Magnetic Control of Fission Plasma Propulsion. In: 31st AIAA Joint Propulsion Conference and Exhibit San Diego (1995)
- Giffin, N.: Cosmic Rays. TRIUMF Radiation Protection Training Course (1996)
- Golombek, M.: Radioisotopic Thermal Generators, probably with plutonium (2006),
<http://www.Space.com>
- Houts, M.G., Poston, D.I., Ranken, W.D.: HPS: A Space Fission Power System Suitable for Near-Term, Low-Cost Lunar and Planetary Bases. In: Space 1996, The Fifth International Conference and Exposition of Engineering, Construction and Operation in Space (1996)
- Houts, M.G., Poston, D.I., Trellue, H.R., Baca, J.A.: Planetary Surface Reactor Shielding Using Indigenous Materials. In: Proceedings of Space Technology and Applications International Forum (STAIF 1999), American Institute of Physics Conference Proceedings, vol. (458), p. 1750 (1999)
- Huerta, P.M., Rajesh, P., Mu, T., Pimblott, S.M., Laverne, J.A.: H atom yields in the radiolysis of water. Radiation Physics and Chemistry 77(10-12), 1203–1207 (2008)
- International Commission on Radiological Protection, Relative Biological Effectiveness (RBE), Quality Factor (Q), and Radiation Weighting Factor (wR) Publication No. 92, p. 80. Elsevier, London (2003)
- Izhevyanov, O.L.: The fuel pin test using electrical heaters (New Mexico Engineering Research Institute, Albuquerque, New Mexico) personal communication (March 1995)
- Jet Propulsion Laboratory, Mars Rovers Probing Water History at Two Sites (2004),
<http://www.jpl.nasa.gov/news/news.cfm?release=2004-253>
- Johnson Space Center, Shielding Effectiveness Against GCR at Solar Minimum (2005),
<http://srag-nt.jsc.nasa.gov/RadDocs/TM104782/techmemo.htm>
- Koenig, D.: Rotating Film Radiators for Space Applications in Society of Automotive Engineers SAE/P-85/164, 85–94 (1985)
- Lipinski, R.J., Lenard, R.X., Wright, S.A., Houts, M.G., Patton, B., Poston, D.: Fission-Based Electric Propulsion for Interstellar Precursor Missions, United States Department of Energy OSTI (1999)
- Liss, A.: Russia plans Mars nuclear station in BBC. Moscow, 18 August (2003)
- Malik, T.: Plucky NASA Rovers Complete Fifth Year on Mars in Space 2009/01/03 (2009),
<http://www.space.com>
- Mars Science Program Synthesis Group, Mars Exploration Strategy 2009-2020, NASA Jet Propulsion Laboratory, Ed. Dan McCleese (2003)
- Marshall, M., Robinson, E., Park, T., Brownlow, L., Elliot, J.: Nuclear Power Options for Mars Polar Robotic Outpost. In: STAIF 2004 (2004)

- Mattick, A.T., Hertzberg, A.: The Liquid Droplet Radiator: An Ultra lightweight Heat Rejection System for Efficient Energy Conversion in Space. *Acta Astronautica* 9(3), 165–172 (1982)
- Mattick, A.T.: Experimental Test of Liquid Droplet Radiator Performance. *Transactions of the Third Symposium on Space Nuclear Power Systems*, Albuquerque, NM (TM-1), 1–2 (1986)
- Mattick, A.T., Hertzberg, A.: Advanced Radiator Systems for Space Power. In: *38th Congress of the International Astronautical Federation*, Brighton, UK IAF-87-230 (1987)
- McGinnis, S.J.: Nuclear Power Systems For Manned Mission To Mars, *Astronautical Engineering* Monterey, California, United States Naval Postgraduate School, p. 105 (2004)
- Mondt, J.F., Truscello, V.C., Marriott, A.T.: SP-100 Power Program. In: *Eleventh Symposium on Space Nuclear Power and Propulsion*, Albuquerque, NM (1995)
- Mondt, J.F., Truscello, V.C., Marriott, A.T.: SP-100 Power Program. In: *11th Symposium on Space Nuclear Power and Propulsion*, Albuquerque, CONF-940101, pp. 143–155 (January 1994)
- NASA, SP-100 Power Source (2004),
<http://spacelink.nasa.gov/NASA.Projects/Human.Exploration.and.Development.of.Space/Human.Space.Flight/Shuttle/Shuttle.Missions/Flight.031.STS-34/Galileos.Power.Supply/SP-100.Power.Source>
- NASA, The Vision for Space Exploration (2004), http://www.nasa.gov/pdf/55583main_vision_space_exploration2.pdf
- NASA, Mars exploration program (2008),
<http://mars.jpl.nasa.gov/missions/>
- National Research Council, E.E.B., Committee on Advanced Space Based High Power Technologies. In: *Advanced Power Sources for Space Missions* Washington, pp. 33–34. National Academy Press (1989)
- Ponomarev-Stepnoi, N.N., Talyzin, V.M., Usov, V.A.: Russian space nuclear power and nuclear thermal propulsion systems. *Nuclear News* 43(13), 33–46 (2000)
- Popa-Simil, L.: Long life single load reactor fuel. In: *Proceedings ICAPP 2006*(1) pp. 140–148 (2006)
- Popa-Simil, L.: The Nuclear Power Demand and Limitations for Deep Space Exploration in Space Nuclear Conference Boston, MA, ANS (2007)
- Poston, D.I.: A Computational Model for an Open Cycle Gas Core Nuclear Rocket in Ph. D. Dissertation. Dept. of Nuclear Engineering, Univ. of Michigan, Ann Arbor (June 1994)
- Poston, D.I., Amiri, B., Chorback, C., Ellis, T., Hehr, B., Yocom, D.: Control Element Options for Compact Fast-Spectrum Space Reactors in Control and Safety Strategies for Space Nuclear Reactors Conf., vol. 1, pp. 537–541 (1995)
- Poston, D.I., Houts, M.G.: Nuclear and Thermal Analysis of the Heatpipe Power and Bimodal Systems. In: *Proceedings of Space Technology & Applications International Forum (STAIF 1996)* DOE CONF-960109 (US Department of Energy report), pp. 1083–1093 (1996)
- Raag, V.: Unicouple thermoelectric converters at Thennotrex, Waltham, Massachusetts, personal communication (August 1995)
- Rajesh, P., LaVerne, J.A., Pimblott, S.M.: High dose radiolysis of aqueous solutions of chloromethanes: Importance in the storage of radioactive organic wastes. *J. Nucl. Mater.* 361, 10–17 (2007)

- Ring, P.J., Sayre, E.D.: Material Requirements, Selection, and Development for the Proposed JIMO Space Power System. In: STAIF, Albuquerque, NM, February 2004, pp. 808–812 (2004)
- Staub, D.W.: SNAP 10A Summary Report in Atomics International Report NAA-SR-12073 (1967)
- Students for the Exploration and Development of Space, Radiation and the Human Mars Mission (2004),
<http://66.102.7.104/custom?q=cache:h2znA1SzvrQJ:www.seds.org/pub/info/mars/RadHuman.mcw+rad+human&hl=en&ie=UTF8>
- Thayler, M.R., Binns, W.R., Dowkontt, P.F., Hink, P.L., Klarmann, J., Lijowski, M.L., Olevitch, M.A., Cohen, C.M.S., Cook, W.R., Cummings, A.C., Kecman, B., Leske, R.A., Mewaldt, R.A., Stone, E.C., Dougherty, B.L., Radocinski, R.G., Wiedenbeck, M.E., Christian, E.C., Von Rosenvinge, T.T.: The Cosmic Ray Isotope Spectrometer (CRIS) Experiment Flight Performance. In: American Physical Society, APS/AAPT Joint April Meeting, Columbus, Ohio, abstract #C8.05, April 18–21 (1998)
- Zinkle, S.J., Wiffen, F.W.: Radiation Effects in Refractory Alloys. In: STAIF 2004, Albuquerque, NM, pp, 733–739 (February 2004)

Chapter 8

Distributed Power Sources for Mars

George H. Miley, Xiaoling Yang, and Eric Rice

University of Illinois, Urbana and Orbital Technologies Corp., Madison, USA

8.1 Introduction

One of the fundamental needs for Mars colonization is an abundant source of energy (Shaban and Miley 2003). The total energy system will probably use a mixture of sources based on solar energy, fuel cells and nuclear energy. Here we concentrate on the possibility of developing a distributed system employing advanced nuclear energy, specifically a mixture of small fusion devices and low energy nuclear reaction (LENR) cells (Miley 1997; Miley et al. 2002a). The fusion devices would provide small central units in the 500 kWe - 1 MWe level. The LENR units would serve as small portable sources ranging from Watts up to kiloWatts. All units would be designed to minimize radiation emission and radioactive waste.

It should be stressed that both the fusion device and LENR cell described here are in a relatively low Technical Readiness Level (TRL), much more effort is needed to fully qualify them. The purpose of this presentation is not to preset a detailed description of the physics behind the concepts, rather to provide a vision of what could be accomplished if such technologies prove out. Indeed, even if the exact embodiments discussed here are blocked, it seems quite probable that at least one of the alternate approaches being researched world-wide in these fields will succeed and, in doing so, provide an energy option for Mars colonization similar to that presented here.

To illustrate a fusion concept that could be developed with the low power levels desired we have selected the Inertial Electrostatic Confinement (IEC) approach (Miley 1997, 1994). Since plasma losses in the IEC scale with transport in velocity space, small units appear feasible. This is in sharp contrast to conventional magnetic confinement devices such as the Tokomak where plasma transport losses scale as the surface to volume ratio, forcing very large units in the thousands of MWe class. Likewise, inertial confinement fusion (ICF) requires large multi-mega joule laser or heavy ion beam drivers, again forcing very large power units. Another key requirement for ease of use is to minimize radiation involvement and

neutron production, necessitating use of D-He³ or p-B¹¹ fuels. P-B¹¹ is ideal since the fuel resources are plentiful and the reaction results in production of energetic alpha (helium) particles with only minor gamma radiation (Miley 1976). To burn such fuels requires the ability to operate with a non-Maxwellian plasma energy distribution in order to reach a reasonable fusion energy gain without excessive input power and to maximize the ion to electron temperature ratio so that Bremsstrahlung losses are not excessive.

There are other “alternate” confinement concepts such as the Field Reversed Configuration (FRC), Spheromak, Z-pinch, Tandem Mirror; Dense Plasma Focus (DPF) etc are also candidate concepts. However, based on what is presently known, the IEC meets all of the key requirements: namely, simple construction, high power to weight, good scaling over a range of power levels, and ability to maintain a stable non-Maxwellian plasma energy distribution. Thus it is used here to illustrate the fusion option.

While the IEC appears ideally suited for small units from a theoretical base, the experimental database for such devices is limited and leaves open a number of scale-up issues. One complication is that small IEC units seem to necessitate the use of a grid structure. This introduces issues of grid lifetime, but some prior studies suggest that proper design and selection of the operational mode can provide reasonable operational life times (Chacón et al. 1997). In addition, in order to minimize radioactivity while operating under the less demanding confinement requirements, D-³He fuel would be used in first generation Mars plants (p-B¹¹ has many advantages, and would be employed in second generation plants if the confinement physics for the IEC works out to be as favorable as anticipated). The use of D-³He brings in a need for developing a ³He supply infrastructure (Sved et al. 1995). While this could be initiated through Mars mining, other methods for ³He breeding could be employed during initial stages of operation (Miley 1998).

The LENR cells for small portable power units would supplement the larger IEC power units. They appear to have minimum radiation involvement, but the main issues in their development concerns R & D related to energy conversion, waste heat management, auxiliary start-up power, and life time (Miley et al. 2002a). Again, the database for such cells is very limited so that a number of issues remain open in relation to scale-up for practical operations including a selection of electrode materials and electrolyte.

With respect to materials for both of these power units, it would be desirable to develop manufacturing techniques that involve indigenous materials available on Mars. If that is possible, the energy network could be manufactured during the colonization process using a minimum of initial power sources to start the process. That would greatly reduce the cost for transport of units to Mars could add great flexibility for future expansion of the colony and its needs expand. The IEC can be operated in a mode to provide a high power exhaust plasma stream (Jet plasma) which appears particularly well suited for certain operations such as oxygen and super Greenhouse gas production (Fig. 8.1). Such IEC units are termed a “plasma torch”. In this presentation these issues will be discussed along with several conceptual designs for both types of units.

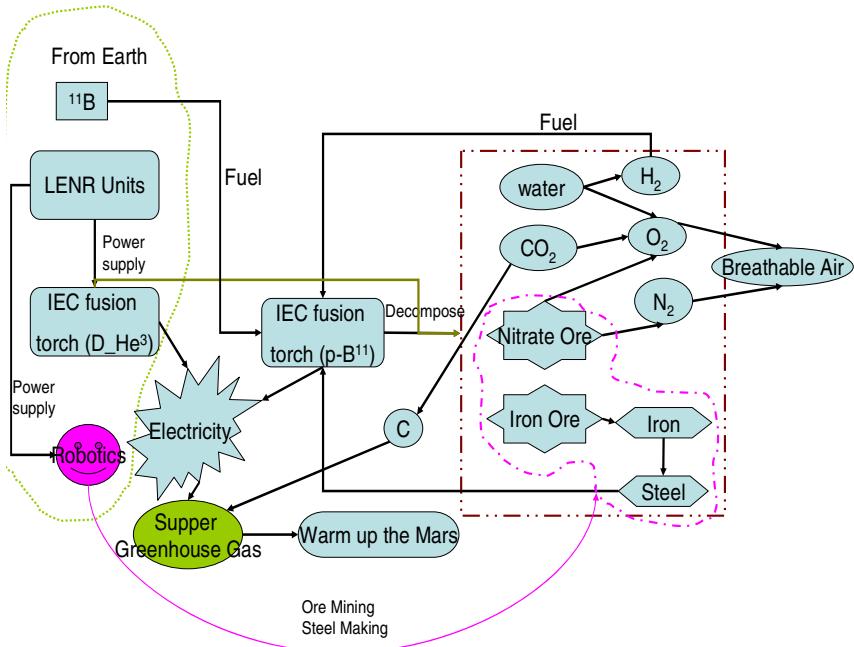


Fig. 8.1 Create a habitable Mars by Nuclear Power Units

8.2 IEC Power Units

8.2.1 Introduction

Fusion energy offers an attractive source of power for space applications. Inertial Electrostatic Confinement (IEC) fusion in particular appears to be one of the most promising approaches. A NASA study has identified an IEC-powered propulsion unit as one of the highest specific power concepts among various fusion design options (Williams and Borowski 1997). Also, the IEC is one of the few fusion concepts that permit small size and modest power level units. In view of its attractiveness, this study considers an IEC for use as a space power unit to supply 1-MWe for “on-board” uses or for local uses at a landing base. Once operation, this base unit would supply power needed for oxygen production and ultimately for manufacturing additional IEC fusion plants.

Inertial electrostatic confinement fusion is a method of electrostatically confining fusion plasma, which was first proposed by Salisbury (1947) and Farnsworth (1966). Early studies were carried out by Hirsch (Hirsch 1967) in the 1960s, but little was done after that until recent experiments by R. W. Bussard at EMC² and G. Miley at the University of Illinois (Miley et al. 1994a; Bussard et al. 1994). As the result of the Illinois research, a version of the IEC was developed by Daimler-Benz Aerospace as a small, portable neutron source for neutron activation analysis (NAA) in industrial process control (Sved 1992). The Daimler-Benz unit employs

D-D reactions to produce 2.5-MeV neutrons at about 10^8 n/sec. steady state. This in fact represents the first commercial use of confined fusing plasma. Due to strategic decisions, however, these units were only used in-house and not sold commercially as initially planned. While an attractive low level neutron source, this device is far from energy breakeven, having a Q-value (fusion energy out/energy in) of $\sim 10^{-5}$. Thus many physics and technology issues require research and development before it is certain that a power plant such as envisioned here or by the EPRI and NASA studies cited above can be achieved. Fortunately, such research could be done rapidly using a series of relatively small-scale experiments. Since confinement losses occur through velocity space processes, IEC scale-up requires an increased input power but does not necessitate a significant increase in physical size. In fact, the sizing will ultimately be determined by heat transfer limits, much as discussed here in relation to the space power unit.

8.2.2 IEC Concept

As shown in Fig. 8.2, the gridded IEC employs a spherical geometry with a concentric spherical grid located inside a vacuum vessel. The inner grid, the cathode, is placed at a large negative potential (80-100 kV) with respect to the vessel. When backfilled with a low pressure gas, the high electric field ionizes the gas and accelerates the positively charged ions towards the center of the device. As these ions converge in the center of the sphere, a dense core region forms where a high fusion rate occurs. Because of space-charge build-up of the ions and electrons in the core region, virtual anodes and cathodes form in a spherical potential well structure (Miley et al. 1998). This structure greatly enhances the ion confinement and increases the fusing ion density in the center spot.

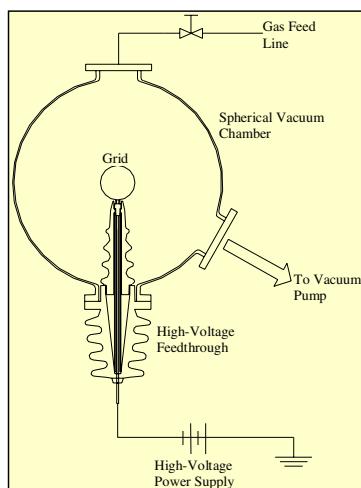


Fig. 8.2 Schematic of Gridded IEC Experiment

The non-Maxwellian beam-beam type reactions, plus a lack of cyclotron radiation due to the elimination of B-fields makes the IEC attractive for burning advanced fusion fuels, like D-³He and p-¹¹B. In view of its advantages (reduced neutrons and radioactivity) and its relatively large reactivity, D-³He is employed as the IEC fuel for the initial plant to be shipped to Mars. As discussed latter, subsequent larger plants built on Mars will be designed for p-B11 operation. That avoids fuel (He³) transportation problems for the main power producing units which consume the most fuel, and also provide aneutronic fusion to avoid possible neutron activation of materials in the proposed materials processing operations using these power plants.

This is the first conceptual design for an IEC for use as a small power unit and as such, it must be viewed as a preliminary design. Prior studies have focused on high power IEC uses for space propulsion and for hydrogen production (Miley et al. 1993, 1997a). A difficulty in projecting the use of the IEC as a power unit is that the experimental database, although growing, is as yet inadequate. Thus, the extrapolation to a power unit design contains many uncertainties. Still, the conceptual design study presented here is most useful in terms of identifying key issues that need further study and to illustrate the potential advantages of development of small fusion power units for space colonization.

The design of an IEC reactor includes four main components: the IEC itself, a direct energy converter (DEC), the step-down electronics, and an energy storage/pulse-forming unit. In addition, cooling systems and waste heat radiation play a major role. A block diagram of the system is shown in Fig. 8.3.

In the present design study, a 1-MWe power plant concept featuring a pulsed-power design is considered. The results suggest small-scale IEC reactor designs are conceivable, with very attractive operational characteristics. If higher power

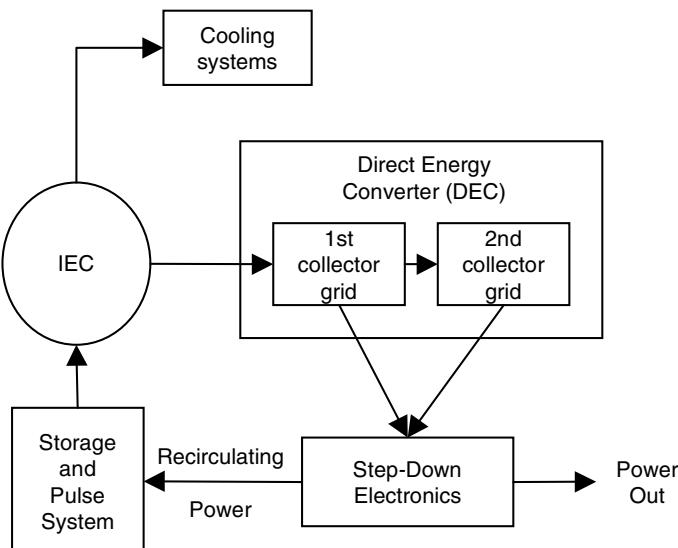


Fig. 8.3 Block Diagram of the Power Flow in an IEC Power Plant

levels are desired, a higher input power appears feasible, or, alternately, multiple units can be employed.

8.2.3 IEC Experiments

University of Illinois experiments in small 30-cm and 60-cm diameter IEC devices have been quite successful, achieving a measurement of the potential well during low current operation (≤ 60 mA) and also achieving strong ($\leq 10^8/\text{sec}$) 2.45-MeV steady state neutron emission when deuterium fill gas is used (Farnsworth 1966; Miley et al. 1998). The level of operation, however, is far below that needed for net power output. To scale up to energy breakeven, higher ion currents of order of kAs are required. As a first step, experiments have been initiated that employ a pulsed power supply (Modified Marx bank) to obtain high peak currents (~ 5 A), with equivalent time-averaged neutron rates of 10^9 D-D neutrons/sec (Jurczyk et al. 1997). The data base from these studies, combined with theoretical studies, (Chacón et al. 1997) is used as the basis for the IEC power plant design study discussed next.

8.2.3.1 Scaling and Cooling Issues

The power flow analysis of the IEC reactor components requires knowledge of how the fusion rate scales with injected ion current. The steady state and pulsed studies cited above, combined with theoretical studies, allow an estimate of the scaling law at high injected ion currents. A “minimum” scaling of the fusion reaction rate with I^2 (I = ion current) is predicted for beam-beam reactions in a deep double well (Miley et al. 1997b). There are, however, several theories and simulation results that support even steeper scaling rates with current due to greater compression associated with standing wave formation (Bussard 1991). However, in the present study, to be conservative, an I^2 scaling is employed. Pulsed operation is assumed to be combined with state-of-the-art pulsed power technology to achieve high peak currents, thus taking advantage of the strong scaling with current.

In this pulsed design, the pulse width is held in the ms range, so that the potential well structure is in quasi-equilibrium over most of the pulse period. Because the fusion energy output scales non-linearly with input current, a very large energy gain is obtained during the pulse. A peak pulsed current of 1 kA would produce a peak fusion rate of 10^{19} fusions per second or a peak output power of 100 MWe. With the 1/100 duty cycle, an output of 1-MWe is produced with 4.8-MW recirculating power. The pulsed power technology to produce current pulses of 1 kA or more falls well within the present state-of-the-art, such as power units for light ion beam fusion experiments (Van Devender and Cook 1986). Thus a pulsed power plant design of this type appears feasible and offers improved energy gain and a compact design.

The scaling law determines ion current requirements; however, the physical size of the reactor is determined by the limitations associated with grid heating. We assume this initial IEC for use on Mars would use D-He³ fuel since it provides a high fusion rate at modest temperature while minimizing neutrons and hence radioactivity involvement (Miley 1976). Thus, this appears to be a good way to get

stated with a Mars plant. As discussed later, subsequent plants would be converted to cleaner burning p-B¹¹ which is aneutronic and also avoids the need for developing an infrastructure to deliver He³ from lunar mining. In the design of the 1MWe D-He³ reactor, the small inner grid is of main concern for cooling because it receives a large amount of power which must be dissipated as heat, either through radiation or by forced convection. The grid is made of tungsten, which has a melting point of 3680 K. The inner grid consists of tungsten tubing through which coolant can be pumped to remove excess heat. The heat radiated from the two IEC grids, as well as the heat from the DEC grids, is eventually absorbed by the chamber wall. Thus, this wall must also be designed with cooling systems to dissipate the heat, which eventually goes to a space heat radiator. For the 1-MWe plant, 4.5 MW is received by a wall surface area of 310 m².

The strategy for development of IEC fusion power on Mars involves shipping the components needed for the 1 MWe plant described above and use it to develop oxygen for human habits and power the initial manufacturing capability to ultimately manufacture a series of larger plants needed for other Mars operations such as materials processing and terraforming.

Gun Injected Design for Larger Power p-B¹¹ Units

The key to developing subsequent high power units following the initial 1 MWe unit is to use external ion guns to form and inject ions into the spherical IEC chamber. This eliminates the need for a grid and allows differential pumping between the guns and the main chamber so that the high vacuum needed in the chamber itself can be achieved. University of Illinois researchers have studied such a system, both theoretically and experimentally. Theoretical studies confirm that, if correctly designed, such IEC plasma can exist stably and has sufficient confinement time for aneutronic fusion. This assumes, however, very precise control is maintained over the energy and angular momentum of injected ions and a balanced supply of electrons is provided. A RF ion injector capable of such operation and has also developed corresponding electron feed techniques. A graded index field is used to increase the ionization efficiency. A key component is the magnetic focusing lens at the extraction port. This allows very efficient differential pumping between the high pressure in chamber and the low pressure IEC chamber.

The RF ion injector is shown attached to an IEC chamber in Fig. 8.4. These studies did not yet include differential pumping so that number of recirculating passes by an ion was low, roughly 2. The injected ion current, I , was about 50 mA. Still, based on measurements of neutrons using deuterium fuel, the Q (fusion energy gain/energy in) was respectable for such a small device, order of 10^6 . Based on these very encouraging results, we can predict the behavior of unit with larger numbers of injector guns and differential pumping.

The Q scaling for such a device goes as $\beta I^2/\alpha^2$ where α is the radius of the dense core, I is the ion current and β is the number of ion recirculations. With a 12 gun design, we expect to increase β to ~200, I to 1800mA, and reduce α by 10 compared to existing unit of Fig. 8.2. This predicts an increase in Q (compared to prior gun experiment) of $\sim 10^5$, giving approximately $Q \sim 12.6$ ($Q=1$ is “breakeven”). As already discussed, this reactor would run with an applied voltage slightly

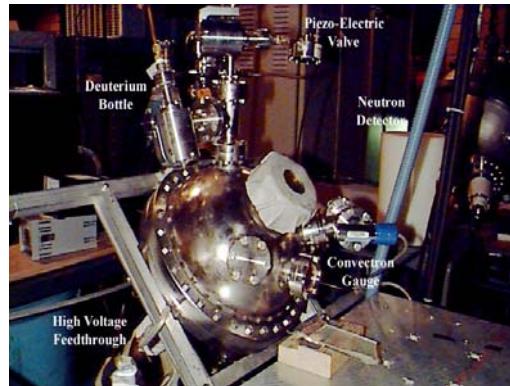


Fig. 8.4 Photo of IEC experimental device.

above 150 kV for efficient burning of p-¹¹B. Fortunately, as pointed out earlier, such high voltage operation has already been achieved in several IEC experiments. They did not, however, have the gun injection and differential pumping reused to obtain the required ion recirculation currents needed for burning p-B¹¹. Still, if the voltage operation is combined with proper injection/vacuum technology, a p-B¹¹ reactor could be achieved according to the physics outline here. Thus all the technology seems to be in place to develop a massively-injected p-B¹¹ IEC reactor which is manufactured and run on Mars. Of course, much more R&D must be done on Earth to prove this concept so a proven design is ready for Mars manufacture.

8.2.3.2 Projected Parameters

A summary of the 1-MWe pilot power plant design is given in Table 8.1. The size shown corresponds to the radius of the spherical vacuum chamber wall, exclusive of cooling systems, for each design. Weights are shown for the IEC power alone and for the entire power system, cf. Fig. 8.2. Most of the total weight involves the recirculation power equipment, the cooling system, and the shadow shield. Despite its small size and relatively low current, a modest fusion power gain of about 5 is projected, based on the scaling consideration noted earlier. Still, this leads to a larger recirculating power than that required for a larger high gain unit.

Table 8.1 Characteristics of IEC Power Plant Analysis for an Attractive 1-MWe Unit Utilizing Pulsed Power Operation

Gain	5
Recirculating power (MW)	0.48
Size/radius (m)	2.1
IEC weight (tonne)	0.5
Total weight (tonne)	4.2

8.2.3.3 Direct Energy Converter (DEC) and Step-Down Electrodes

The bulk of the energy output from D-³He fusion reactions is in the form of high-energy charged particles, a 14.7-MeV proton and a 3.54-MeV alpha particle. In addition "side" reactions produce some neutrons, but the energy fraction carried by neutrons can be held to <<10%, so their effect is minimized. To collect the energy of the charged particles and convert it to useful electrical power, a direct energy converter (DEC) and step-down electronics to condition the output electrical output are required. The DEC and conditioning circuit will be quite similar to that described in detail in a companion paper on IEC space propulsion (Burton et al. 2003).

8.2.3.4 Mass Power Density Projections

When compared to other power reactor designs including fission, fusion, and fossil fuel plants, the IEC fusion power plant offers a quite high specific power. Figure 8.5 compares mass power densities for many reactor designs. The IEC characteristics are shown alongside the mass power densities of other reactor designs that were compiled by Krakowski (Krakowski 1991). The mass power densities shown correspond to output power levels of 1000-1200 MWe. Two points for the IEC have been added based on earlier studies of 25 MWe and 1,200 MWe units (IEC-25 and IEC-1200, respectively, in Fig. 8.5). The difference between the two demonstrates

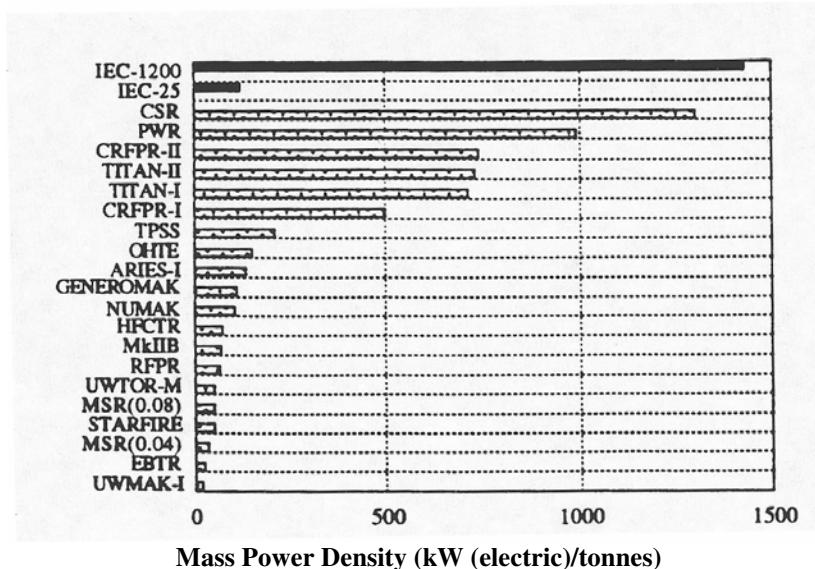


Fig. 8.5 Comparison of Mass Power Density Projected for the IEC to Recent Fusion Reactor Designs. Nomenclature and references for the other reactors shown are given by Burton et al. (2003). IEC-1200 and IEC-25: 1200 MWe and 25 MWe D-³He IEC reactors, respectively; PWR: pressurized water reactor; CRFPR through OHTE are advanced reversed field-type fusion reactors; ARIES-I through UWMK-I are various "conventional" tokamak-type devices.

the strong dependence of mass power density on the power level. Small units like the present 1-MWe plant have a lower mass power density than large units, but the relative ranking is expected to remain relative to the various types of devices if they were also scaled down in power. Thus, the IEC offers a very attractive approach compared to other concepts.

For reference, note that the high power version of the IEC has a higher mass power density comparable to a light water reactor fission reactor (PWR in Fig. 8.5). The mass power density is based on the fusion “power core” components, which, for the IEC, are taken as the IEC unit plus the DEC. Due to the drastic differences in design features between the IEC and the other plants, the definitions of components to be included in the “core” are somewhat ill-defined. Still, the trends shown should be valid, so the present 1-MWe IEC is expected to be very competitive compared to small fission reactors.

8.3 The IEC Fusion-Plasma Torch for Materials Processing

8.3.1 Introduction

Nuclear sources, both fission and fusion, have ability to generate extra heat, which could have multi-usage, such as high temperature electrolysis or thermo-chemical decomposition (Schultz et al 2003; Brown et al. 2002). Numerous studies have

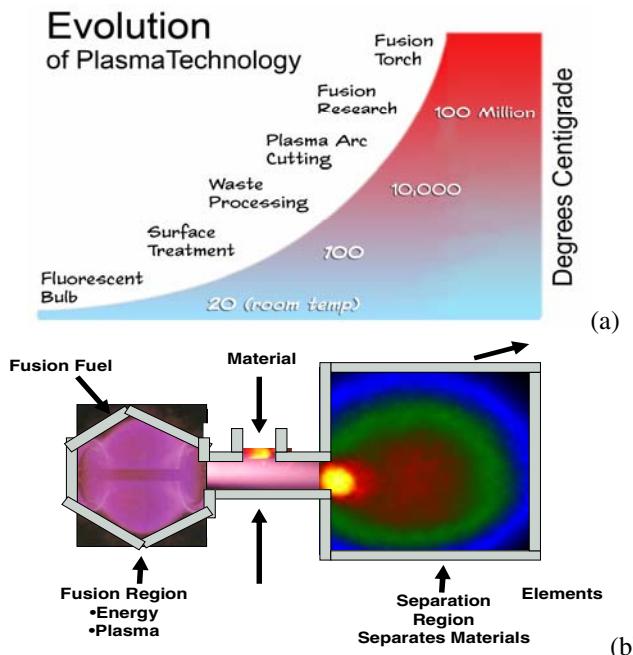


Fig. 8.6 (a) The fusion torch and research devices such as the electrically drive IEC plasma are orders of magnitude hotter than arc type plasmas currently used for waste processing. (b) The scheme of a conceptual fusion torch.

envisioned a fusion reactor for the heat source. Among them, IEC fusion device is inherently small and simple to construct. Besides, it can achieve unbeatable high temperature compared with other fusion devices (see Fig. 8.6(a)). Figure 8.6(b) shows the scheme of a conceptual fusion torch. Materials are fed into the chamber where the high temperature is maintained by fusion torch. The ionized material is transferred to the separation chamber for further separation.

Recently, an IEC fusion torch concept for material recycling has been proposed (Gough and Eastlund 1969), upon normal fusion torch concept for waster recycling (Gough and Miley 2008). On top of that, we envision a broad application of this torch on Mars' colonization, e.g. decomposition of CO₂ or Ores.

8.3.2 Concepts and the Previous Proposed Applications

The original fusion torch concept was vague about the confinement type and fuel employed (Gough and Eastlund 1969). More recently the authors and collaborators proposed an IEC-fusion-torch system using the hydrogen-boron fusion fuel cycle (p-B¹¹). When burned, unlike D-T fusion which releases 80% of its energy in neutrons, p-B¹¹ yields 3 energetic alpha (helium ion) particles. This fuel is then ideal since neutron activation and tritium contamination of the materials are avoided. However, burning it requires much hotter plasmas and better confinement that is not possible with conventional confinement concepts such as the Tokomak (Miley 1976). Fortunately, research now shows that p-B¹¹ fusion is possible in the IEC configuration (Gough and Miley 2008). The IEC is well-suited for this demanding task due to its characteristic non-Maxwellian plasma and potential trap physics which make sustainment of recirculating high energy ion beams. In addition, as already noted, the hot plasma can be conveniently coupled out via a jet-like electrostatic “divertor” into the processing region.

Indeed, the required ion temperature conditions for burning p-B¹¹ have already been achieved in experimental IECs at several labs (Miley 1976, 1999). However, the confinement time and the stability of its non-Maxwellian plasma must be maintained at higher power levels. This appears possible for a carefully controlled ion injected type IEC and this approach is currently under study at several IEC labs, including the University of Illinois. There are now a half dozen small but active p-B¹¹ research programs in the United States.

Actually, on Earth, municipal solid wastes are being processed by using plasmas arc technology. Plasma recycling with energy recovery is a world wide activity involving at least nine companies. An off-shot of an IEC torch – IEC plasma jet could generate very hot process plasma (~10³ times the arc plasma temperature) in a configuration where the plasma flows out through an exit nozzle and interacts with the ore. The ultra-high plasma temperature combined with the large volume of flowing plasma is a key to achieving very efficient ore destruction and energy recovery. This jet with 3 orders of magnitude above arc type plasmas which run at a few eV could translate the enormous temperature advantage into the ability to completely vaporize and dissociate ores on Mars, increasing the production of useful syngas products for electrical energy generation or heating. These destructed gas sources could also be utilized by other energy generation units, such as fuel cells (Luo et al. 2006).

In addition to materials processing, the IEC torch can also be used for water dissociation to produce hydrogen (Gough and Miley 2008). In this process, water is completely ionized in the high temperature flowing plasma. The ions are then separated out and ultimately recombined into hydrogen and oxygen using a modified electrostatic direct collection section inserted into the diverted plasma column (Gough and Miley 2008; Miley 1976). In the process, the excess energy carried by the ions is electromagnetically converted into electrical output, greatly improving the overall plant efficiency. The other unburned fusion fuel species and reaction products (“alphas particles”, i.e. helium) are also divided into separate streams and their excess kinetic energy is also directly converted to electricity by the electrostatic converter. This approach potentially offers several key advantages over high temperature electrolysis. Very pure hydrogen and oxygen are obtained with high conversion efficiency. Also, excess (waste) energy involved in the ionization process is directly converted to electricity at a high efficiency, allowing a hybrid hydrogen-electrical plant.

On Earth, the use of p-B¹¹ fusion has many practical advantages for industrial processes. In addition to minimal involvement of radioactive substances, the fuels required are readily available. Hydrogen for the reaction would be extracted as a small side stream from the main hydrogen production in the plant. The Boron required is abundant and ubiquitous on Earth. For fusion, the isotope B¹¹ is needed, and fortunately this isotope represents 80% of natural boron. However, by now, the availability of Boron on Mars is still unknown. Thus, the Boron supply for Mars’s IEC has to come from Earth. The shipping amount of Boron will depend on the energy needed and the energy produced by p-B¹¹ fusion. The detailed calculation will be addressed later in the manufacturing section.

8.3.3 IEC Jet Torch Experiments

A bench scale experimental device is illustrated in Fig. 8.7, along with an exhaust plasma channel where sample materials to be vaporized are injected into the plasma stream. The resulting ions created by the interaction of the test sample in the plasma flow into and are separated by a quadrupole mass analyzer. For this initial proof-of-principle demonstration, the IEC is run in an electrically-driven mode rather than using a fusion reactor as would ultimately be done in a full scale processing plant.

We have already known the basic physics behind the IEC device. What is different here is that Miley and co-workers found increasing the size of one of the grids allows extraction of plasma “jet” from the central IEC core plasma. This jet is used in Fig. 8.7 to form a flowing plasma channel where the materials, such as water (Miley et al. 2008), are injected and ionized. In this bench test (intended for electrical drive rather than fusion per se), helium is employed in the main chamber to form the plasma and a ~50 kW electrical power supply is used to form and drive the plasma action. (Miley et al. 1994a; Gu and Miley 2000).

In summary, the initial bench experiments using the set up of Fig. 8.7 are intended to demonstrate the efficient ionization and separation of hydrogen and oxygen created when water is injected into the flowing plasma. Later experiments

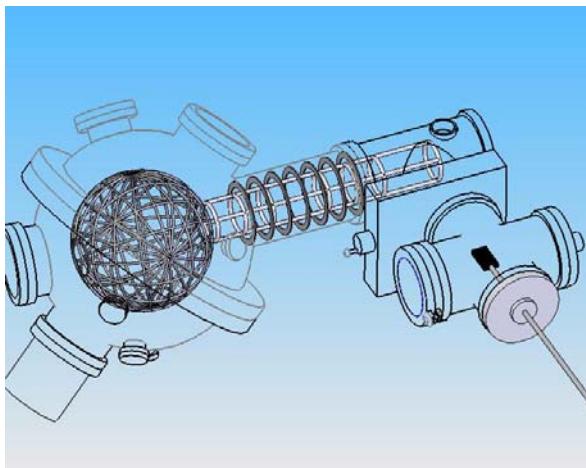


Fig. 8.7 Sketch of proposed 2-kW experimental unit to demonstrate the IEC plasma jet. The jet plasma enters a chamber on the right where a sample has been inserted (sample show in withdrawn position in sketch). Off-gases produced exit through tubing at the top of the interaction chamber and enter a gas collection system Sludge is collected on a tray at the bottom of the interaction chamber for removal and analysis. Baffling

are planned to test insert of various materials, plastics, etc, instead of water. A next major step in the R&D would add an electrostatic energy conversion section to demonstrate the efficient recovery of excess energy invested in the ion formation (Miley 1976). Following the two steps, development would concentrate on a third generation device designed to demonstrate plasma generation/ heating by fusion (vs. electrical drive).

8.3.4 The Application of IEC on Mars Colonization Issues on Developing Habitable Mars Atmosphere

Today, Mars is a cold dry desert world with a thin atmosphere; there is no known life form on Earth that can grow on the Martian surface. However, Mars was not always this way. There is considerable evidence that three to four billion years ago Mars had more Earth-like conditions. For example, a likely aqueous environment was showed from Mars Exploration Rover data at Meridiani (Squyres et al. 2004).

Similarly, Mars might have had a high-oxygen, low-CO₂ atmosphere previously which is much like the present Earth. In the present thin Martian atmosphere, all we know for certain is that there is a small amount of each of these volatiles, but that at one time Mars must have had a much thicker atmosphere. The real thickness could be estimated in variety ways, based on the abundance of the rare gases, theories of atmospheric evolution, the size of the fluvial channels on Mars, and simply by scaling from Earth, but the agreements on how much of each of these volatiles were or are still present on Mars haven't been made. Table 8.2 shows the amount of material needed to make Mars habitable compared to the

range of estimates of the initial abundances for carbon dioxide, nitrogen, and water. Obviously, more Nitrogen, Carbon dioxide and oxygen are needed for converting Mars to a habitable planet. But it is interesting that the estimates of the amount of volatiles initially present on Mars overlaps the range needed for habitability. Then where did these volatiles go? It is possible that some of these volatiles have left the planet permanently, flowing out into space due to Mars' low gravity. However, this has probably not been an important loss term. For example, consider water. The current rate at which water is escaping from the top of the Martian atmosphere corresponds to the loss of a layer of water only about 2 meters thick every three billions years. As it is estimated that Mars had much thicker water needed to make a wetter Mars in the past, it probably still has them locked up in the subsurface and they could be used in the future for creating a biosphere as well. A similar situation could be estimated for other volatiles in previous Mars atmosphere.

Table 8.2 Inventory of Important Volatiles on Mars (McKay and Stoker 1989)

	CO ₂ (atm)	N ₂ (atm)	H ₂ O (meters)
Required for plants	>2.00 (for warming)	>0.01	>500
Required for humans	>0.20 (for O ₂)	>0.30	>500
Present Mars atmosphere	0.001	0.0002	7x10 ⁻⁶
Range of current estimates	0.20-2	0.002-0.30	6-1000

To make Mars habitable, there are two steps we need to follow. First is to raise the Mars temperature, and the second is to form breathable air for human being and all the other animal being (see also Chap. 26). The criteria for choosing gases to be released on Mars involve more than just being efficient greenhouse gases. The compounds must be easily made with elements likely to be abundant on Mars and they should persist in the atmosphere for a long time before being destroyed by Sunlight. Most importantly, they should have no deleterious side effects such as destroying ozone. Nitrogen may be in short supply on Mars and this argues against using large amounts of ammonia or nitrous oxide. Fluorinated compounds look like a good choice. These include CF₄, C₂F₆, C₃F₈ and SF₆. They are all very efficient at absorbing thermal radiation, have long lifetimes in the Martian atmosphere (expected to be thousands of years) and do not adversely affect ozone. Furthermore, carbon, sulfur and fluoride are all abundant elements on Mars. Thus, the most promising mix is a few parts per million of these fluoride based compounds.

After Mars is getting warmer, the release of carbon dioxide from the polar caps and regolith could be triggered to create a thick carbon dioxide atmosphere and augment the greenhouse effect even more, further releasing carbon dioxide and water vapor. The positive feedback would be enough to create a thick warm atmosphere. The thick carbon dioxide atmosphere would be capable of keeping Mars warm without the need for continued artificial greenhouse gas production.

At this stage, plants start to grow and slowly release oxygen through photosynthesis. However, it will be very slow to release enough breathable oxygen for human being, and it is estimated that it would take over 100,000 years to produce a breathable atmosphere on Mars. So, one key to speed up Mars' colonization is to fast oxygen production.

On the other hand, to make Mars' atmosphere breathable, a buffer gas, nitrogen, is also needed. As seen in Table 8.2, the amount of nitrogen needed to make a breathable atmosphere is at the upper limit of the range of estimates of the total initial inventory of nitrogen on Mars. If Mars did start out with that much nitrogen and it is still on the planet, the only possibility is that it is tied up as nitrate in the soil and subsurface. We don't usually see nitrite accumulation in Earth's soil because it is a very soluble salt easily removed by water and it is readily transformed by microorganisms into nitrogen gas. Thus, the only significant accumulations of nitrate on Earth occur in the one place that is so dry that neither liquid water nor microorganisms have been able to recycle the nitrate and has been so dry for long enough that nitrate has accumulated. This place is the Atacama Desert in Chile, possibly the driest place on Earth and one of the oldest deserts with an estimated age of 15 million years. Due to Mars' hyper-aridity and lack of biological activity, nitrogen would slowly have been converted to nitrate by lightning, volcanic-included electric discharge, or meteors it would have accumulated in sediments with no way to recycle it. Thus, a key target for future missions will be to assess the amount of nitrate in the Martian soil, because ferrying nitrogen from Earth is clearly impractical, since the amount need to reach a pressure of 0.3 atmospheres is over 1015 tonnes (compare to the Shuttle lift capability of 40 tonnes). It would also require over several hundred thousand comets of size 1 km radius composed of pure nitrogen to provide this material. If Mars does not have the necessary nitrogen it is not within near-term capabilities of humans to bring it there.

8.3.4.1 Super Greenhouse Gas Generation

On Earth, super greenhouse gas is produced by converting some of the organic wastes, most of which are not present on Mars. Thus, new ways of making them on Mars are required. We know that among all kinds of super greenhouse gases, fluorinated compounds are a good choice. These include CF_4 , C_2F_6 , C_3F_8 and SF_6 . They are all very efficient at absorbing thermal radiation, have long lifetimes in the Martian atmosphere (expected to be thousands of years) and do not adversely affect ozone. Also, carbon, sulfur and fluoride are all abundant elements on Mars.

As for now, hundreds of methods could be used to make fluorinated compounds. Some of them involve F_2 production, and others use fluorine ore directly. On Earth, industrial production of fluorine entails the electrolysis of hydrogen fluoride in the presence of potassium fluoride. This method is based on the pioneering studies by Moissan in the 1880s. Fluorine gas forms at the anode, and hydrogen gas at the cathode. Under these conditions, the potassium fluoride (KF) converts to potassium bifluoride (KHF_2), which is the actual electrolyte. This potassium bifluoride aids electrolysis by greatly increasing the electrical conductivity of the solution:



The HF required for the electrolysis is obtained from calcium fluorides upon treatment with sulfuric acid:



If the above method is incorporated on Mars, IEC will act as a major electric power supply for electrolysis process. In some cases, CaF_2 could be directly used for CF_4 production (Peter 1968):



where carbon supply could be got from CO_2 decomposing product on Mars, and chloride is also abundant in Mars ore. Although, by now the presence of CaF_2 in Mars' soil hasn't been reported, Allen et al. (2000) reported that if microbes ever exist on Mars, CaF_2 could be produced by them, and it is highly possible that CaF_2 is in Mars' soil. The article on JPL-NASA website (Steigerwald 2009) implies that Mars had or have microbes which can release methane. In this case, it is also possible that some other microbes, which can help form CaF_2 , could have existed.

8.3.4.2 Use IEC Torch to Accelerate Mars Habitability

As we know, to develop breathable Mars' atmosphere, more carbon dioxide, oxygen, and nitrogen are needed. While possibly enough carbon dioxide could be achieved by melting dry ice with warmed planet by super green house gas, the release of enough oxygen and nitrogen remains an issue to be addressed. Making an oxygen rich atmosphere to complete the terraforming process is much more difficult. While the energy required only corresponds to 170 year of Martian Sunlight which equals 26281401.6 GJ, the chemical transformation of carbon dioxide to oxygen is intrinsically more difficult than merely thermally warming of the planet. The most likely process for converting an atmosphere worth of carbon dioxide to oxygen is global-scale biological photosynthesis. On the Earth, the efficiency with which plants can produce oxygen from Sunlight is about 0.01%. It is unlikely to be better on Mars. Genetic engineering could improve on this process. It is estimated that it would take over 100,000 years to produce a breathable atmosphere on Mars. While this may sound long, we should keep in mind that the same process on Earth took over two billion years. Besides, to authors' knowledge, an efficient way of producing nitrogen directly from Mars hasn't been proposed yet. To accelerate the formation of breathable air, in this section, we propose to address this issue using IEC fusion torch.

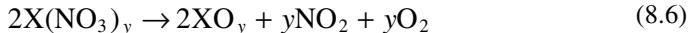
Just as in the water decomposition experiments on the bench scale IEC described earlier, CO_2 could be absorbed directly into the IEC jet exhaust plasma and decomposed into oxygen and carbon, in which carbon latterly could be used for super greenhouse gas CF_4 production. We already calculated that the total energy needed to produce enough oxygen for breathable air through CO_2 decomposing is 262814 TJ, while a 1000 MW IEC could produce this much energy within

10 years, which dramatically advance the time of mankind's landing on Mars. Consider energy released per p-B¹¹ fusion:



The total amount of Boron we needed is about 1 tonnes for CO₂ decomposing, which could be shipped from Earth.

Decomposing nitrate into nitrogen and oxygen may be more complicated. This process must go through two steps. The first step is a thermal heating step which is to convert nitrate into metal oxide, nitrogen dioxide, and oxygen:



The second step is nitrogen dioxide is further decomposed into N₂ and O₂. The initial thermal energy needed for the heating step may be achieved from CO₂ decomposition, during which, electricity could be generated by cascading the high electrical conductivity exit gas through a magneto hydrodynamic process (MHD) followed by a steam Rankin cycle. If the energy equivalent of the carbon is added to the MHD electrical production, a net energy output from the overall system appears possible despite the electrical input required to generate the IEC plasma. The extra energy could be used for nitrate heating step. Once the process started, the heating energy would be slowly achieved from the net energy generated by NO₂ decomposition.

The other immediate application we could find for the application of the torch would be to decompose other abundant ores on Mars. We know that Mars has abundant resources of different ores, such as basalt, andesitic rocks, and silicon glass. Most of Martian surface is deeply covered by a fine iron (III) oxide dust that has the consistency of talcum powder. The IEC fusion torch would be able to decompose them as well to produce pure iron, which is used for steel production and realizing self producing of IEC power units.

8.4 LENR Power Units

8.4.1 Introduction

Cold Fusion first gained attention in 1989 when Pons and Fleischmann reported that nuclear fusion was produced during electrolysis of heavy water on a palladium (Pd) electrode (Voss 1999). They reported excess heat production with a total energy production over extended run times that could not be achieved with chemical reactions, thus only seemed possible source through nuclear processes. They further reported measuring small amounts of nuclear reaction byproducts, including neutrons and tritium (Fleischmann and Pons 1989). These reports raised hopes of a cheap and abundant source of energy (Browne 1989), however, when many laboratories tried but failed to reproduce the Pons Fleischmann results, cold fusion fell into disrepute. Now, after almost 20 years of "low key" research at many labs around the world, the basic physics of the processes involved are becoming better understood and it again appears that this process is destined to have a major impact in the energy field.

The original Pons-Fleishmann (P-F) type reaction involved DD Fusion, where the reaction channel passes through deactivation of the ${}^4\text{He}$ by transferring energy to the lattice which ultimately appears as heat. In association with excess heat, various cold fusion researchers have also reported observing gamma rays, neutrons, and tritium (${}^3\text{H}$) production, with the quantities in excess of background levels. In addition, a number of researchers have reported transmutation reactions that involve interaction between deuterium/hydrogen and atoms in the host lattice, typically heavy metals (see Fig. 8.8). Since such reactions do not involve “fusion” reactions, but mainly “transmutation” type nuclear reactions (often involving fission of the intermediate compound nucleus), this type of reaction has gained the name Low Energy Nuclear Reactions (LENRs).

<u>D-D Reactions</u>		<u>% branching</u>	
		<u>hot fusion</u>	<u>“P-F” type</u>
	$T + p$	50	< 0.1
$D - D$	$\xrightarrow{\quad} \xrightarrow{\quad}$		
	$He-3 + n$	50	$< 10^{-6}$
	$\xrightarrow{\quad}$		
	$He-4 + \text{gamma}$	$< 10^{-5}$	99+

<u>LENR</u>	
$p + \text{metal} \rightarrow$	<i>products or “fission” product array</i>

Fig. 8.8 Comparison of the LENR reactions and the D-D reactions. Occurring in hot fusion and in Ponds-Fleishmann type cold fusion.

Transmutation reactions can be broadly classified according to their products. Some experiments have observed a large array of reaction products with mass numbers ranging across the periodic table. These reactions are traced to multi-body events leading to a heavy compound nucleus, which can both decay, and fission into an array of elements (Steigerwald 2009). The other set of experiments lead to one or few distinct isolated products (Miley et al. 1997c). These reactions may or may not involve multi-bodies but the net result is direct formation of the reaction products as opposed to the disintegration of a compound nucleus.

In the present discussion we will focus on LENR type of reactions since our effort now is to develop this type of system into a small power producing cell suitable for mobile and distributed small power applications on Mars.

Earlier work on LENR (transmutation type) reactions at the University of Illinois involved thin films on microspheres, while the recent studies converted it to thin films coated on a flat alumina substrate as illustrated conceptually in Fig. 8.9. In this design, a double-layer Pd/Ni thin film, at 8000 Å and 1000 Å respectively, was used as a cathode. The electrical current flow is parallel to the thin film surface so that a high current density and high proton flow rate are obtained along with a high deuterium density (Iwamura et al. 2002). It is reported that the excess heat evolution during electrolysis runs was detected compared with smooth Pt by a

high sensitivity open-type calorimeter (Miley et al. 1994b, 2002b). Besides, long exposure of photographic films to electrodes removed after extensive runs indicate low levels of soft x-ray and/or beta particle emission. More importantly, a possible close correlation was found between the product emission and the excess heat by computing the excess power from both (Miley et al. 1994b).

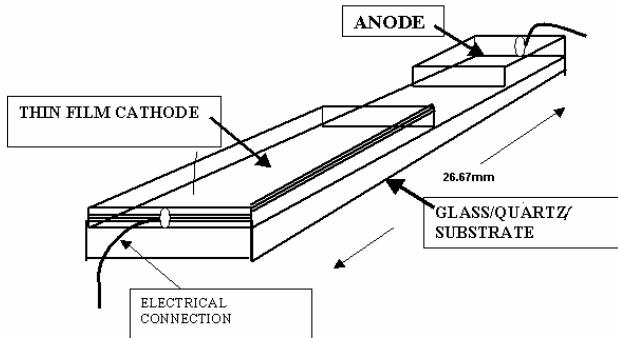


Fig. 8.9 Recent Work Uses a Unique Integrated Thin film Plate-Type Electrode

8.4.2 Small Portable LENR Power Units

The observation of MeV particle emission, combined with the transmutations and excess heat measurements, provide extremely strong evidence for nuclear reactions occurring in the thin films during electrolysis. The results, combined with recent observations of localized D/H condensation sites (Lipson et al. 2005), have made the production of LENR small portable power units possible.

The localized D/H condensation sites in Palladium, which actually D/H condensed in dislocation loops, were achieved by cyclically loading and unloading D/H into Palladium. The Superconducting Quantum Interference Devices (SQUID) magnetic measurements described in Lipson et al. (2005) show “clusters” have characteristics of a type-II superconductor. The hydrogen clusters are strongly condensed in dislocation cores and can only come out at temperature above 300 °C. The dislocation density calculation give hydrogen densities approaching $10^{24}/\text{cc}$, which yields $x_{\text{eff}} \sim 1.8$, suggesting superstoichiometric hydride formation in the deep dislocation cores (see Fig. 8.10). Although excess heat evolution in this type of electrode hasn't been tested yet, with such high D/H density, the density of nuclear reaction is expected to be high.

Although these clusters have extremely local high density, the low fractional volume of the clusters (which is where the fusion producing reaction would occur) limits total reactions to low levels. Material with abundant nano or micro structures may solve this problem. The idea is that the dislocation loops mainly form at the near surface of the material. Nano or micro structures feature with large surface area, thus providing more rooms for dislocation formation. If high volumetric

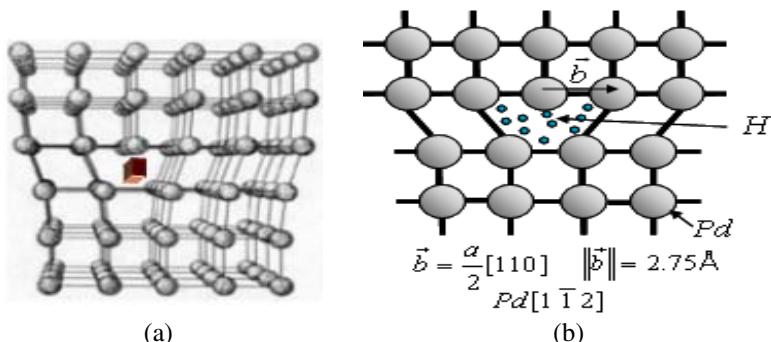


Fig. 8.10 SQUID magnetic measurements show clusters have characteristics of a type-II superconductor. Cluster regions can have hydrogen densities approaching 1024/cc. (a) 3D scheme of typical edge dislocation core (see orange bar) in the Pd crystal. (b) Cross-section of A.

densities of cluster sites can be created via methods outlined earlier, a high reaction rate per cc should result given a competitive power cell at small size. We term the electrode designed to achieve a “massive cluster electrode” (MCE) for controlled cluster reactions. Besides in palladium, dislocations can also form in other materials such as Nickel with directly hydrogen charge, and higher hydrogen loading at dislocation cores than in the bulk material has been reported. Hydrogen interactions at dislocations in Silicon have been extensively studied, but few effort have been made on seeking its application on nuclear reaction field. Once scientists crack the underlying mechanism of hydrogen/deuterium condensation in the defects of solid materials, development of these low-Z host materials is just a time issue.

Based on the excess heat measurements, local power densities exceeding a kW/cc are possible, promising very high energy density power units. Since the radiations emitted (protons, alphas, and x-rays) are not very penetrating (do not escape the cell structure) and no long lived radioactive reaction products are observed, LENR power units would be a remarkable “green” nuclear technology. A sketch of a small D-cell equivalent LENR battery is shown in Fig. 8.11.

The cell shown has a deuterium gas “fuel” tank attached. Gas loading is used rather than electrolysis for compactness. Heat flow is directed to the outer casing through a thermoelectric element using an insulation and heat sink design. Modular sections connected in series allow a 1.5 V output at 0.1 A. This type of battery must be used in devices where natural convection air cooling or other heat flow dissipates heat fro the battery casing. The battery run time is determined by the amount of deuterium stored refilling the depleted tank is provided by pump down and gas injection through a filtered line connected to a “filling station”. The unit shown is designed for 1000 Ahr per gas fill. The main technological step needed before construction of this battery is to finish development of the nano-layered electrode structure described earlier.

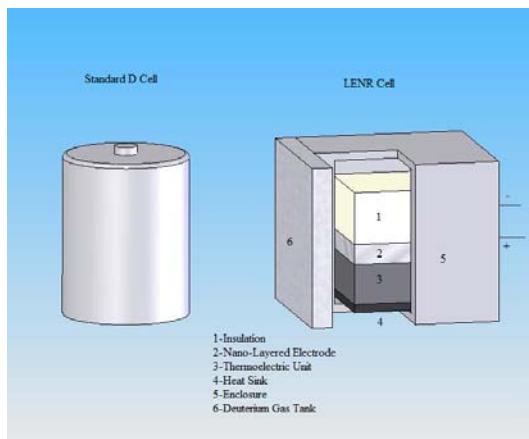


Fig. 8.11 Small LENR Battery -- design based on present experimental data base

8.5 Manufacture IEC and LENR Power Units on MARS

8.5.1 IEC Power Units

To develop IEC on Mars, the main material we need is steel. On Earth, Steel is manufactured by the chemical reduction of iron ore, using an integrated steel manufacturing process or a direct reduction process. In the conventional integrated steel manufacturing process, the iron from the blast furnace is converted to steel in a basic oxygen furnace (BOF) in which oxygen has to keep flowing. Steel can also be made in an electric arc furnace (EAF) from scrap steel and, in some cases, from direct reduced iron. Other than BOF, EAF needs much less oxygen to provide supplemental chemical heat. Mars is full of iron ores (Landis 2009). Thus, we won't have any problem of getting raw materials (see Chap. 20). We already know that when decomposing material on Mars, IEC could produce a big amount of electricity, which could be used in electric arc furnace to reduce iron. Also, IEC torch itself has an electric-arc-like effect between its anode and cathode, and is able to melt and reduce iron directly. Thus, once the first IEC is manufactured, more IEC could be easily produced. Therefore, the question remains how to produce first IEC torch.

With current Mars' atmosphere, it is difficult to light fire because there isn't enough oxygen to burn material. Also, if we do so, it will slow down the timetable of producing breathable air. Thus making steel by BOF is hardly practical, and EAF is the way to go. To initiate IEC running, initial electrical power supply is needed. To achieve this initial power supply, batteries or fuel cells may be choices. Utilizing these small power supply IEC may produce tens of times more power. We then will have continuous supply of electricity to make steel for fabrication of IEC torch.

The power value of the IEC that is about to ship to Mars depends on the shipping ability of the spaceship. As shown in Fig. 8.2, to make a running plant, some

functioning parts have to be attached to center IEC device. Those include cooling systems, energy converters, and so on. For a 1-MWe IEC plant, the IEC part alone is about 0.5 tonnes while the total weight goes to 4.2 tonnes. This isn't too bad, because most of today's rocket can do it. Assume we first ship to Mars a 1MVe IEC along with a robotic system used to help manufacturing another IEC. The task of this IEC and the robots is to produce oxygen, and simultaneously produce enough energy for iron ore mining, steel refining, and manufacturing another IEC with possibly larger power. This process is called self-duplication. Part of produced oxygen will be utilized for steel making. Assume produce another IEC with 100MW power, we may need $100^x \times 4.2$ tonnes steel with $0 < x < 1$ (here we introduce an exponential factor x because in real manufacture process the amount of steel needed is usually not exactly a multiple of the amount of steel used for a 1MW unit). Assume $x = 0.7$, then we need c.a. 106 tonnes steel. The energy needed to produce this amount of steel is 6 MJ per tonnes multiplied with 106 tonnes which yield to 636 MJ (Mohsen and Akash 1998). When taking into account the mining and refining process, the total energy needed is assumed to be 3 times of basic energy 636 MJ which yield 1908 MJ. One 1MW IEC could produce energy 3600 MJ per hour, meaning that the steel needed for next 100MW IEC could be made within less an hour. However, consider the time needed for setting up the ore-mining, steel-refining and further manufacturing instruments, the total time scale may extend to months or several years. The amazing part is that once this steel-making flow process is set up, it will take much less time to manufacture next one, for example, 1000 MW IEC. As we calculated previously, once we get a 1000MW IEC, it will only take about ten years to form enough oxygen for breathing. By considering manufacturing process, the total time needed for producing breathable air could be as short as 20 years (see Table 8.3).

Table 8.3 Mars' breathable air forming progress

Task	Progress Years				
	1	2~5	6~10	11~15	16~20
Bring first 1 MW IEC and robot system to Mars, and get ready to operate.	—	▲			
Construct iron ore mining, steel mining and manufacturing flow process line	—		▲		
Manufacture a new 100 MW IEC on Mars	—		▲		
Manufacture a new 1000 MW IEC on Mars	—		▲		
Construct Nitrate ore mining and decomposing system	—		▲		
Decompose CO ₂ and Nitrate Ore to form breathable air	—		▲		▲

8.5.2 Small Portable LENR Power Units on Mars

By now, analysis of data from Mars by the Odyssey Gamma-Ray Spectrometer (GRS) has produced maps for 6 elements (H, Si, Cl, K, Fe, and Th) (Boynton et al. 2007). Current models of the planet's interior imply a core region about 1,480 kilometers in radius, consisting primarily of iron with about 14–17% sulfur. In June, 2008, the Phoenix Lander returned data showing Martian soil to be slightly alkaline and containing vital nutrients such as magnesium, sodium, potassium and chloride. On July 31, 2008, NASA announced that Phoenix confirmed the presence of water ice on Mars, as predicted on 2002 by the Mars Odyssey orbiter. During the initial heating cycle of a new sample, TEGA's mass spectrometer detected water vapor when the sample temperature reached 0 °C (Turenne 2008) because liquid water is hard to exist on the surface of Mars with its present low atmospheric pressure, except at the lowest elevations for short periods. So, Mars are full of materials we could use to achieve LENR. Developing Small portable LENR power units on Mars would surely help to locally construct convenient portable power supply. An immediate application of these small power units is to provide power supply for initiating IEC operation.

In 2004, President George W. Bush announced his vision for exploring the Moon and Mars. Much of the program matched the outline sketched out in early reports, involving sending astronauts back to the Moon by 2020 and travel on to Mars after that. Although in some ways, the Moon will be harder than Mars, because Mars environment is more close to Earth than Moon's, still, the Moon is ultimately much more forgiving because it is much closer -- 250,000 miles away, while Mars is 34 million miles from Earth at its closest point. If someone needs help on the Moon, it takes three days to get there. By contrast, Mars will be several months away even with the help of advanced -- and as yet nonexistent -- propulsion systems. Thus, after successfully terraforming Moon, some equipment, such as first IEC torch or some LENR cells or both, may be transferred to Mars directly from Moon. As a matter of fact, by the time of Moon terraforming, cargo spaceship system (Orbital Science Co.) may have been fully developed between Moon and Mars or Moon and Earth, which will highly facilitate crew and goods ferrying. The final decision of how exactly we are going to bring these equipments to Mars should be made based on various factors, such as the cost, the detailed plan of Mars' colonization, and so on.

8.6 Summary

The colonization of Mars, although technically definitely feasible, seems to be economically unattractive to most people. But this is only so if you look at Mars as nothing but a vast amount of worthless wasteland just like the Spanish did with North America in the 17th century. Rather, Mars is the only place in the solar system, except for Earth of course, where plants can grow with no artificial illumination and no massive radiation shielding (see Chap. 18). Besides, Mars also has large quantities of water, which will help for exploration, base-building, settlement and, of course, terraforming. These are what make Mars unique and also

why, in past few decades, NASA has been putting huge efforts trying to know better and understand more about Mars, including geography, atmosphere, and temperature. After number of failures, to date, with series successful landing of Mars rovers, mass of information have been achieved from Mars, which, to a great extent, helps NASA to launch more detailed plan for Mars colonization.

Based on the information we have learnt about Mars, we envision possibilities of developing Nuclear Power Plant directly on Mars. In this chapter, two types of fusion power plants, IEC torch and LENR small power units, were proposed. To reach the first step of Mars colonization, forming habitable environment, IEC boron fusion torch may act as a very promising candidate due to its small size, extremely high energy and low radiation. Small portable units, such as LENR power cell and fuel cell, may be a perfect supplement to supply staring power for IEC device. Solar energy, in virtue of vast absorption by super greenhouse gas, could probably used in robotic system or other industry and human use, depending on the amount of its remains.

References

- Allen, C., Albert, F., Chafetz, H., Combie, J., Graham, C., Kieft, T., Kivett, S., McKay, D., Steele, A., Taunton, A., Taylor, M., Thomas-Keppta, K., Westall, F.: Microscopic physical biomarkers in carbonate hot springs: Implications in the Search for Life on Mars. *Icarus* 147, 49–67 (2000)
- Boynton, W.V., Taylor, G.J., Evans, L.G., Reedy, R.G., Starr, R., Janes, D.M., Kerry, K.E., Drake, D.M., Kim, K.J.S., Williams, R.M., Crombie, M.K., Dohm, J.M., Baker, V., Metzger, A.E., Karunatillake, S., Keller, J.M., Newsom, H.E., Arnold, J.R., Brückner, J., Englert, P.A.J., Gasnault, O., Sprague, A.L., Mitrofanov, I., Squyres, S.W., Trombka, J.I., d'Uston, L., Wänke, H., Hamara, D.K.: Concentration of H, Si, Cl, K, Fe, and Th in the low- and mid-latitude regions of Mars. *J. Geophys. Res.* 112, E12S9 (2007)
- Brown, L.C., Besenbruch, G.E., Schultz, K.R., Showalter, S.K., Marshall, A.C., Pickard, P.S., Funk, J.F.: High efficiency generation of hydrogen fuels using thermochemical cycles and nuclear power. In: AIChE 2002 Spring National Meeting, New Orleans, LA (2002)
- Browne, M.: Physicists debunk claim of a new kind of Fusion. *New York Times* (1989)
- Burton, R., Momota, H., Richardson, N., Shaban, Y., Miley, G.: Fusion Ship II- A fast manned interplanetary space vehicle using inertial electrostatic fusion. In: El-Genk (ed.) *Proceedings of Space Technology and Applications International Forum (STAIF 2003)*, AIP Conference Proceedings, vol. 654, pp. 553–562 (2003)
- Bussard, R., Jameson, L., King, K.: Ion-Acoustic waves and ion wave group trapping in IEC systems. *American Physical Society Bulletin* 37, 1582 (1992)
- Bussard, W.: Fusion Technology 19, 273–293 (1991)
- Chacón, L., De Mora, J., Miley, G.: Engineering issues of gridded inertial electrostatic confinement (IEC) devices. In: *Proceedings of IEEE/NPSS SOFE 1997 Symposium*, San Diego, CA, vol. 2, pp. 737–740 (1997)
- Farnsworth, P.: Electric discharge device for producing interactions between nuclei. U.S. Patent No. 3,358,402 (1966)

- Fleischmann, M., Pons, S.: Electrochemically induced nuclear fusion of deuterium. *J. Electroanalytical Chem.* 261, 301–308 (1989)
- Gough, W., Eastlund, B.: The fusion torch-closing the waste cycle from use to reuse. WASH-1132 Division of Research, U.S. Atomic Energy Commission, Washington, D.C (1969)
- Gough, W., Miley, G.: The IEC fusion –plasma torch a path for closing the materials cycle. In: Proceedings of ANS Topic Meeting, on Fusion Energy, San Francisco (2008)
- Gu, Y., Miley, G.: Experimental study of potential structure in a spherical IEC fusion device. *IEEE Transactions on Plasma Science* 28, 331–346 (2000)
- Hirsch, R.: Inertial-electrostatic-confinement of ionized fusion gasses. *J. Appl. Phys.* 38, 4522 (1967)
- Iwamura, Y., Sakano, M., Itoh, T.: Elemental analysis of Pd complexes: effects of D2 gas permeation. *Japan Society of Applied Physics* 41, 4642–4650 (2002)
- Jurczyk, B., Gu, Y., Miley, G.: Resonant ion driven oscillation (RIDO) concept. *Am. Phys. Society Bull.* 42, 1818 (1997)
- Krakowski, R.: Progress in commercial magnetic fusion energy reactor design *Fusion Technology*, vol. 20, p. 121 (1991)
- Landis, G.A.: Meteoritic steel as a construction resource on Marsstar. *Acta Astronautica* 64(2-3), 183–187 (2009)
- Lipson, A., Heuser, B., Castano, C., Miley, G., Lyakhov, B., Mitin, A.: Transport and magnetic anomalies below 70 K in a hydrogen-cycled Pd foil with a thermally grown oxide. *Physical Rev. B* 72, 212507 (2005)
- Luo, N., Miley, G., Mather, J., Burton, R., Hawkins, R., Gimlin, J., Rusek, T., Narayanan, S.: NaBH4/H2O2 fuel cell for lunar and mars exploration. *AIP Conference Proceedings* 813, 209–221 (2006)
- McKay, C., Stoker, C.: The early environment and its evolution on Mars: Implications for life. *Rev. Geophys.* 27, 189–214 (1989)
- Miley, G.: Fusion Energy Conversion. American Nuclear Society, LaGrange (1976)
- Miley, G.: ^3He sources for D- ^3He fusion power. *Nucl. Instrum. Methods Phys. Res. A* 271, 197–202 (1988)
- Miley, G.H., Yamamoto, Y., Javedani, L., Gu, Y., Satsangi, A., Heck, R., Nebel, R., Turner, L., Bussard, R.W., Ohnishi, M., Yoshikawa, K., Momota, H., Tomita, Y.: Hydrogen generation with an electrostatic confinement power source. In: *Hydrogen Generation with an Inertial Electrostatic Confinement Power Source*, The First Int. Conf. on New Energy Systems and Conversions, Yokohama, Japan (1993)
- Miley, G.: The inertial electrostatic confinement (IEC) as a unique space power source. In: Specialist Workshop on Advanced Space Propulsion. Joint Propulsion Laboratory, Pasadena (1994)
- Miley, G., Javedani, J., Nebel, R., Nadler, J., Gu, Y., Satsangi, A., Heck, P.: An inertial electrostatic confinement neutron / proton source. In: Haines, M., Knight, A. (eds.) *AIP Conference Proceedings*, vol. 299, pp. 675–689 (1994a)
- Miley, G., Hora, H., Batyrbekov, E., Zich, R.: Electrolytic cell with multilayer thin-film electrodes. *Trans. Fusion Technology* 26, 313 (1994b)
- Miley, G.: The inertial electrostatic confinement approach to fusion power. In: Panarella, E. (ed.) *Current Trends in International Fusion Research*. Plenum Press, New York (1997)
- Miley, G., Gu, Y., DeMora, J.M., Stubbers, R.A., Hochberg, T.A., Nadler, J.H., Anderl, R.A.: Discharge characteristics of the spherical inertial electrostatic confinement (IEC) device. *IEEE Trans. on Plasma Science* 25, 733 (1997a)

- Miley, G., Gu, Y., Ohnishi, M., Yamamoto, M., Hasegawa, M., Yoshikawa, K.: Potential well structure and scaling studies for the IEC, Abstracts. In: International Sherwood Fusion Theory Conference (1997b)
- Miley, G., Narne, G., Williams, M., Patterson, J., Nix, J., Cravens, D., Hora, H.: Quantitative observation of transmutation products occurring in thin-film coated microscoppheres during electrolysis. Progress in New hydrogen, 629–644 (1997c)
- Miley, G., Bromley, B., Jurczyk, B., Stubbers, R., DeMora, J., Chacon, L., Gu, Y.: Scaling of the inertial electrostatic confinement (IEC) for near-term thrusters and future fusion propulsion. In: El-Genk, M. (ed.) Proceedings of 15th Symposium on Space Nuclear Power and Propulsion (STAIF 1998) Woodbury, New York, AIP Conference Proceedings, vol. 420, pp. 1373–1375 (1998)
- Miley, G.: A D-3He IEC power unit for space applications. In: Proceedings of Space Tech. & Applications International Forum, U. of New Mexico, Albuquerque, NM (1999)
- Miley, G., Castano, C., Lipson, A., Kim, S.-O., Luo, N.: Progresss in development of a low energy reaction cell for distributed power applications. In: Proceedings of ICONE 10: 10th International ASME Conference on Nuclear Engineering, Arlington, Virginia, Track, vol. 8, pp. 1–7 (2002a)
- Miley, G., Heinrich, H., Lipson, A., Kim, S., Luo, N., Castano, C., Woo, T.: Progress in thin-film LENR research at the University of Illinois. In: Proceedings of ICCF-9, Tsinghua University, Beijing (2002b)
- Miley, G., Gough, W., Leon, H.: Large-scale hydrogen production using a fusion torch process. In: Proceedings Fuel Cell Seminar, Phoenix, Arizona (2008)
- Mohsen, M., Akash, B.: Energy analysis of the steel making industry. International J. Energy Res. 22, 1049–1054 (1998)
- Peter, K.: Process for continuous production of carbon tetrafluoride. U.S. Patent No. 3,386,989 (1968)
- Salisbury, W.: Method and apparatus for producing neutrons. U.S. Patent No. 2,489,436 (1947)
- Schultz, K., Brown, L., Besenbruch, G., Hamilton, C.: Large-scale production of hydrogen by nuclear energy for the hydrogen economy. GA-A24265 (2003)
- Shaban, Y., Miley, G.: Distributed power sources for Mars colonization. In: El-Genk, M. (ed.) Space Technology and Applications International Forum, American Institute of Physics, Conference Proceedings, pp. 1211–1218 (2003)
- Squyres, S.W., Arvidson, R.E., Bell, J.F., Brückner, J., Cabro, N.A., Calvin, W., Carr, M.H., Christensen, P.R., Clark, B.C., Crumpler, L., Des Marais, D.J., d'Uston, C., Economou, T., Farmer, J., Farrand, W., Folkner, W., Golombek, M., Gorevan, S., Grant, J.A., Greeley, R., Grotzinger, J., Haskin, L., Herkenhoff, K.E., Hviid, S., Johnson, J., Klingelhöfer, G., Knoll, A.H., Landis, G., Lemmon, M., Li, R., Madsen, M.B., Malin, M.C., McLennan, S.M., McSween, H.Y., Ming, D.W., Moersch, J., Morris, R.V., Parker, T., Rice Jr., J.W., Richter, L., Rieder, R., Sims, M., Smith, M., Smith, P., Soderblom, L.A., Sullivan, R., Wänke, H., Wdowiak, T., Wolff, M., Yen, A.: The opportunity Rover's Athena Science Investigation at Meridiani Planum, Mars. Science, 1698–1703 (2004)
- Steigerwald, B.: Martian Methane Reveals the Red Planet is not a Dead Planet (2009), http://www.nasa.gov/mission_pages/mars/news/marsmethane.html (Accessed April 22, 2009)
- Sved, J.: The commercial IEC portable neutron source. Trans. ANS 77, 306 (1992)

- Sved, J., Kulcinski, G., Miley, G.: A commercial lunar helium-3 fusion power infrastructure. *J. British Interplanet Society* 48, 55–61 (1995)
- Turenne, V.D.: NASA spacecraft finds water on Mars. *Los Angles Times* (2008),
<http://latimesblogs.latimes.com/lanow/2008/07/nasa-spacecraft.html> (Accessed July 31, 2008)
- Van Devender, J., Cook, D.: Inertial confinement fusion with light ion beams. *Science* 232, 831 (1986)
- Voss, D.: What Ever Happened to Cold Fusion. *Physics World* (1999), <http://physicsworld.com/cws/article/print/1258> (Accessed May 1, 1999)
- Williams, C., Borowski, S.: An assessment of fusion space propulsion concepts and desired operating parameters for fast solar system travel. In: *Proceedings 33rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*, Seattle, WA (1997)

Chapter 9

Advanced Nuclear Compact Structures for Power Generation on Mars

Liviu Popa-Simil

LAVM Co., Los Alamos, USA

9.1 Introduction

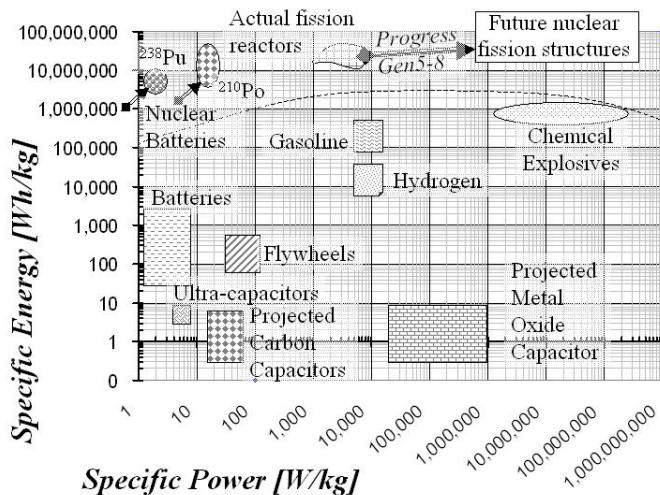
Mars represents the first big extraterrestrial step of mankind. This is a very important historic moment representing the beginning of the cosmic age. This step and the steps that follow cannot be done without strong, compact reliable power sources. Nuclear power is the strongest contender for life support systems on Mars. The importance and advantages of nuclear power sources is shown in Fig. 9.1.

9.2 Radioisotope Thermo-Electric Power

Thermo-electric nuclear power production relies mainly on alpha generating radioisotopes to create the heat supply and the efficiency of the thermoelectric material. The actual efficiency of thermoelectric material is low, being between 4-8%, while the operating temperatures are also low, limiting the maximum temperature difference to 500°C between the facets of the thermopile.

The limiting factor being the total power produced in the world, which limits the total power, and amount of nuclear isotopic materials being used as possible heat sources. The maximum production of about 80 TWDay/year (Toffer 1992) seems to have been reached by 1984, and has been in a slight decline since then. 80 TWDay/year implies that the potential total amount of ^{239}Pu in the waste fuel of about 50t/year.

The amount of minor actinides left in the burned fuel (about 1 t/year) also implies an estimated maximum thermal power of about 1MW/year possible of being obtained if all of the planet's nuclear production will be effectively used. This simple estimation shows us that the total available power possible derived from isotopic sources is drastically limited, and stresses the need for advanced conversion systems in reactors and generation systems, while at the production level is reinforcing the call for better and more specialized production systems.

**Fig. 9.1** Various sources of energy/power**Table 9.1** Test Results from a KE Reactor for a successful ^{238}Pu production demonstration test carried out in 1967 (Toffer 1992)

Neptunium Tests		Target Type	
g^{238}Pu	/ g^{237}Np	Water Core	Graphite Core
0.125		0.112	
Quality (^{238}Pu /total Pu)		0.877	0.894
^{236}Pu in ^{238}Pu		0.39; 0.52	0.51

Table 9.2 Alpha rays of various plutonium isotopes

Isotope	α_1 MeV	α_2 %	Half Life MeV	%	[y]
^{238}Pu	5.50	72	5.46	28	87.5
^{239}Pu	5.16	88	5.11	11	24,390
^{240}Pu	5.17	76	5.12	24	6.58
^{241}Pu	β				13.2
^{241}Am	5.49	85	5.44	13	458

The most valuable isotopes for space applications are ^{238}Pu and ^{241}Am due to their long lifetimes of 87.5 years and 458 years, respectively with specific powers of 546 W/kg and 187 W/kg respectively, but only very limited amounts are available. This shows one path of obtaining ^{238}Pu in nuclear reactors with main production features given in Table 9.1. The features of each of the decay of this isotope are given in Table 9.2.

9.2 Reactor Concepts and Designs for Mars

9.2.1 Previous Developments

Among the many ideas, concepts, and designs, the most realistic solutions are (McGinnis 2004):

- The Pellet Bed Reactor System (a relatively new design, whereas the Potassium Rankine system has been around since the 1960s).
- The SP100 derivative (an evolutionary design from the SP100 era of the late 1980s and early nineties).

9.2.1.1 Pellet Bed Reactor Concept (PeBR)

The PeBRR utilizes a common reactor type for both Nuclear Thermal Propulsion (NTP) and Nuclear Electric Propulsion (NEP) applications. The system's overall specific mass is 3.29 kg/kW_e.

The Nuclear Subsystem

The PeBR is a helium gas cooled, fast spectrum reactor. The fuel used is in the form of 10mm diameter pellets. The core is separated into three 120-degree sectors. Each sector is fully contained and can be operated individually or in concert with one or two other sections. Each of the 120-degree sections has the power of 16.3 MW_t and is filled with fuel pellet (similar to a “pool of balls”). Each fuel pellet contains TRISO-type fuel made of micro-spheres dispersed in a graphite matrix. The microspheres contain a UC-ZrC fuel kernel (approx. 400-500 micrometers in diameter) with three coatings:

- An inner coating of low density, pyrolytic graphite 15 microns thick
- An intermediate coating of high density graphite 5-10 microns thick
- An outer coating of ZrC 10 microns thick

The fuel beads serve as the first pressure vessel for the containment of fission products. UC-ZrC fuel has a melting temperature of 3693 ± 20 K with low carbon content and 3123 ± 50 K with high carbon content.

The shield is made of LiH and W in a double-layered shadow shield located outside of the reactor pressure vessel. The reactor is cooled by recirculated helium and maintained below 500K and helped by potassium radiators that were attached on the shields to remove the decay heat from the vessel via conduction to potassium IR heat sinks into space.

Two independent control systems are implemented with 15-segmented Be₂C/B₄C control drums spaced equally within the radial Be₂C reflector and 6 B₄C safety rods at 0.2 m radial distance from the center. The control drums keep the reactors subcritical during launch and in conjunction with the control rods are sufficient to prevent criticality in a water submersion casualty.

Power Conversion Subsystem

The three individual sectors of the PeBR system are connected to a closed Brayton power conversion system using the helium gas coolant to produce 5MW_e .

Heat Rejection Subsystem

The heat pipe radiator has 5300 m^2 working at 645°K , with an emissivity of 80%, radiating a power density of about 7.85kW/m^2 , totaling 41.6MW .

The advantages of this reactor design with main parameters presented in Table 9.3 are:

- Launching safety with or without fuel
- Capability of refueling in orbit
- Multiple layers of Redundancy
- Negative worth of temperature
- Passive decay heat removal system
- Non-nuclear testing possibilities

The drawbacks are related to low efficiency and material issues that makes the safe operating temperature lower than predicted.

Table 9.3 Summary of design parameters

Total Thermal Power	50MWt	Maximum Fuel Temperture	2200K
Total Electric Power	15MWe	Overall System Efficiency	29.1%
Total System Mass	49300kg	Specific Mass	3.29kg/kWe

9.2.1.2 Potassium Rankine System (KRS)

This fast reactor system has been developed since the 1960s by the competition among three different organizations: NASA sponsored 2001 Rocketdyne design, ORNL 1983 design based on the Medium Powered Reactor Experiment (MPRE) design of 1966, and “1993 NASA internal study on a NEP architecture for a manned Mars mission using” KRS based on a scale up of SP-100. Features are summarized in Tables 9.4 to 9.7.

When placed next to each other these similar KRS designs show different assumptions that translate in to grossly different specific masses(McGinnis 2004). Each system utilizes the same radiator type and working fluid, the same fuel material, and the same shielding materials, yet the specific masses range from $4\text{-}10 \text{ kg/kW}_e$. They differ due to the burn up assumption of the UN fuel for the Rocketdyne system. They assume a 25% burnup, yet BWX Technologies has shown that UN performs well between six and ten percent burnup (Carmack 2004). Beyond 10% fuel burnup, fission product buildup causes swelling and cracking of the cladding structure. The other two designs show a significant difference in their specific masses plus a difference in the required amount of power to send a

Table 9.4 Nuclear subsystems. Comparative analysis

Rocketdyne 2001	ORNL	NASA
Liquid metal cooled and hexagonally cermet fueled with UNW-25Re	Fuel pin, UN pellet with T-111 alloy (Ta-8W-2Hf) cladding with a W liner between the UN and T-111 cladding.	1993 Design Reference Mission (DRM) utilizes NTP as the benchmark propulsion system UN pin type, SP-100 derived
One reactor and three power conversion systems producing 10MWe	burnup is up to 6%.	Power 24MWt
Cooled with liquid lithium operating temperature is 1550K based on refractory metals limit	Operating temperature of 1365K supplies 99% quality liquid potassium to the turbines	Lithium cooled reactor outlet temperature is 1375K and utilizes a potassium for power conversion
The shield is a monocoque shape with a 17.4 degree cone half angle. It is made with a reactor side of tungsten 8.2 cm thick and a payload side of Be ₂ C/B ₄ C 62.6 cm thick.	The shield is made off alternating layers of LiH and W with varying thickness between 50-70 cm. The dose rate 20m away from the shield is: - 1000 neutrons/cm ² and - 1.06 Sv of gamma.	Cermets fuel has a higher strength and higher thermal conductivity than the pin type. There is more data and experience with pin fuels than cermets
The dose at 100m is: 0.05 Sv/yr		

Table 9.5 Power Conversion Subsystem

Rocketdyne 2001	ORNL	NASA
KRS utilizes three turbines with one backup cooling the reactor concurrently; Turbine inlet temperature limit is 1350K	KRS seven stage turbine has a 83% efficiency, spinning at 1050 rpm	two reactors utilize three 2MWe KRSs for a fifty percent redundancy feature
The output voltage is 10,000V _{DC}	The output voltage is 1000V at 2000Hz	The output voltage is 1400V _{DC} at 2000Hz ;total power 8MWe for the manned mission and 4MWe for the unmanned;
This temperature reduces the amount of tantalum necessary in the primary loop and allows niobium alloys which are lighter to be used.		1400V _{DC} was shown to be lighter than a more efficient 5000V _{DC} PMAD system

Table 9.6 Heat Rejection Subsystem

Rocketdyne 2001	ORNL	NASA
The heat pipe radiator is a carbon/carbon potassium working fluid composite structure with Nb-1%Zr coating.	The primary radiator has three manifolds of 17.8 meters each operating at 1020K They are arranged in three flat plate manifolds, 120 degrees apart	The primary heat rejection temperature is 975K in the potassium heat pipe radiators
The auxiliary radiators use both mercury and water for working fluids. Its rejection temperature is between 1000-1025K total heat transfer area is 899m ²	Each heat pipe is 2.5 cm diameter with a 1.7m inner set and 4.42 m outer set. Combined the total heat transfer area is 660m ²	
Advanced radiator studies have shown a possible mass savings of 6480kg for a 10MWe 2yr mission system. Micrometeoroid impact analysis was done on the radiator structure		Mass estimates are based on a 5kg/m ² which are SP-100 based

Table 9.7 Comparisons

Parameter \ System	Rocketdyne	ORNL	NASA
Power Output(MWe)	10	5	4
Lifetime (years)	2	5	5
Reactor Power (MWt)	52	28	24
Overall Efficiency(%)	19.2	17.9	16.7
Number of Systems	1	1	2
Fuel Burnup %	25	6	6
Fuel Type	Hex-Cermet	Pin	Pin
Payload distance from Rx (m)	100	20	60
Voltage (kV _{DC})	10	1	1.4
Frequency (kHz)	0.46	2	2
Radiator Type	C/C K-heat pipe	C/C K-heat pipe	C/C K-heat pipe
Radiator Temperature (K)	1000	1020	975
Waste Heat (MW)	42	23	20
Radiator Area (m ²)	899	660	636
Waste Heat/Area (kW/m ²)	46.72	34.85	31.45
Radiator (kg/m ²)	3.87	8.33	6.57
Number of Power Converters	4	1	3
% PC redundancy	25	0	50
Reactor mass (kg)	4500	3500	3810
Shield mass (kg)	6930	11000	9760

Table 9.7 (continued)

Parameter \ System	Rocketdyne	ORNL	NASA
Primary, Auxiliary loop (kg)	11702	in shield	in shield
Power Conversion System (kg)	12060	4500	7860
Heat Rejection System (kg)	4435	5500	4180
Structure (kg)	included	included	4650
Power Conditioning System (kg)	468	included	11250
Total mass (kg)	40095	24500	41510
Specific Mass (kg/kWe)	4.01	4.9	10.38

manned mission to Mars (the NASA mission requires two 4MW_e power sources). 5MW_e is not sufficient to meet the duration in space requirement for a manned Mars mission.

9.2.1.3 Scaleability of the Sp-100 Program

The Mars mission's need for a larger power range made the SP-100 the benchmark for future reactor designs. In 1991, the space-power-system study scaled up the SP-100 to 5, 10, and 40 MW_e systems. The fast reactor has highly enriched UN fuel, is lithium cooled by a radiator with 6kg/m² specific mass. The design life span was 10 years of operation instead of an operational life of ten years and full power lifetime of 7.3 years (Newkirk 1991). The operational temperature was increased from 1350K to 1400K due to some turbine inlet material advances. Table 9.8 presents a summary of the final design parameters (General Electric Company 1988).

Table 9.8 Performance of scalable reactor SP-100 derivatives

System	SP-100	Rankine	Brayton
Power (MW _t)	2.4	25	25
Output power (MW _e)	0.1	5	5
Lifetime (years)	7.3	10	10
Reactor mass (kg)	700	4200	4200
Shield (kg)	1037	3930	3930
Structure (kg)	538	incl in S/S	incl in S/S
Primary Heat Transport System (kg)	500	1550	1550
Power Conversion System (kg)	409	31140	56140
Heat Rejection System (kg)	1027	4180	4180
Power Conditioning System (kg)	399	incl in PC	incl in PC
Total Mass (kg)	4610	45000	70000
Specific Mass (kg/kWe)	43.8	9	14

9.2.1.4 Power System Designs Supplementary Issues

The operational concerns that have been considered are

Restart Capability

Restart capability is a must for any operational reactor because it has to allow for any shutdown. An automatic reactor shutdown mechanism must be in place to protect the reactor from an accident, which would render the reactor useless. SNAP-10A did not have a restart capability, but its power was not critical for life support. The procedure for restarting a reactor must be quick and effective. Presently, it is believed that the loss of the USS Thresher (with all hands on April 10, 1963) might have been avoided if the reactor could have been restarted in real time. Now recovery startup procedures have been implemented later on all US naval reactors. An additional problem is using external power to keep the liquid metal coolant from solidifying, since the heat of decay is usually not enough to keep the coolant molten. For a manned Mars mission, the communication lag between earth and Mars requires one member of the crew to have adequate knowledge to determine the cause of any shutdown, understand the implications, and be able to correct it quickly enough to safely restart the reactor in order to limit thermal stresses. Some expert computer systems may be designed to assist in this operation, but the ultimate decision must be made by the crew.

Backup Power Requirements

Nuclear submarine reactor designs are the best example of backup systems because of durability and flexibility in the power source design. The reactor continually charges a battery bank, and a diesel engine is always available to provide power on or near the surface of the water. If the reactor should shutdown, the battery provides enough power to maintain crucial systems while the reactor's problems are determined. Battery power is sufficient to bridge the gap between reactor shutdown and subsequent restart, or for switching over to diesel power. Space nuclear power does not have the luxury of being "near the surface" (McGinnis 2004), which is why reliability is very important, and cannot be assured redundancy because redundant faulty designs do not combine to make an overall safe design.

NASA believes that launching a cargo mission first, months earlier using the same power and propulsion systems as for the human mission, is mandatory for a safe manned flight.

Backup power sources could be RTGs, fuel cells, or solar cells (a backup chemical system would defeat the purpose of using NEP). RTGs could be used in the primary piping as a heat source for maintaining the coolant above melting temperature. Fuel cells could use the same hydrogen used for propulsion or for shielding.

For a manned mission, the redundancy rules require at least two independent reactors and a supplementary backup power source to be used for life support and to ensure operability and have the capability to support hotel loads while supplying power for a reactor restart and maintaining the coolant temperature.

Power Management and Distribution

One of the vital systems in space is the power management and distribution system because it allocates the correct power and maintains the balance between the

sources and consumers. The trend of electric voltages is to increase from 120Vdc to 200Vac for the period up to 2020 to 5000Vac by 2030 (Mason 2001), while for propulsion a minimum of 10 kV seems to be required. The specific mass of this complex yet necessary system may scale 1.75kg/kW_e (Metcalf 1993).

Launch and Assembly Safety

The most critical moments for a mission in space is launching and landing that is why the safe placement of the reactor in space and the safe assembly of the exploration vehicle while in space pose the highest concern for policy makers and the public.

The safety and flexibility of the space reactor's design allows fuel to be launched separately from the vehicle. Reliable launch vehicles are always a concern due to the expense of the payload. No launch vehicle can be considered 100% reliable. Launch failure poses the most immediate danger and the most important design feature is to have the fuel burn up in re-entry if the launch vehicle does not make it into orbit. The most dangerous type of accident considered to date is water submersion when the reactor has to be prevented from becoming critical. One solution is to add nuclear poison plugs. SP-100 was designed to remain intact upon reentry, and the two RTGs lost in the American space program did not release any radioactivity to the environment, while the two Russian nuclear powered spacecraft re-entered the atmosphere and produced severe contamination to environment. Cosmos 954 fell on January 24, 1978 in Canada and scattered radioactive debris in an uninhabited area. One piece of Cosmos 954 gave off 2 Sv/h (Space.com 2006), and Cosmos 1402 fell in the Indian Ocean south of Diego Garcia on January 23, 1983. Recently, it was reported that the Russian RORSATs leaked 360 pounds of NaK coolant in a 900 km orbit, posing a space debris hazard (Space.com 2003).

Attitude Determination and Control System

Using space nuclear power, a boom or truss structure is used to reduce the radiation emitted from the reactor without using appropriate mass shielding. Large truss structures require complicated navigation control systems and an analysis must be done in order to decide whether a large control moment gyro (CMG), large flywheel system, or thrust vector control (TVC) will be used. This is not the case for the reactors delivered to the Martian surface, which are designed to be loaded and reach criticality on the surface.

Shield Requirements for Necessary Extra-Vehicular Activity (EVA)

Most space reactors do not require any maintenance to be performed on them in operation. The crew and any sensitive equipment stay in the shadow of the shields of the reactors. Docking to other spacecraft or exploration vehicle must consider the shadow space distribution versus the exposure time.

Artificial Gravity (AG)

Artificial gravity is desirable for long duration space flights because the lack of gravity reduces bone density over time. AG brings the benefits of astronaut health

and ease of testing for reactor components. Centripetal forces of up to few revolutions per minute are a substitute for mass gravity, but add complications because part of the ship must be despun in order to provide stable propulsion or during vehicle rendezvous. For a nuclear reactor that is designed to operate in 1g the payload have to act as a counterweight and to have a despun system to off load the astronauts. The other solution is to design the reactor to operate independently of gravity, making it independent of the artificial gravity systems. Paramagnetic AG is in the very early development stages, but aims towards the same solutions for the nuclear reactor.

9.2.1.5 Mass Reduction

The biggest cost of the nuclear power in space is the cost of putting the power source in orbit costing between \$10-50k/kg (SpaceX Launch Services 2008). Because the nuclear reactor mass is large, the cost of putting it on Low Earth Orbit (LEO) has a high cost; therefore, means of mass reduction are in high demand.

Shield Mass Reduction

Because the shield weighs more than $\frac{1}{4}$ of the total system mass, it is beneficial to reduce its contribution as much as possible. The shield mass may be lowered by reducing the thickness of the lithium hydride and tungsten while compensating by optimizing the astronaut shield for cosmic radiation and the truss length. In this way, a part of the mass is moving from the shield to truss, and an integrated power system and habitat module design is required.

Increasing Power Conversion Efficiency

A higher efficiency would reduce the system and radiator mass along with the mass of nuclear fuel required for a given power level.

1. Thermal Photovoltaics

The downside of direct energy conversion is it has low efficiency, but static energy conversion may be considered in the 1-10MWe range. A new static energy conversion system, thermal photovoltaics (TPV), converts heat energy $>0.5\text{eV}$ into electrical energy. These cells operate in the band where reactors reject heat the most, increasing the potential conversion efficiency of the system up to possibly 40% (McGinnis 2004). TPVs are made of low band-gap semiconductor materials, such as Indium Gallium Arsenide. Low Band Gap semiconductors are unstable at high temperatures and require the TPV cells to be maintained under 500K, but this also increases the IR heat sink.

2. Increasing Operating Temperatures

Raising operating temperatures increases conversion efficiency but decreases the reliability due to materials issues. Each system has to be thoroughly matched in operation temperatures and safety requirements (Rockwell International 2001).

Even higher temperatures, above 1550K, bring serious material issues, solved by thickness and subsequent mass increases that make the gain in performances irrelevant.

3. Turbine Material Advances

New inlet turbine blade materials, such as ceramics, increase the turbines efficiency and are a potential source of indirect mass savings.

4. Overall Material Advances

All kinds of materials are envisioned to be used in structures at high temperature and high thermal conductivity. Carbon foils and alloys may possibly improve the primary piping, boiler/reheater, and liquid piping, replacing the heavier refractory alloys as T-111, ASTAR-811C and Nb-1Zr in use today.

5. Advanced Radiator Concepts

There are several high-flow-heat sinks used as heat pipe radiators exhibiting IR emissivities of 85% at 1000K, rejecting up to 48 kW/m².

New higher emissivity coatings, or other concepts such as the liquid droplet radiator and the pumped loop radiator increase the heat flow from the surface.

9.2.1.6 Recommendations and Observations

Early Decisions

Each subsystem's modus operandi, advantages, and disadvantages have to be considered for the decision to send a nuclear fission source in space. Constraints such as funding limitations add also to the aspects of different types of fuels, shielding, heat rejection, and power conversion techniques; therefore, the earliest decisions in the design process are the most critical. The first reactor to be produced to go in space is under Project Prometheus for the Jupiter Icy Moons Orbiter, JIMO (NASA 2008). The technologies developed and implemented for Prometheus will be the basis for the development of a multi-megawatt power system for future manned exploration of Mars. Other recommendations on viable Multi-megawatt designs are given in Table 9.9.

Table 9.9 List of recommendations on viable Multi-megawatt designs

Specific mass	< 4kg/kWe	Stirling power conversion	>200kWe
Operating temperatures	>1550K	Burnup	> 8%
Radiators specific mass	<5kg/m ²	Restart/backup capability	Yes

Recommendations on Technology Investment Decisions

The following developments are needed:

- Nuclear fuels development, production, and irradiation testing
- Heavy Lift Launch Vehicle (HLLV) development
- Refractory alloy manufacture, production, irradiation, and creep testing.
- Potassium Rankine duration testing.
- Long duration, high power, electric propulsion development and testing.
- Launch vehicle development (even if shuttle derived, is a long process).

- Refractory elements production, which are the backbone for space nuclear power systems.
- Electric propulsion systems, for the mission to Mars need to be developed further.
- High power, long duration systems are needed both for propulsion and power.

9.3 The Present State of Art

Following the natural development of the space race, several countries have started to develop concepts to provide human missions to Mars. The NASA KRS concepts have been developed and updated using many ideas based on LANL developments. One of the most advanced designs is HOMER (Poston 2000), built at Los Alamos National Laboratory in a group with a long tradition in space nuclear applications.

9.3.1 HOMER (Heat Pipe Operated Mars Exploratory Reactor)

HOMER contains stainless-steel (SS)-clad, uranium dioxide (UO_2) fuel pins that are structurally and thermally bonded to SS/sodium (Na) heat-pipes conducting heat from the fuel pins, out to an ex-core power conversion system (Poston 2000).

HOMER is a very simple, modular, system with almost all core components made of Stainless Steel; therefore, material and fabrication costs are low. Fuel may be added to or removed from the core easily making the system's transport and storage easier and less costly. Passive sub-criticality is done by permanently placing boron carbide in the core. By taking advantage of the neutron spectral shift during water immersion scenarios, about 80 kg of mass is added to the reactor, but it simultaneously provides absolute safety and the most reliable system operation. HOMER also has passive removal of decay heat and is essentially non-radioactive before operation.

HOMER's thermal performance is very good, being designed with a large thermal margin. Maximum reactor power is determined only by material temperature limits. For SS, the peak temperatures have to remain below 1250 K. During nominal operation, the peak SS temperature is 1120 K. The power level of an HPS is limited largely by the conductivity and maximum operating temperature of the cladding material, which is why refractory cores have much higher power levels than super-alloy cores. HOMER is designed to produce full power even if a heat-pipe has failed and there is no contact between the failed module and other modules. Full power still can be provided with multiple not-adjacent heat-pipe failures, but power has to be lowered to avoid exceeding temperature limits. Another advantage of the system is that the thermal performance of the entire system can be well established with resistive heater testing.

9.3.1.1 Nuclear Performance of HOMER

HOMER's nuclear performance is given by the robustness of its design. Subsystem components (power conversion, control drives, electronics, etc.) will have to be shielded as necessary. The cold-clean k_{eff} is 1.03, the temperature defect and

burnup reactivity are each less than 1% being a fast reactor where the temperature reactivity feedback is small (and negative), and the neutron lifetime is relatively short. Zero-power criticales will be required to verify all nuclear characteristics that pertain to safety, which may require an additional system.

Nuclear irradiation effects normally will make SS harder decreasing its ductility, but the stresses and fluences are low in the HOMER design; therefore, 5 years at full power produces a peak fluence of $3 \times 10^{21} \text{ n/cm}^2$.

An important factor to consider is the effects of the Martian atmosphere. HOMER is designed to take advantage of the Martian atmosphere, since it is not hermetically sealed and exposed to CO₂ special effects of carburizing /oxidizing SS and giving it increased IR emissivity.

Some protection may still be required to prevent dust buildup. Exposure to CO₂ improves thermal performance of SS by providing a thermal conduction path in gaps where pins/heat-pipes do not have good thermal contact. The atmosphere can also be used as a heat sink by flowing CO₂ in a radiator. Martian gravity may aid in heat-pipe performance by allowing a more efficient use of the Rankine conversion cycle. The atmosphere and/or other indigenous materials may be used for shielding in place there by reducing transported mass. Problems will be caused by dust clogging up or coating control drives and other components for any systems deployed on Mars as well as hot structural parts embitterment by carburetion in the CO₂ atmosphere and will all require encapsulation.

The thermal power scalability of most HPS reactors is determined by material temperature limits and fuel to heat pipe ratio. At very high powers, the performance starts to become limited by heat-pipe or fuel heat conductivity loss due to burnup.

HOMER already has a high heat-pipe-to-fuel-pin ratio, giving increased reliability that may be adjusted depending on the mission.

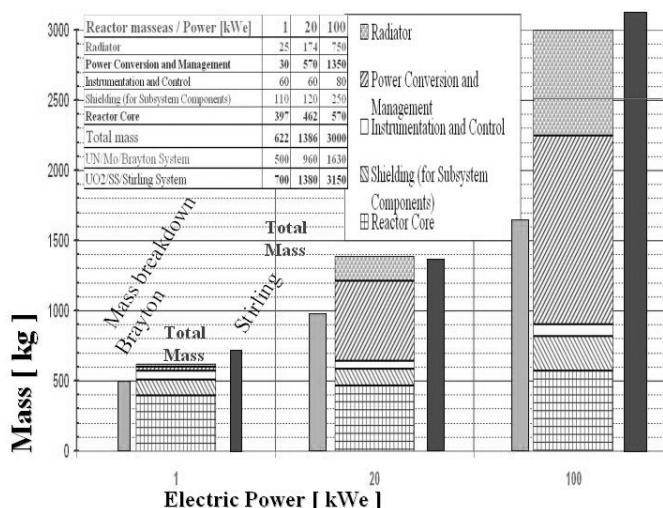


Fig. 9.2 HOMER masses versus Electric Power

Actually, HOMER has a practical thermal power range of 25 to 500 kWt. If refractory metals can be used, this raises the upper limit to >1 MW. The fact must also be considered that the delivered electric power has to be matched with the removal of low-grade heat without impeding the life support systems. However, at the moment it does not seem to be a problem on Mars.

Rough mass estimates at various powers are given by using the same methodology as that used above for the 20 kWe design (Poston 2000). Scaling was based on how each component depends on reactor size and power level, with 20% contingency absorbed into each component. The bar graph in Fig. 9.2 shows the mass of each component at various power levels. Figure 9.2 shows that an improved HPS design, using UN instead of UO₂ fuel, molybdenum cladding, instead of SS, and Brayton conversion (35 % conversion efficiency and a radiator temperature of 573 K) instead of Stirling, as power increases, the advantage of the Brayton system becomes even greater. It is possible that at low powers (<5 kWe), a static power conversion system [thermoelectric, thermionic, or alkali metal thermal-to-electric conversion (AMTEC)] could be optimal for certain applications (Lipinski 1999).

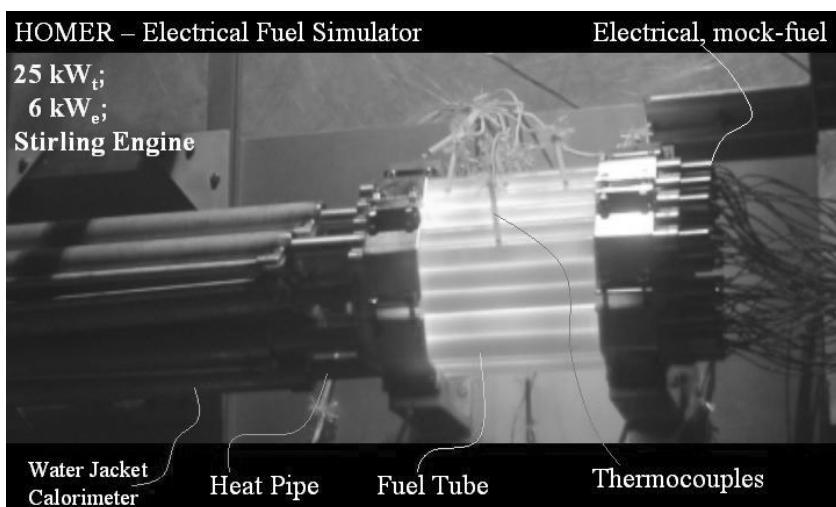


Fig. 9.3 HOMER tests (courtesy of LANL)

Finally, HOMER is a robust, low technical risk fission system that can provide electric power on the surface of Mars and is ready to use technology. The system relies on the Heat-pipe Power System (HPS) concept that has been under development for several years at LANL (Poston D. L 2000). HOMER is designed to be a near-term, low-cost system by using existing technology, designing the prototype system for electrically heated testing, and keeping the system simple. Hardware testing currently is being successfully conducted proving HOMER as a viable and attractive option for Mars surface power. Figure 9.3 shows a prototype HOMER

reactor on the test bench using the electrical fuel simulator to generate the required heat, and testing the heat-pipe to gas heat exchanger and the heat flow throughput. Figure 9.4 shows HOMER on the test bench coupled to a Stirling engine that uses the power by driving an alternator. All of the thermal conditions in space and on Mars are simulated using mock electric fuel. The nuclear fuel and the reactor control operation are also tested. The fuel sent in space with the reactor has to be tested in sub-critical short-pulse conditions only in order not to accumulate radioactive material.

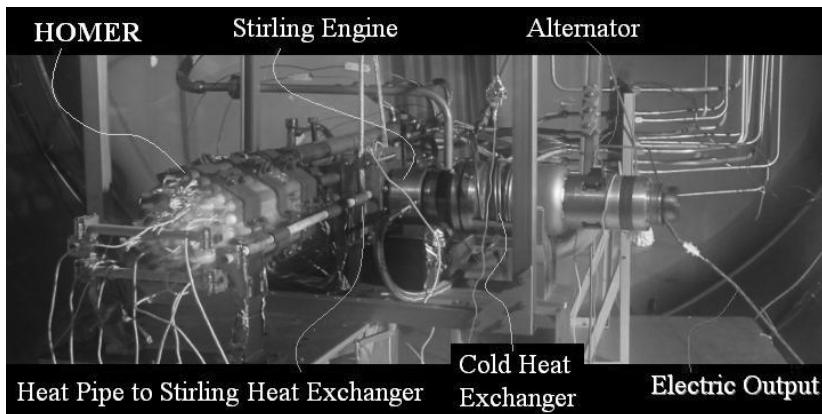


Fig. 9.4 HOMER and the Stirling Engine at test (courtesy of LANL)

9.3.2 Near Future Developments – with Available Technology

Near future developments are related to enhanced versions of HOMER meant to improve performance, safety and reliability.

9.3.2.1 The “SAFE” Reactor

The new design, shown in Fig. 9.5, is called SAFE. This reactor uses the same concept of 4:1 fuel pins to heat-pipe concept used in HOMER, but has triple heat-pipe to He-Xe gas heat exchangers, and 3 redundant Brayton engines. This structure allows a variable power production from few kWe up to 100 kWe with the same high efficiency, an optimized Be neutron reflector, and reactivity control drums made of special air-foil shaped Be₄C/Be drums to provide enhanced sensitivity in its normal domain of operation. The entire structure is made from high temperature resilient alloys, more reliable than the stainless steel used for HOMER.

The SAFE prototype reactor is shown in Fig. 9.6, courtesy of Los Alamos National Laboratory, illustrating the main features of the new design. The heat-pipes to He-Xe gas heat exchanger are being tested using the electrical heated fuel pins simulator. This makes the test easy and radiation free allowing all the components to be tested independently.

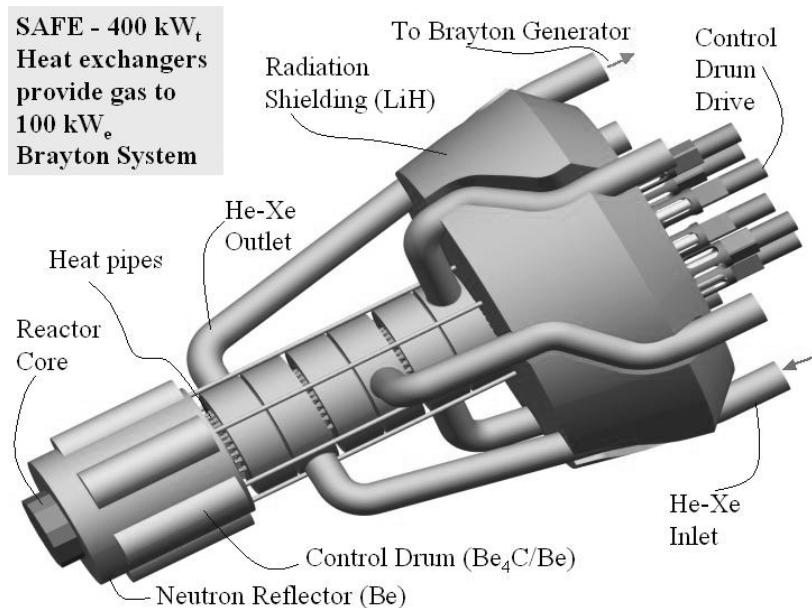


Fig. 9.5 SAFE reactor design (courtesy of LANL)



Fig. 9.6 SAFE reactor (courtesy of LANL)

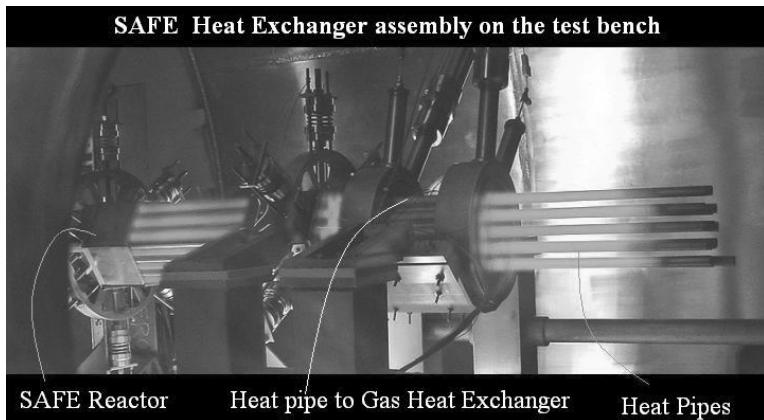


Fig. 9.7 SAFE heat exchanger test (courtesy of LANL)

Figure 9.7 shows the SAFE reactor assembly during the tests to see how the temperature distribution is varying along the heat pipe as consequence of the fact that the heat flow is extracted by the exchanger and driven to the inlet of the Brayton turbine. An enhanced Brayton cycle is designed in order to assure certified conversion efficiency greater than 30%. These reactors are ready for the next applications in the Mars program after 2010.

9.4 Future Novel Reactor Concepts for Application on Mars

A **novel nuclear reactor concept** has been proposed that is radically different than the current or projected technology in structure, nuclear fuel formulation, fuel burning efficiency, energy extraction method, fission products management, and radiation management.

This new concept might have tremendous advantage over the current state-of-the-art designs, driving to an overall lower cost, greater safety, significantly reduced waste, and high versatility (covering all the range from large to small power systems), but is not completely proven and has no test data.

The whole concept is developed around a novel approach to the control principles applied to the entities involved in nuclear energy release processes such as: neutrons, fission products, knock-on electrons, recoils, vacancies, and phonons.

The development is gradual and from macro scale to micro and nano-scale with specific scale related developments as:

- micro-hetero structured fuel “Cer-Liq-Mesh”
- nano-structure direct conversion of nuclear energy to electric energy – also called “Conductor-Insulator-conductor-insulator (“CIci”)
- nano-cluster-sinter (NCS) structure delivery of isotopic enriched materials with minimal chemical processing

- nano-structured radiation guides (NRG) for electronically controlled nuclear reflectors and ultra-light radiation shielding

These developments may be used separately or all together, forming various applications.

9.4.1 Nuclear Power Generations 1-4

In DoE's "Road Map to Generation 4" the features and evolution line was established as a line of development of terrestrial nuclear power. Unfortunately, none of these reactors is suitable for operating on Mars, mainly due to the lack of water and temperature levels under freezing that force many features out of their operating range. Using DOE's road map as a reference, new added features drive towards new generations of nuclear reactors, bringing enhanced performance and tolerance of broader environments, which makes them suitable to operate on Mars.

Nature provides models for energy to follow

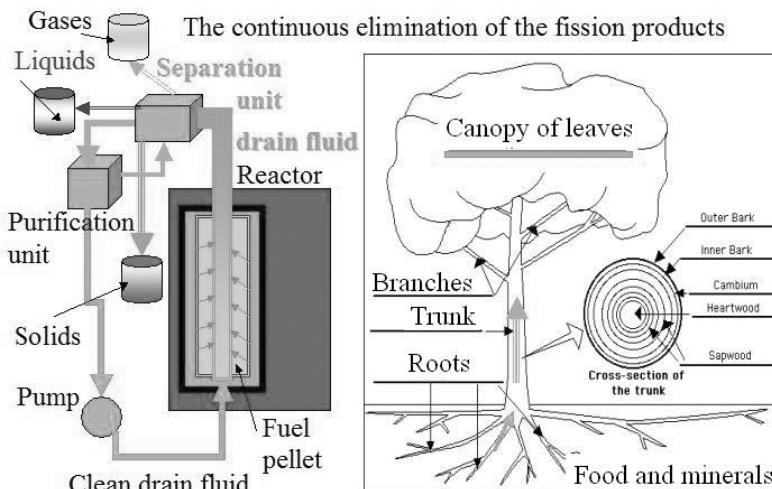


Fig. 9.8 Nature provides models for energy to follow

9.4.2 Generation 5

Is characterized by the micro hetero structure making a novel fuel called "cer-liq-mesh" that maintains the nuclear-thermo-mechanical-electric energy conversion cycle, and adds enhanced performance and features. By analogy, nuclear reactors behave like living-beings because all the time they are alive, they are eliminating the poisons and waste materials while feeding on fresh fuel, and having a low critical mass as shown in Fig. 9.8.

The Micrometric hetero-structured fuel enhances the fission products management and leads to:

- Reactors free of fission products accumulation
- Up to 20 times longer fuel life (20-30 years)
- Enhanced fuel thermal conductivity, with lower radiation damage.
- Higher possible operating temperatures ($1000 \text{ K} < T < 2500 \text{ K}$) for higher efficiencies ($50\% < \text{Eff} < 85\%$)
- New waste cycles with less waste (>100 times reduction)
- Higher burnup (up to $>700 \text{ GWday/ton of fuel}$)

The development of such a reactor structure requires a deeper understanding of fission and fusion energy transport along with materials behavior in radiation dominated structures. Having high efficiency is possible by being cooled in Mars' atmosphere, as the excess thermal energy is about $\frac{1}{4}$ of the total energy.

9.4.2.1 Micro-structured Fuel Mesh or “cer-liq-mesh”

It is a generic name that means replacing solid ceramic fuel with a new type of hetero-structured fuel.

To minimize damage in the fuel lattice by fission products by nuclear interaction, it is necessary to make micro-beads shorter than the fission products range (about 15 microns in urania), known since 1956 as the fuel dispersion theory. The beads are immersed into a liquid metal, which stops and collects the fission products without being damaged, instead of in a metal as in actual cer-met fuels. The new structure is generically called “cer-liq-mesh”. The liquid metal stopper flows

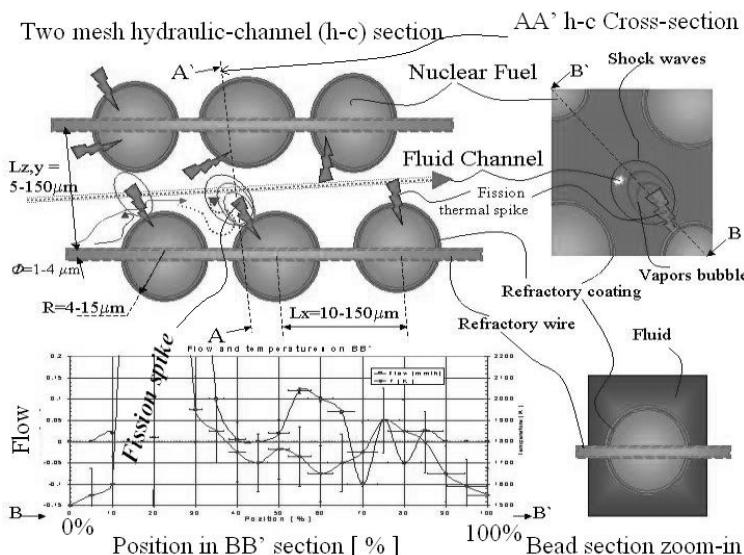


Fig. 9.9 Cer-liq mesh fuel details

smoothly and continuously eliminates fission products from the active region in order to avoid the “poisoning” effect that impedes the burning process and also reduces the fission products isotopic distribution simplifying the waste cycle as shown in Fig. 9.9.

This type of fuel has several times higher thermal conductivity, which allows a higher operation temperature, associated with higher conversion efficiencies up to 70%. It also allows higher power density, reducing the reactor linear dimensions by a factor of 2-3 with associated weight reduction by 3-10 times.

Brief on Micro-structured fuel performances

Several micro-structure fabrication methods are possible of being used (Popa-Simil 2007b). On terrestrial applications, the most preferred drain liquid is based on LBE because it generates a hard neutron spectrum (Little 1999; Romano 2004), while the fuel buoyancy is almost zero, but for space applications it increases the payload. In LBE, fission products and their combinations will have positive buoyancy traveling against the pressure gradient field. For space applications, NaK or liquid Li metal are preferred, but their corrosion resistance has to be considered (Saito 1999). The structure will be lighter and in gravity the fission products will have negative buoyancy (being heavier than the liquid) and will sink while the gaseous fractions rise in the opposite direction. The fuel is primarily made of beads with a random spheroid shape, 10-20 microns large, sitting on a 1 – 3 micron thick wire mesh as shown in Fig. 9.10. The mesh layers may be spaced by 20-80 microns to each other. This sheet of nuclear fuel beaded mesh may be processed similar to a textile fabric, being grouped in packs looking similar to a felt, and using its elastic properties may create mechanical structures.

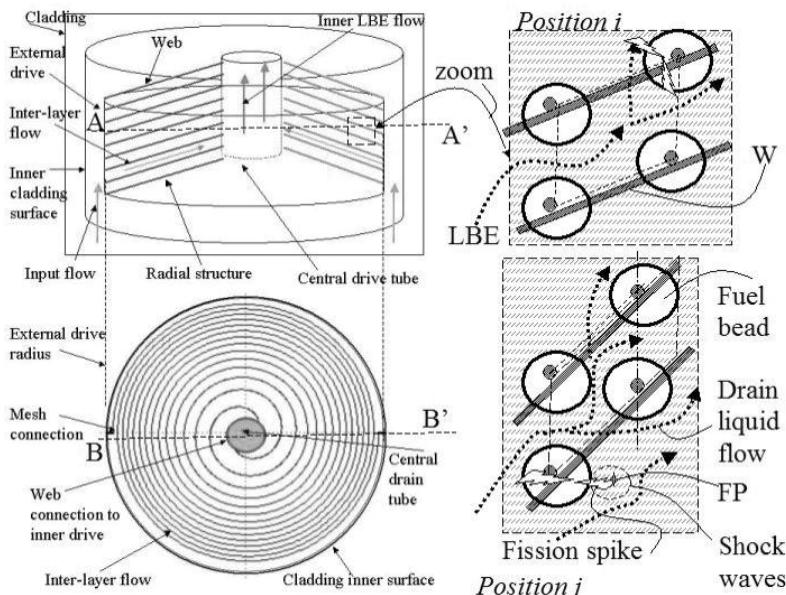


Fig. 9.10 Potential fuel pellet structures

Carbon nano-tubes as an alternative to Tungsten or metal wires is also considered together with a coating layer to enhance the micro-flow performance by adjusting the liquid-structure interface adhesion forces (Ajayan 2005). In the top left of Fig. 9.10, a longitudinal cross-section through an example pellet is presented. The mesh creates conical shapes supported on a radial structure that maintains and adjusts the interspacing between the layers. In the larger tube, the radial supports are near perpendicular to the cladding, maximizing the inter-bead gap as shown in Fig. 9.11. When the diameter of the tube decreases, the radial structure tilts and accommodates the fuel bead fabric layers inside eliminating part of the drain fluid. Under the longitudinal section through the pellet is cross section BB' showing the radial arrangement. The right side of Fig. 9.10 shows a section zoom in a fuel voxel and at the AA' level, which shows the fuel compaction from the upper to the lower zoom corresponding to the lower, larger cladding diameter respectively. The beads on adjacent layers in the left side voxel magnification form a hydraulic channel where the liquid gently flows and drains the fission products. The fission process is accompanied by a thermal spike, that makes a vapor bubble similar to a micro-cavitations process, which generates compression shock waves. Position j in the image shows a section in a smaller diameter of the reactor tube, where a hydraulic channel shrinks the beads towards each other, reducing the amount of drain liquid.

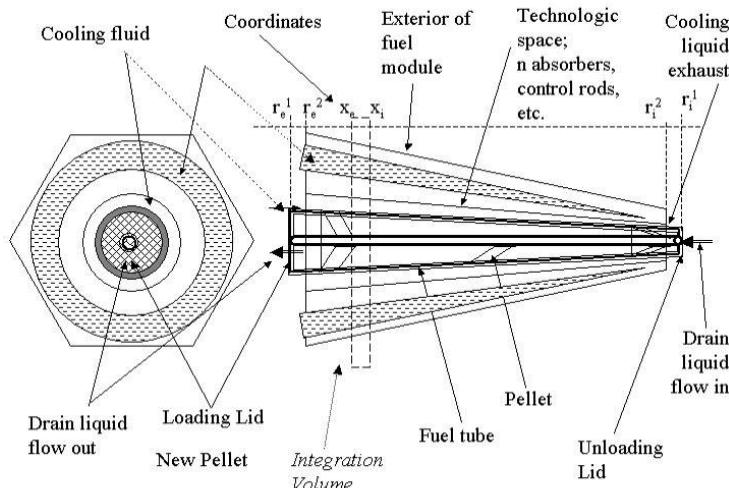


Fig. 9.11 Fuel tube conic shaped and pellet compression

Compressible fuel structure

The meshes provide elastic properties, which can be used to create elastic compressible fuel structures presented in Fig. 9.11. By compressing the spent fuel after a burnup period, the reactivity is kept constant by eliminating drain fluid among the beads and increasing the fuel-drain-liquid dilution ratio. The pellet is made of disk-shaped, fuel-beaded webs connected at the central drive tube to an external

drive tube or other structure(Popa-Simil 2005). In fact, as the fuel works in the reactor, the fission products leave the bead structure stopping into the liquid metal. In this way, the fuel beads are losing mass, and the reactivity of the assembly decreases due to the fuel consumption and because of the fission products nearby. Fuel reactivity adjustment by compression represents a coarse reactivity adjustment. Control rods or reflectors will perform the fine reactivity control. Fission product extraction made by the drain liquid drives the need for large overcriticalities compensated by absorption rods or by poisoned fuel.

Simulations show that fuel swelling and cracking is smaller due to the low recoil rate and the effect of the liquid's higher thermal conductivity, which shields the fuel bead and provides for a flat thermal field distribution. Due to the fact of fuel compressibility, the reactivity varies up to ten times and pushes the burnup factor up to 90% of the initial mass. It also has the potential to increase the fuel lifetime up to 10-15 years.

Cer-liq-mesh is a Nuclear Fuel with increased thermal conductivity.

Due to the presence of the liquid metal inside the ceramic fuel structure, and of the metallic meshes, the average thermal conductivity of the pellet is improved. The discharge of the fission power, mainly into the high conductivity material, helps to improve the thermal field and heat flow.

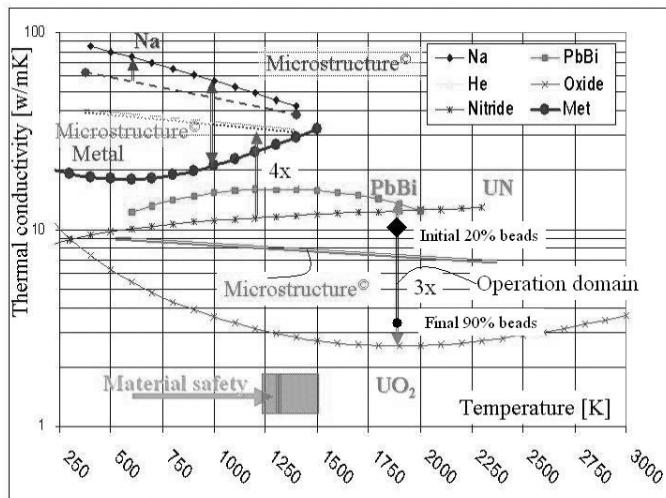


Fig. 9.12 Thermal conductivity versus temperature

Using experimental data fit formulae(Wallenius 2004) for thermal conductivity picked for various nuclear reactors fuels (Ronchi 1999; Inoue 2000; Degueldre 2003; Fink 2004) and cooling agents, the improvement brought to thermal conductivity by sinking the fuel micro-beads into the cooling agent was estimated as function of temperature and liquid-fuel ratio as given by the compression factor.

As shown in Fig. 9.12, the main trend is to increase the thermal conductivity and to reduce its variation with temperature. As an example, for Urania-LBE the thermal conductivity increases 3-5 times while its variation with the temperature reduces up to 5 times and decreases the variation with temperature. Better results are obtained if NaKE or Li are used as the fuel becomes lighter, but the boiling point and bead's material related sensitive phase transformation points also limit the maximum operation temperature (Douglas 2004; Vermoyal 2004).

For very high temperature applications over 1000 C, some other metals or molten salts may be used as drain and cooling agents, like Zr, Ag, etc (Bakker 2004; Romano 2004). Special attention is given to cer-met-liquid metal structures able to operate at ultra high temperatures based on carbides like PuC, UN, UC, WTiC₂, Zr, Ag, etc. These structures may operate at ultra high temperatures, which would allow very high-energy conversion efficiency up to 80%. These advances in mobile power sources might trigger the development of flexible power grids with higher reliability to weather adverse actions or accidents and make them more suitable for Mars.

Nuclear Poisons Free Fuel

The release of fission products and their impact on the high burnup rate, is improved in the new structure where the liquid carries away the fission product (Lanning 1997). In systems where the fluid drains the fission products the flow rate is chosen small enough to allow the unstable short-lived fission product to continue the decay chain inside the reactor's hot zone shielding, but high enough to drive the high poisons outside of the hot zone as fast as possible. The ¹³⁵Xe case used as example in Fig. 9.13 is obtained by the beta decay of ¹³⁵Te as the primary

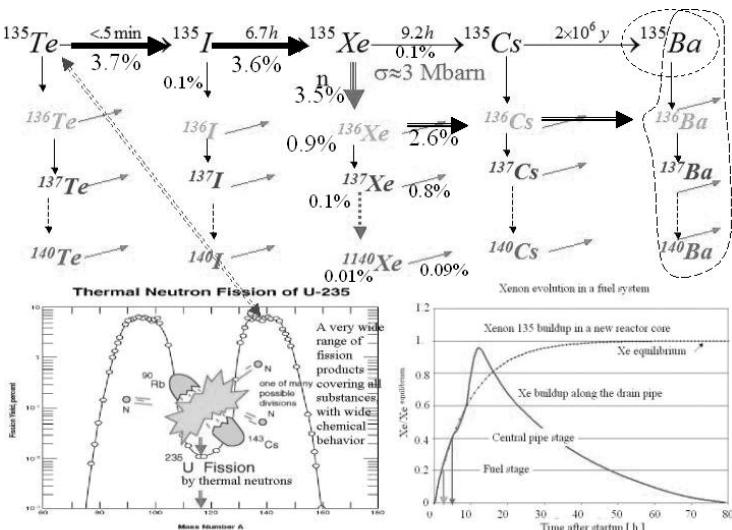


Fig. 9.13 Complexity of the fission products transmutation – the ¹³⁵Xe case

fission product. If the ^{135}Te can be immediately removed from the reactor zone, only the first line of the 135- isomer decay will operate generating a ^{135}Xe peak that finally ends in the stable ^{135}Ba as presented in the chart nearby.

Because fission products are not removed from the neutron flux present in the reactor core, all the fission products are exposed to n-capture nuclear reactions that produce a large variety of isotopes, all having their specific decay chains lifetimes. Another disadvantage of the fission products is the wide span of their chemical reactivity, covering almost all the possibilities. As an example, ^{135}Te is electro-negative and ends in ^{135}Ba , which is electro-positive. Inside the fuel, these nuclear disintegrations and transmutation embitter the structure by the sudden changes in chemical bounds accompanied by energy release and subsequent defect formation. The drain fuel is practically insensitive both to fission product collision damage and to its further decays. In the case of the ^{135}Xe buildup presented in the chart of Fig. 9.13, the time form the pickup of the Xe atom until it is out of the neutron flux area and the Xe buildup starts is about 6 hours. During this time, the Xe follows the normal curve of solid fuel buildup because of its very large n absorption cross-section that forces the majority of the nuclear transmutation channels to go mainly through ^{136}Xe . Being a gas, it separates itself easily from the rest of the fission products. Therefore, the Xe separation will contain at least two Xe isotopes, with the ratio dependent on the drain liquid flow rate details. The Xe peak shown in the chart occurs in the isotopic separator and not inside the nuclear reactor. This is one of the mechanisms that save neutrons, which may be redirected for fuel breeding or high burnup and desired transmutation. Nuclear waste management is enhanced due to the fact that the fuel bead has dimensions smaller than the fission products range, and most of them get collected into the drain liquid, which slowly carries them outside the reactor hot zone into a separator zone. Restart capability is improved, in spite of the fact that ^{135}Xe buildup in fast spectrum reactors is less important than in thermal reactors. The already partitioned and stabilized fission products out of the reactor hot-zone present less of a hazard than inside, while the neutron economy is better.

Compact nuclear structures

Figure 9.14 shows the amount of fissile material and the smallest dimensions of various isotopic nuclear reactors calculated for spherical reactors (Loaiza 2004). The only purpose of the connection lines is to help link related data.

The figure shows that the entire nuclear reactor does not contain enough fissile material to produce an explosive device, and due to the combination as XW, XWC, XC, (where X stands for fissile material as ^{235}U , ^{239}Pu , ^{241}Am) the extraction is difficult and produces a low yield. The used fuel contains almost no fissile material left due to ultrahigh burnup and no breeding product due to an isotopic enriched structure.

Near-Perfect Burning Nuclear reactors

Continuous fission and breeding products handling makes the residual radioactivity small, and the potential danger form an accident is minimized. If the elastic

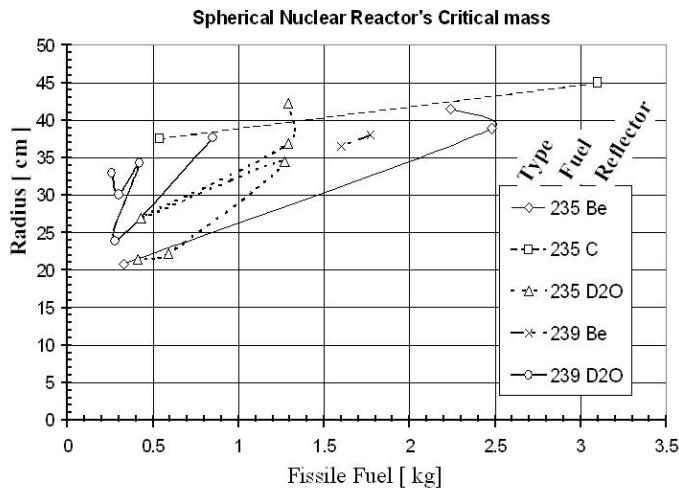


Fig. 9.14 The criticality

properties of the fuel are exploited together with the fission products drain capabilities, this technology drives us towards a new reactor core design giving near perfect burning (Popa-Simil 2005a, 2006b, 2007b).

If only the fission products drain capabilities are used, the fuel may offer longer life and easy reprocessing (Popa-Simil 2005b). There is an important conceptual difference between the treatment of the fission products as done inside the Triso fuel (present in KSR) and this concept. Present thinking is that it is safer to put more coatings on as confinement layers to keep the fission products in place inside the fuel. That makes the fission products “powerful” and in short time is able to disqualify the fuel. The new concept, is an application of a natural principle that “the space takes over the power”, eliminates all confinement layers induced inner barriers inside the cladding leaving the fission products to reach the cladding, or to flow through the cladding outside the nuclear reactor into a waste treatment system. The advantage of the lack of internal confinement barriers is compensated by the low osmotic pressure of the fission products. Outside the hot zone special getter coatings and other chemicals are stabilizing and containing the fission products at a lower temperature decreasing their hazard by orders of magnitude.

This structure may assure an increase in fuel burnup providing a less hazardous nuclear reactor structure even for the most severe Chernobyl type accidents (Popa-Simil and Muntele 2007).

On Mars this structure might be used in an enhanced SAFE reactor design where the pin fuel is replaced by a specialized encapsulated mesh fuel where the fission products are stored at the bottom of the reactor in a specially designed trap. The drain liquid will be Na, or NaK that has the effect of increasing the thermal conductivity by 4-6 times with improved reactor longevity. Without increasing the fissile mass, the extraction of fission products may increase the fuel life by several times. Fuel recovery may be simple, with the possibility of being performed on

Mars, by opening and removing the fission products while compressing the fuel and re-using the same fissile material.

9.4.2.2 Pulsed Thermoelectric Machine

An interesting alternative conversion using a new type of thermoelectric transport is described, consisting of pulses of charge carriers which "fly" periodically through an external circuit from the hot end of the sample to the cold end, with a

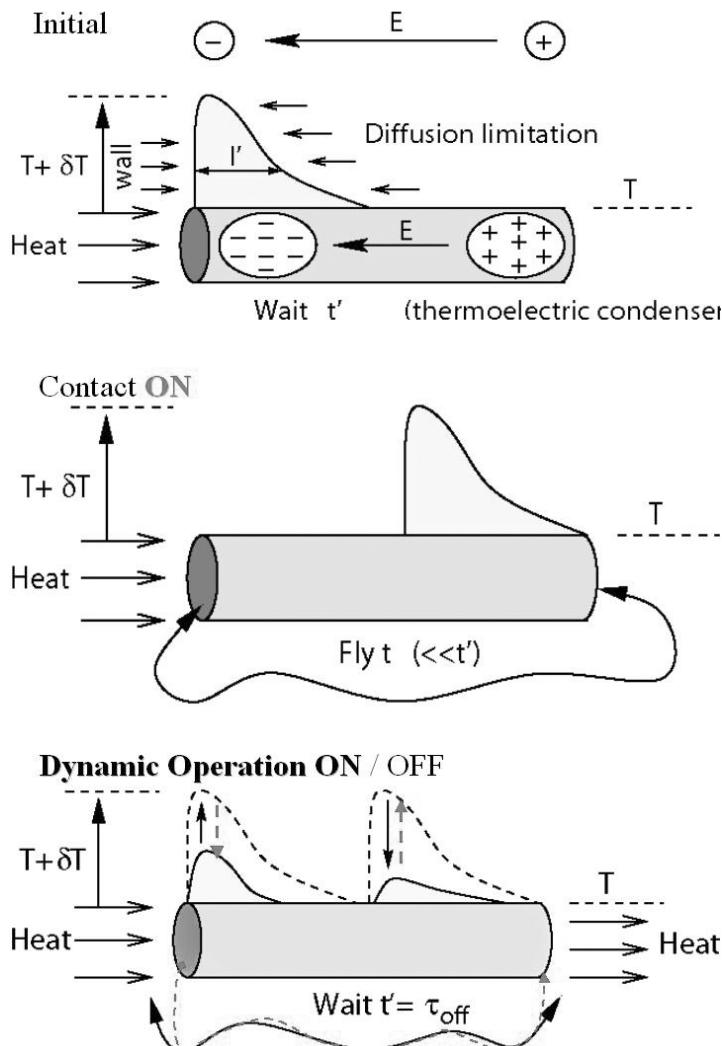


Fig. 9.15 Operation principle

predetermined duty cycle of the electric connections, while continuously maintaining the thermal connection (Apostol 2007a). It is shown that such a non-equilibrium ideal thermal "machine" works cyclically with the same efficiency quotient as the ideal efficiency quotient of the thermoelectric devices operated in an equilibrium transport regime, but the electric flow and power are increased, as a consequence of concentrating the charge carriers on pulses of small spatial extent. The machine is reversible, in the sense that it can operate either as a thermoelectric generator or as an electro-thermal cooler.

The machine operates in two strokes: in the first, the wire is connected (the electric circuit being ON) and in the second, the wire is OFF. Figure 9.15 shows the important features of this pulsed thermal-to-electricity converter. In the sketch it is a thermo-electric condenser warming up from a heat source taking a temperature δT , that produces an electron excitation, and more electrons are passing in excited states charging the capacitor negative. In a short thermal interval the heat energy is completely in equilibrium with the entry stage, propagation by diffusion, or thermal conductivity towards the cold source in the right. After an optimal time interval, the two ends of the system are put in electric contact through the wire and the electric power is transferred to the cold end of the heat capacitor. The charges pass through the wire and charge the other capacitor. During the contact time the polarization charge from the hot end decreases the electric charge being transferred to the cold end, increasing its temperature and potential. The heat flow converted in excitation electrons at the hot end is transferred as electricity to the cold end and the device remains positively charged at the hot end. The load returns the electrons back to the hot end to complete the cycle.

The theoretical calculations (Apostol 2003) made for this new type of thermoelectric transport (Apostol 2001) show very high power density parameters even when the conversion efficiency is obtained by:

$$\eta = \frac{\eta_C}{\eta_C + 1 + \frac{K}{Q^2 \sigma T}} = \frac{\eta_C}{\eta_C + \frac{1}{ZT}}$$

where $\eta_C = \delta T / T$ is the Carnot efficiency, and ZT is the "figure of merit" (Apostol 2008). The power frequency and circuit dependency was theoretically formulated as:

$$\bar{P} = P_{eq} \left(\frac{l}{l'} \right) \frac{\tau_{off}}{\tau_{ON} + \tau_{off}} = \frac{l}{\sqrt{vA}} \sqrt{f(1 - \tau_{ON})} P_{eq}$$

where τ_{ON} is the time the circuit is connected conducting the charges towards the cold end, P_{eq} is the electric power produced by the pulse-operated thermoelements which is higher, by the same factor 10:1, than the electric power of the same

thermo-electric elements operated at equilibrium. It is easy to see that the equilibrium power $P_{eq} = j_{eq}^2/\sigma l$ (per unit area of the cross-section) can also be written as $P_{eq} = \sigma U^2/l$, or $P_{eq} = U^2/r$, where r is the internal electric resistance. An overview of the numeric calculations is given in Table 9.10.

Table 9.10 The predicted values (Apostol 2007b)

Description	Metals	Semiconductors
electron density	$n = 10^{22} \text{ cm}^{-3}$	$n = 10^{17} \text{ cm}^{-3}$
chemical potential	$\mu = 1 \text{ eV}$	$\mu/T = -5.7 \left(\frac{\mu}{T} \right) = \ln \left[n \left(\frac{h^2}{2\pi n T} \right)^{3/2} / 2 \right]$
Fermi velocity	$v = 10^5 \text{ m/s}$	$v = 10^4 \text{ m/s}$
room temperature	$T = 300 \text{ K}$	300 K
Hot end temperature	$\delta T = 100 \text{ K}$	
mean free path	$\Lambda = 100 \text{ nm}$	$\Lambda = 100 \text{ nm}$
δ -pulse electric flow	$j = 10^5 \text{ A/mm}^2$	$j = 400 \text{ A/mm}^2$
δ -pulse	$\delta n = \left(\frac{\pi^2}{4} \right) \left(n T \delta T / \mu^2 \right)$	$\delta n = n \left(\frac{3}{2} - \frac{\mu}{T} \right) \delta T / T$
Equilibrium flow	$j_{eq} = 1000 \text{ A/cm}^2$	$j_{eq} = 0.4 \text{ A/cm}^2$
sample length	$l = 1 \text{ mm}$	$l = 1 \text{ mm}$
Voltage	$U = 0.4 \text{ mV}$	$U = 60 \text{ mV}$
the thermo power coefficient	$Q = -\left(\frac{\pi^2}{6e} \right) \left(\frac{T}{\mu} \right) \approx 4 \left[\frac{\mu V}{K} \right]$	$Q = 600 \text{ } \mu\text{V/K}$
electric power of the pulse	$P = 4 \text{ kW/cm}^2$	$P = 2.4 \text{ kW/cm}^2$
the equilibrium power	$P_{eq} = 400 \text{ mW/cm}^2$	$P_{eq} = 24 \text{ mW/cm}^2$
optimal pulse length	$l' = l/10$	$l' = l/30$
the current	$j = 100 \text{ A/mm}^2$	$j = 12 \text{ A/cm}^2$
the power	$P = 4 \text{ W/cm}^2$	$P = 0.72 \text{ W/cm}^2$
the maximal power	$P_{max} = 2 \text{ W/cm}^2$	$P_{max} = 0.36 \text{ W/cm}^2$
increase factor	5	15
operating frequency	$f = 500 \text{ kHz}$	$f = 50 \text{ kHz}$
Time ON /electric length	$(\tau_{on} = 1 \text{ } \mu\text{s} \text{ and } l_e = 10 \text{ cm})$	$\tau_{on} = 10 \text{ } \mu\text{s} \text{ (for } l_e = 10 \text{ cm)}$

The operating frequency is determined, as well as the on- and off- times, as functions of the sample type, the extension of the pulses and the length of the external circuit. The electric flow and power are higher for pulse-like transport than for equilibrium transport, by a factor of 5-15, as a consequence of concentrating the charge carriers on pulses of small spatial extent. Such a pulse-like operating thermoelectric device works cyclically, with higher efficiency than the equilibrium operating thermo-elements.

9.4.3 Generation 6

The thermo-mechanical intermediary conversion cycle shown in Fig. 9.16 is possible of being removed by a solid state-compact nuclear reactor battery that loads directly from nuclear fission energy and discharges it as electricity shown in Fig. 9.17.

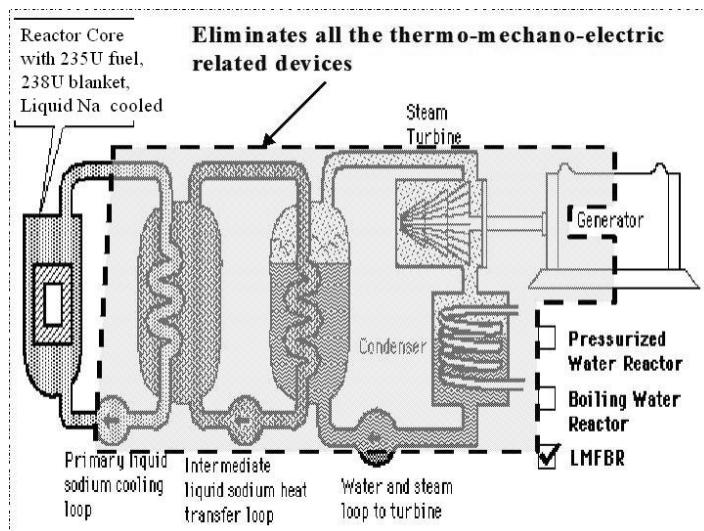


Fig. 9.16 The present reactor structural change

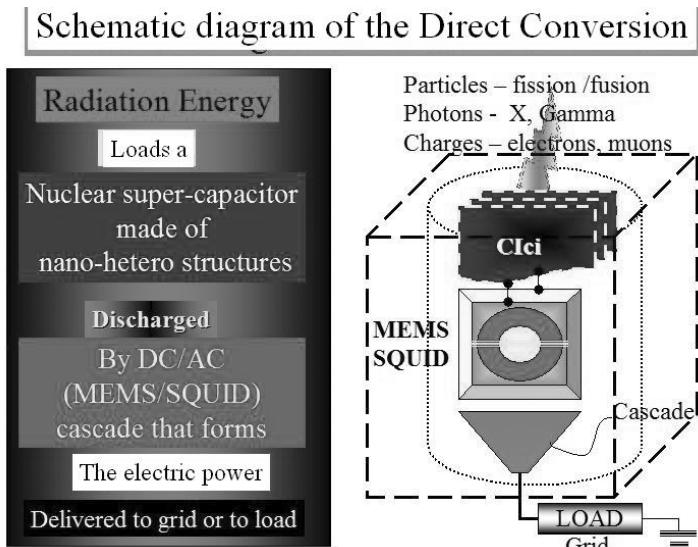


Fig. 9.17 Generation 6 operation principle

The development of the nano-hetero-structures handling the knock-on electron showers may drive to:

- Direct nuclear energy conversion into electricity with applications in fission, fusion, hybrid structures, active shields and remote power antenna for space shuttles – without using thermal turbines and heat exchangers with efficiency reaching over 80%.
- Cryogenic nuclear reactors, for high efficiency conversion (efficiency > 98%), Self-recovering structures from radiation damage, Mobile, nuclear high-power (<1GW) sources (battery like devices) and nuclear compact isotopic batteries using fission products and actinides.

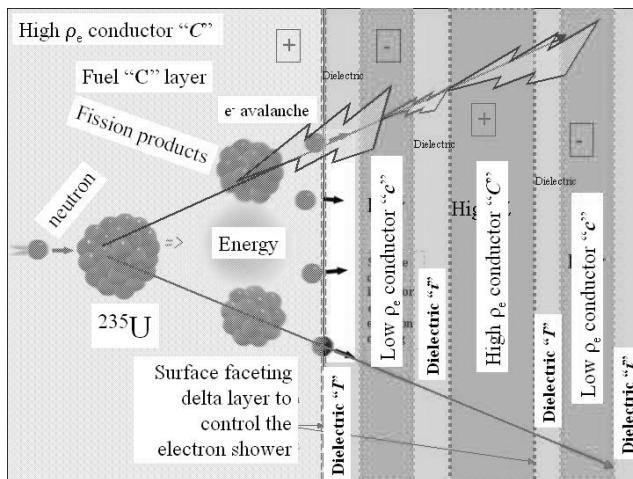


Fig. 9.18 Nano-layer structure

9.4.3.1 Direct Conversion from Nuclear Energy-to-Electric Energy

The direct nuclear-to-electric energy conversion a concept is based on the controlled collection of knock-on electrons resulting from the movement of the nuclear particles inside a nano-structured lattice, shown in Fig. 9.18. Moreover, by having the energy extracted directly as electricity, there will be minimal to zero energy left for heating the reactor. Such a reactor will run cold, even cryogenic.

9.4.3.2 History of Direct Conversion

The interest in direct conversion appeared in early 1940s. Table 9.11 lists the direct energy conversion development.

Table 9.11 Direct energy conversion development

No.	Inventor	Description
1	Kallmann (1940)	measuring the intensity of radiation of slow neutrons by means of an ionization chamber Ra-n- ¹⁰ Be
2	Linder (1946)	thermo-ionic fission, followed in 2000 by (Polansky 1999; Brown 2000)
3	Broxon (1948)	differential ion chamber
4	Rappaport (1953)	nuclear batteries
5	Willson (1955)	beta ray collecting
6	Schwarz (1958)	beta-voltaic batteries utilizing fission products as well as neutron activated isotopes
7	Schuylar (1958)	Semiconductor metal junction
8	Hoff Jr. (1970)	junction electron-voltaic semiconductor
9	Adler (1976)	
10	Ato Yasuro (1976)	pn magnetic field to the converter
11	Schenectady (1977)	deep diode atomic battery made from a bulk semiconductor crystal containing three-dimensional arrays of columnar and lamellar P-N junctions
12	Van Dine (1987)	selective region of semiconductor layer
13	Tsang (2004)	Liquid Electronics Advanced Power Systems (LEAPS)-GTI financed by United States Defense Advanced Research Projects Agency (DARPA)
14	Ritter (1979)	Gamma photoelectric effect
15	Young (1997; 1999)	charged particle powered battery
16	Lal (2002; 2005)	Micro-electromechanical Systems (MEMS) piezoelectric converter
17	Shanks (1998)	liquid source solid interface
18	Snyder (2004)	junction-type battery with Gallium liquid source
19	Gadeken (2004)	n-type pores with tritium gas
20	Winston (2004)	piezoelectric switch element. with tritium light source
21	Blanchard (2005)	miniature devices
22	Shimabukuro (2005)	solid state micro-power electrically insulating substrate; a semiconductor layer on the substrate
23	Yarygin (2005)	silicon-germanium thermoelectric converters with the radial-ring geometry at 1700 K
24	Anghiae (2002)	Magneto-Hydro-Dynamic vapor core reactor

These developments show the growing interest to develop nuclear power and unleash its exceptional capabilities over a wide range of applications.

9.4.3.3 Brief on Micro-structured Fuel Performance

In the future, *nuclear generator power density could be increased up to several GW/liter*. Nuclear reactors developed using micro-nano-structured fuel will look like a solid-state nuclear battery, with no turbines and no heat exchangers.

This concept relies on the relationship between the fission products and the electrons and is a byproduct of the answer to the question: *How does fuel heat up?*

This also brings up a secondary question: *How extract that energy without heating the fuel?* These theoretical developments appeared in the late 80s stimulated by a failed experiment, where several tenths of micro-amps were generated by depositing a few micro-amp beam into a micro-layer structure (Popa-Simil 1986).

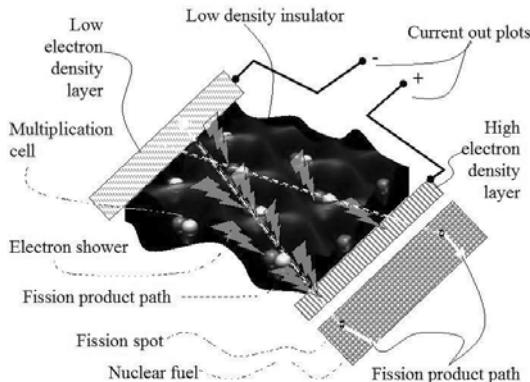


Fig. 9.19 Nano-Cluster structure

A potential explanation to that experiment was found after 1990 using RBS experiments and composite micro-foil sandwiches.

The nano-structure is made from many layers of conductor-insulator as showed in Fig. 9.19 a nano-clustered version of the structure in Fig. 9.18. The layers have the advantage of removing the energy produced as electricity by a few orders of magnitude faster than by heat flow, which would provide exceptionally high power densities, up to several GW/liter. Present day power densities are $< \frac{1}{2}$ MW/liter.

It is also possible to generate electricity more efficiently by using superconducting structures (Popa-Simil 2006a; 2007a; 2007d; 2007f). These structures are the most appropriate for space applications, producing the right amount of power for solar system manned exploration, and unmanned missions all the way to Alpha Centauri (Popa-Simil 2005c).

The advantages of these systems are that they allow a fast response to power production which, in turn, allows harvesting of other available energy such as solar and wind (Popa-Simil 2007c).

For a normal material shown in Fig. 9.20 at the breakdown limit and at maximum resistive losses in normal conductors, the power density is about 5 KVA/mm³, which equals 5 GW/liter. At this power density, a structure with 25% fissile material embedded may run less than 1 day and completely consume all the fuel.

Cryogenic structures are anticipated to go at least 10 times higher in power density, thereby opening the way to a new type of miniaturized pulsed power. Mixed solutions of micro-nano structured fuel can be used for synergistic performance gains.

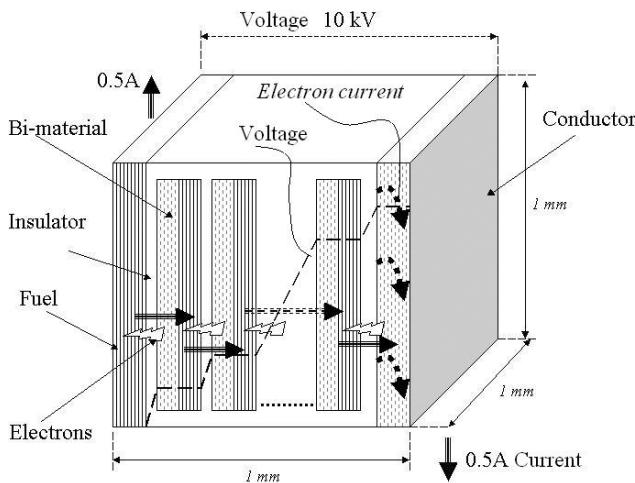


Fig. 9.20 Limits of power density in a planar nano-structure

9.4.3.4 The Challenges of Nuclear Batteries

The direct conversion process relies on harvesting the kinetic energy of nuclear particles crossing a sequence of materials with different ionization stopping power, which generates a differential voltage at the circuit terminals. The current is created by the accumulation of knock-on-electron generated showers that converts the high energy of the radiation represented by a single particle, high-energy alpha or fission product, shares the energy with many particles, and forms the electron avalanche.

There are mainly three versions of structures that enable the high efficiency, direct-conversion, process:

- The planar conductor-insulator (Clci) structure is made by using a number of “Clci” modules encompassing the radiation stopping range in all directions around the radiation source.
- The serial “plasmon-nano-cluster” structure, replaces the sequence of “Clci” modules with a specific conductive nano-clustered insulator.
- The nano-tube-based ultra-capacitor structure represents a morph of the initial “Clci” structure to accommodate the nanotube enhancement. Nanotubes have a special structure, which assures high efficiency energy harvesting.

9.4.3.5 Nuclear Battery Power Density

The battery design depends on the radioisotope type and material used, being a tradeoff between lifetime and power density (see Fig. 9.21).

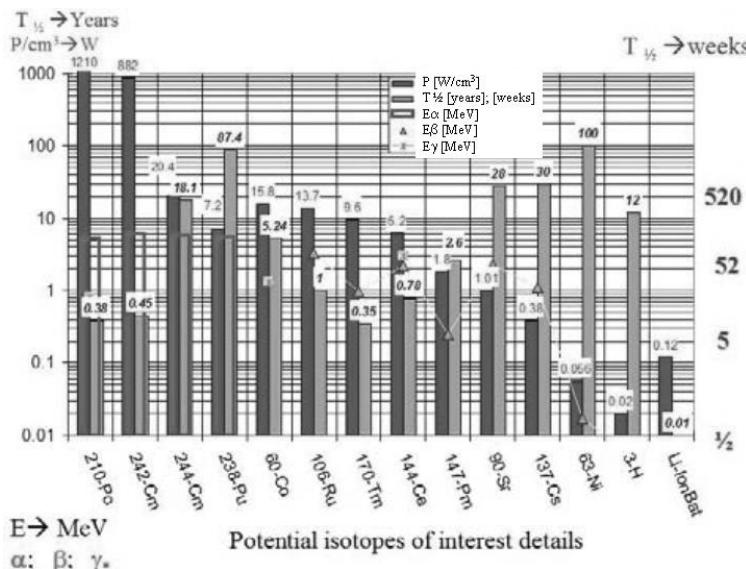


Fig. 9.21 Some of the potential isotopes

For ^{210}Po batteries, the maximum power density may be as high as 250 W/cm^3 , while after 1 year only 15% of the power remains. For ^{238}Pu , the power density may reach up to 5 W/cm^3 , and after 90 years it will have 48% of the initial power. There are about 30 other radioactive alpha emitter radioisotopes that may be considered along with 50 beta emitters. Gamma rays may also be used, but their power density is small. Collateral radiation of these isotopic power sources is typically small and depends on the radioisotope used, and must be considered depending on the application.

9.4.3.6 Nano-Cluster-Sintered Fuel (NCS) for Super-Grade (^{233}U , ^{239}Pu) Production

A new breeding or transmutation procedure leading to super-grade materials has been proposed making use of nano-structures in conjunction with recoiled nuclei, neutron capture, and elastic scattering. The result is a nano-structured system that can deliver isotopic enriched materials with minimal chemical post-processing. These systems may bring advantages in extraterrestrial environments such as Mars where the typical procedures may not easily fit. This process uses nano-grains and the recoil of the compound nucleus to place the newly formed isotope into a suspension-liquid, thus preventing multiple neutron capture and creating super-grade materials.

This process also has the advantage of producing high purity isotopes for medical and technical applications and for nuclear fuel breeding. In addition it separates the processes inside the nuclear reactor and does not perturb the criticality balance or process balance, which makes this process attractive in spite of the more complicated pellet structure as shown in Fig 9.22.

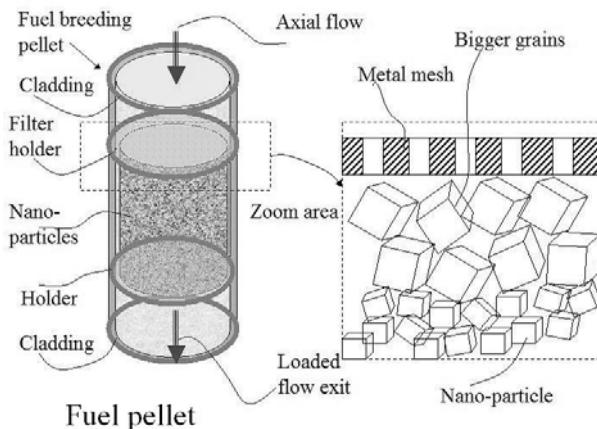


Fig. 9.22 NCS fuel pellet

The NCS process converts ^{238}U , representing 99.3% of uranium ore, and ^{232}Th into the fissile materials ^{239}Pu , and ^{233}U respectively. In addition, the NCS structure may be used for burning poisons or for other material transmutation inside the nuclear reactor.

On Mars the energy demands are small, but the operation conditions are very challenging; therefore, whatever the price of nuclear fuels and their abundance will be on Earth, there will be enough to supply the initial outpost. In the future will be possible to develop mining and fuel processing on Mars making the nuclear power self sustained. The first most prolific resource is Thorium, and that drives to ^{232}Th - ^{233}U cycle.

The thorium and uranium contents of volcanic and ancient terra formations on Mars have been evaluated on the basis of the Mars-5 space probe measurements of gamma radiation from the Martian surface (Bazilevskii 1981). Gamma ray data for two partially overlapping territories containing different proportions of the major Martian geological formations were analyzed under the assumptions that the data represent composition differences between volcanic and terra formations, or between native and weathered rock. In the first case, it was found that in volcanic Martian rocks the thorium concentration is 5 ± 2.5 ppm and the uranium content is 1.1 ± 0.8 ppm and are similar to those of many typical terrestrial and lunar basalts. On Terra formations the thorium concentration is 0.7 ± 0.35 ppm and the uranium concentration is 0.2 ± 0.14 ppm, resembling the anorthosite-norite-troctolite series of the lunar highlands and not terrestrial continental granites. Another important impediment is that the actinide mining procedures developed on Earth are incompatible with the conditions on Mars, where specialized mining procedures have to be developed from scratch.

9.4.4 Generation 7

The incorporation of the electronically controlled reflectors minimizes the critical mass and shielding dimensions making possible the development of ultra-small nuclear power sources. These developments are only in the idea stage presently, being hot issues of active research, and theoretical extrapolations (Avakian 2004) from research on radiation channeling (Kumarov 2004) with very promising possibilities.

As we showed in Fig. 9.14 the nuclear reactors may be made very small, about $\frac{1}{2}$ m in diameter with weights less than 1 ton and powers in accordance with construction materials safety limits. However, the shielding remains the biggest obstacle in their potential applications.

If the nano-guides for nuclear radiation (mainly neutrons and gamma) will work successfully, the nuclear reactor shield size might be reduced from 2 m down to 5-10 cm thus reducing the dimensions of a typical nuclear reactor from about 10 m to less than 1 m for about 1GW of power(Popa-Simil 2007e; 2008).

Figure 9.23 shows the internal structure of a ideal nuclear reactor, where the shielding and the control is made electronic by using electro-reactive channeling structures.

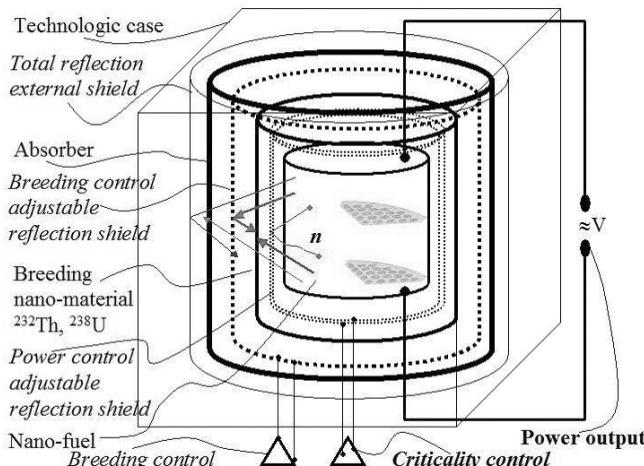


Fig. 9.23 New reactor structure using nano-structured radiation guides (NRG)

The reactor contains two electronically controllable albedo active shields that regulate the criticality and breeding rate along with the operating power. Escaped neutrons are transmitted to an absorber material for transmutation purposes and are totally reflected inside the reactor. In this situation the criticality mass becomes very small – estimated at 50 g of ^{239}Pu or ^{233}U for several MW of power in about a 30 cm radius.

It is known that to have 1MW-day power, it is necessary to burn out about 1 g of pure fuel ($^{233}, ^{235}\text{U}$, ^{239}Pu , etc). The fuel has to be added daily or “exposed” by

withdrawing the absorption rods or control drums to the nuclear reactor. Since the critical mass does not burn, it has to be maintained to assure the continuity of the nuclear fission process, and this increases with the burnup.

Generation 7 opens the way to power autonomous transport vehicles for many years with a single load (1,000,000 miles for a Hummer SUV equivalent will require about 3-10 g of fissile material) or its extraterrestrial Rover version with significantly increased performance making Mars more accessible by car.

The usage of the compact nuclear power structure on flying vehicles by electric jet propulsion opens a new means of transportation, allowing a complete tour around the Mars planet for less than 10 g of fuel in less than 2 days. The propellant may be provided by compressing the rarefied Mars atmosphere in tanks and accelerated in a plasma accelerator jet to achieve the necessary thrust. Electro-turbo-fan mode may be a high-speed low altitude-flying mode that may be also possible.

There are many other applications of electro-turbo-fan nuclear powered systems, such as: stratospheric aerodromes, magnetic levitation rolling super highways, space power and propulsion, etc., all based on the success of electrical controlled nano-shielding. All of these “generation 7” devices are in the very incipient stage of the theoretical concept development phase.

9.4.4.1 Electronically Controlled Nuclear Reflectors for Ultra-Light Radiation Shielding

Figure 9.24 shows the operating principle of a complex nano-structured composite material proposed for electrically controlled shielding, creating a new type of nuclear reflector with more adjustable albedo/transmission ratio and ultra light nuclear shielding, which opens the way towards portable nuclear power sources based on fission and hybrid fusion-fission power sources.

The principle of quantum diverter

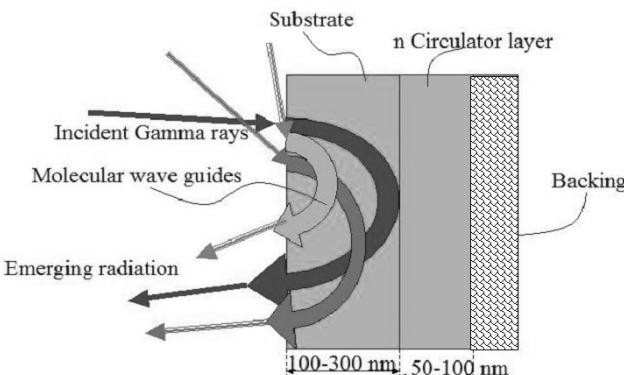


Fig. 9.24 Nano-Structured Radiation Guides (NRG) For Electronically Controlled Nuclear Reflectors And Ultra-Light Radiation Shielding

As Fig. 9.24 shows, gamma rays and neutrons are guided inside molecular wave-guides that can be perturbed by a quantum diverter, which when electrically excited, introduces supplementary quantum states that generates scattering or kicks-out the channel of the trapped, guided particles. This is a subject under the disclosure process – and if successful will enable a very thin shield with the transmission/reflection coefficient electronically controlled, usable for high penetrating radiation and will dramatically improve on reactor design and shielding. It will further reduce the critical mass to less than 1/10th of the amount used in actual explosive devices; a new generation reactor becoming critical with less than 50g of ²³⁹Pu is predicted. Theoretical predictions show that versions of this material may be used for radiation protection suits (attenuating the radiation level by >1000 times), ultra-low-weight radiation shielding for space shuttles and stratospheric airplanes, even for ultra high power modulation from the nuclear reactor source, etc.

9.4.5 Generation 8

Generation 8 would consist of hybrid nuclear reactors using fusion-fission sources based on super-colliders or plasma devices. Fusion has the specific yield of 3 times more than fission. The most interesting reactions are those free of neutrons as shown those given in Table 9.12, because the neutron's energy is harder to harvest.

Table 9.12 Main fusion reactions

Aneutronic Nuclear Fusion Reaction	Energy [MeV]	Specific En. [MeV/nucl.]
D + ³ He → ⁴ He (3.6 MeV) + p (14.7 MeV)	18.3	3.66
D + ⁶ Li → 2 ⁴ He + (22.4 MeV)	22.4	2.80
p + ⁷ Li → 2 ⁴ He + 17.2 MeV	17.2	2.15
³ He + ³ He → ⁴ He + 2 p + 12.86 MeV	12.9	2.14
³ He + ⁶ Li → 2 ⁴ He + p + 16.9 MeV	16.9	1.88
p + ¹¹ B → 3 ⁴ He + 8.7 MeV	8.7	0.73
p + ⁶ Li → ⁴ He (1.7 MeV) + ³ He (2.3 MeV)	4	0.57

Nano-hetero structured material solves both the problem of energy harvesting from fission products and fusion charged particles as well as the problem of neutron energy amplification by fission and harvesting.

Fusion represents the future of space missions because it is less polluting and has no critical mass accumulation on board restrictions. Critical mass accumulation will ultimately limit the future of high power space nuclear systems. Fusion based power systems are the only systems able to operate with manned missions up to the borders of the solar system (Popa-Simil 2007).

Fission systems have limitations on the total energy due to criticality conditions, but they can be used as “small” continuous utility power sources.

9.4.5.1 Complex Multipurpose Fusion and Hybrid Systems

Nuclear power is not good enough to sustain manned intergalactic missions in spite of the huge energy it can deliver, because of the ultra-low conversion efficiency between the energy released on board and that converted to space shuttle kinetic energy. The mass penalty is still unacceptable for deep space missions, without applying the other constraints for manned missions in space.

Another interesting feature shown in Fig. 9.25, is that energy harvesting tiles may be configured in antennae, radiation collectors, or radiation source units (^{238}Pu long-life (200 yr.) battery with about 4 w/cm³) (Popa-Simil 2005c, 2006c).

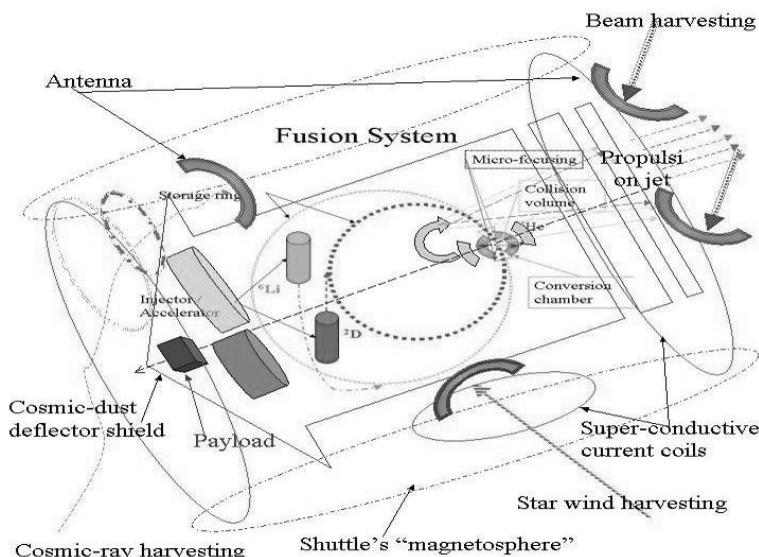


Fig. 9.25 Space power system

Nano-shields are appropriate for shielding life on the shuttle or to create individual protection shields, cutting all of the radiation spectrum energy by more than 100 times. Compared with the previous calculations of the data, the most advantageous design is to place the payload surrounded by liquid hydrogen tanks 1 m in diameter or larger in order to cut the cosmic radiation in half. The new shielding might remove this impediment with a more efficient and lighter structure.

9.4.5.2 Present Fusion Researches

There is an intense research for obtaining functional devices using various fusion reactions as:

- Hydrogen-boron fusion, using dense plasma focus, funded by NASA's Jet Propulsion Laboratory and the Air Force Research Laboratory (Focus Fusion Society 2008).

- An inertial electrostatic confinement device enhanced by magnetic confinement fusion, creating a potential well to produce fusion, is intended to be used by US Navy for jet propulsion (Bussard and Jameson 1995)
- Z-pinch hydrogen-boron reactor at Sandia National Laboratory Z-machine (Haines 2006)
- Colliding beam fusion, based on Field-Reversed Configuration ion beam col-lider with an increased ion density (Rostoker 2006).
- Electron spiral toroid to produce self-confined toroidal plasmoids, without ex-ternal magnetic fields (Seward 2002).
- p-¹¹B fusion in compact spheromak based on high conducting plasma shell con-finment (Koloc 1973; 2002).
- Hydrogen-boron aneutronic fusions by picosecond laser inertial confinement (Belyaev 2008)

To date, none of the fusion research delivered the expected power, have bulky construction, and are not suited for space applications in spite of the fact that there are many possibilities for space propulsion or energy production (Slough 2001).

9.4.5.3 Anti-matter Stimulated Fusion

An interesting electric-power-fusion-based solution was generated by the long-standing challenge of developing a means, to travel to the stars within the life span of a human being (Lewis 1999). In their book Mirror Matter they review several proposed concepts, including rockets utilizing nuclear fusion reactions or antimatter annihilation, as well as rocket less systems such as laser-pushed light-sails (Forward 1988). Each concept faces fundamental technological problems. For example, fusion rockets require massive (several hundred tons) high-powered "driv-ers", being either a laser or particle beam array (ICF systems) or a magnetic torus or mirror (MCF systems). Antimatter systems generally call for quantities of anti-matter far in excess of current or conceivable future production capabilities. Rocket less systems require intense high power laser or particle beams with excep-tional focusing requirements, etc.

A new concept, called Antimatter Initiated Micro-fusion (AIM), which is a hybrid of nuclear fusion and antimatter technologies has been developed (Lewis 1999). Any system designed for deep space missions must meet the following criteria:

- high specific power: $\alpha > 1 \text{ kW/kg}$;
- high exhaust velocities: $v_{\text{ex}}(\text{max}) = 10^4 \text{ km/s}$; and
- continuous power with near zero maintenance for several years.

The parameters of the AIM Star concept and expected performance for stellar missions are given in Table 9.13.

Table 9.13 AIM Star 50 year Mission to 10,000 A.U.

Fuel and fusion reaction	$D-T \Rightarrow n(14.1 \text{ MeV}) + \alpha(3.5 \text{ MeV}),$	$D-\bar{He}^3 \Rightarrow p(14.7 \text{ MeV}) + \alpha(3.6 \text{ MeV}),$
ΔV	956 km/s	956 km/s
V_e	$5.98 \times 10^5 \text{ m/s}$	$5.98 \times 10^5 \text{ m/s}$
I_{sp}	61,000 s	61,000 s
Power	33 MW	0.75 MW
Thrust	55.2 N	1.25 N
dm/dt	$9.22 \times 10^{-5} \text{ kg/s}$	$2.09 \times 10^{-6} \text{ kg/s}$
t_b	0.50 yr = 6 mo	.22 yr
Distance @ burnout	37 AU	1635 AU
α_{ave}	30.5 kW/kg	0.69 kW/kg
N_{pbar}	130 mg	28.5 mg

$\Delta V/V_{ex} = 1.6$ for energy optimization (from the rocket eq., this fixes payload-to-total-mass ratio = 0.2), and 361 kg dry mass to 1444 kg propellant. The authors propose to inject small fusion fuel droplets into a cloud of antiprotons confined in a very small volume within a reaction Penning trap (Lewis 1999). The reaction trap is roughly the size of a shoebox, weighing perhaps 10 kg. It is fed 10^{11} antiprotons on a periodic basis from a portable trap positioned about 1 meter away on axis, safe from fusion debris. Radial confinement within a 0.8 cm maximum diameter orbit is provided by a 20T axial magnetic field. Axial trapping of the 2 cm long cloud of antiprotons is done within a 10 keV space charge electric potential from the electrodes. The key ingredient for heating of the 42 ng D- 3 He liquid droplet is an anti-proton-induced fission fragment that has a range of 45 μm in the droplet. In order to spark the micro-fusion process, 5×10^8 antiprotons are annihilated in a 2% molar admixture of a heavy metal, such as ^{208}Pb or ^{238}U , with the D- 3 He. Annihilation takes place on the surface of the antiproton cloud, peeling back 0.5% of the cloud. The power density released by the fission fragments into the D- 3 He is about $5 \times 10^{13} \text{ W/cm}^3$, fully ionizing the D and 3 He atoms. This is roughly comparable to a 1 kJ, 1 ns, laser depositing its energy over a 200 μm ICF (Inertial Confinement Fusion) target, a system much too massive for driving a small space probe.

It has been shown recently that the fission fragments from antiproton-induced fission are not radioactive, so there is no concern of accumulative radioactive contamination of the engine and spacecraft as the engine burns (Smith 1997).

This antimatter concept is a useful one providing a light fusion reactor with the price of antimatter. This structure may be used in conjunction with a direct conversion tile to harvest the fusion particles' energy.

9.5 Conclusions

Mars colonization represents a challenging mission and has many requirements. The power systems (with emphasis on a large diversity of nuclear systems) represent a part of the problem, but an important one because without sufficient available

power, the mission will have to sacrifice safety, reliability, and flexibility in order to launch. The specific mass /power ratio, with limitation to masses difficult to place in orbit, transport and land on Mars is a very strong constraint. Improving this ratio requires investment in the pertinent technologies.

Material advances in fuels and refractory metals will increase reliability and lower specific mass. Stronger, lighter materials will allow for higher temperatures, which will increase efficiencies and lower radiator masses.

Advances in static energy conversion sources with emphasis on new generations of direct energy conversion have the long term potential to revolutionize space nuclear power and make dynamic conversion obsolete.

Mars needs to be explored smartly by humans on its surface. Humans have the endurance and the knowledge to widen their frontier beyond the moon, but they need an international commitment to join together and accomplish this ever-present challenge by using the exceptional potential and imagination of the entire human race.

References

- Adler, K., Ducommun, G.: Miniaturized Nuclear Battery US Patent 3934162 (1976)
- Ajayan, P., Burlage, H.: Nanotube Foams Flex And Rebound With Super Compressibility in Newswise (2005)
- Anghaie Smith, B., Knight, T.: Direct Energy Conversion Fission Reactor Gaseous Core Reactor With Magnetohydrodynamic (MHD) Generator in DE-FG 03-99 SF21894, Final Report (2002)
- Anthony, T., Schenectady, R., Cline, H.E.: Deep Diode Atomic Battery in US Patent 4024420 (1977)
- Apostol, M.: On Pulse-Like Thermoelectric Transport. *J. Theor. Phys.* 69, 12 (2001)
- Apostol, M.: A Pulsed Thermoelectric Machine. *J. Theor. Phys.* 91, 13 (2003)
- Apostol, M.: Pulsed Thermoelectric Machine. *J. Theor. Phys.* 1, 161 (2007a)
- Apostol, M.: Pulsed Thermoelectric Machine. *J. Theor. Phys.* 161, 12 (2007b)
- Apostol, M.: Generalized Theory Of Thermoelectric Figure Of Merit. *J. Appl. Phys.* 104(053704), 3 (2008)
- Ato, Y., Miyagawa, S.: Method Of Converting Radioactive Energy To Electric Energy And Device For Performing. The Same in US Patent 3939366 (1976)
- Avakian, R.O., Ispirian, K.A., Yaralov, V.J.: Calculation Of Planar Channeling Radiation Produced By 20 Mev Electrons. In: Quartz in NATO Science Series: Mathematics, Physics And Chemistry 199 (2004)
- Bakker, K., Frodo, C.: The Use Of Molybdenum-Based Ceramic-Metal (Cermet) Fuel For The Actinide Management In LWRs. *Nuclear Technology* 146(3), 325–331 (2004)
- Bazilevskii, A.T., Mokaleva, L.P., Manvelian, O.S., Surkov, I.A.: Evaluation Of The Thorium And Uranium Contents Of Martian Surface Rock - A New Interpretation of Mars-5 Gamma-Spectroscopy Measurements. *Geokhimiia*, January 10-16 (1981) (in Russian)
- Belyaev, V.S.: Observation Of Neutronless Fusion Reactions In Picosecond Laser Plasmas. *J. Appl. Phys.* 104, 053704 (2008)
- Blanchard, P.: Nuclear Power Minibatteries in Wired News (2005)

- Brown, L.C.: Direct Energy Conversion Fission Reactor in Annual Report To The U.S. Department of Energy GA-A23593 (2000)
- Broxon, J.W., Jesse, W.P.: Differential Ion Chamber in US Patent 2440167 (1948)
- Bussard, R.W., Jameson, L.W.: Inertial-Electrostatic-Fusion Propulsion Spectrum: Air-Breathing To Interstellar Flight. *J. Propulsion And Power* 11(2) (1995)
- Carmack, W.J., Husser, D.L., Mohr, T.C., Richardson, W.C. (2004) Status Of Fuels Development And Manufacturing For Space Nuclear Reactors At BWX Technologies. In: STAIF, Albuquerque, NM p. 426 (February 2004)
- Degueldre, C., Arima, T., Lee, Y.W.: Thermal Conductivity of Zirconia Based Inert Matrix Fuel: Use And Abuse of The Formal Models For Testing New Experimental Data. *J. Nuclear Materials* 319, 6–14 (2003)
- Douglas, F.J., Schulze, R.K., Neu, M.P.: Surface Chemistry of Pu Oxides. *J. Nuclear Materials* 328, 2–3 (2004)
- Eric, G., Karl, S.: Secondary Emission Type Of Nuclear Battery in US Patent 2858459 (1958)
- Fink, J.K.: Thermal Conductivity And Thermal Diffusivity Of Solid UO₂ in International Nuclear Safety Center (INSC) Preliminary Recommendation (Anl. Gov), 28 (2004)
- Focus Fusion Society, Focus Fusion: The Fastest Route To Cheap, Clean Energy on Web (2008)
- Forward, R.L., Davis, J.: Mirror Matter. Wiley Science, New York (1988)
- Gadeken, L.: Apparatus And Method For Generating Electrical Current From The Nuclear Decay Process of A Radioactive Material in US Patent 6774531 (2004)
- General Electric Company, Sp-100 Project Integration in Meeting Notes Long Beach, Ca, July 19-21 (1988)
- Haines, M.: Viscous Heating Of Ions Through Saturated Fine-Scale MHD Instabilities In A Z-Pinch At 200-300 Kev Temperature. *Phys. Rev. Lett.* 96(075003) (2006)
- Hoff Jr, F.B.: Electron-Voltaic Effect Device in US Patent 3714474 (1970)
- Inoue, M.: Thermal Conductivity Of Uranium-Plutonium Oxide Fuel For Fast Reactors. *J. Nuclear Materials* 282(2/3), 186–195 (2000)
- Kallmann, H.I., Kuhn, E.: Device For Measuring The Intensity of A Radiation of Slow Neutrons By Means of Ionization Chamber in US Patent 2288718 (1940)
- Koloc, P.M.: Method And Apparatus For Utilizing A Compound Plasma Configuration in US Patent No. 4,023,065 (1973)
- Koloc, P.M.: The Engineering Physics Of An Optimized Confinement Concept (The Plasmak). In: ICC 2002, Innovative Confinement Concepts, College Park, MD (2002)
- Kumarov, M.: Channeling Radiation: History Of Discovery. In: Status And Use in SPIE Proceedings Series 5974(597403), pp. 1–15 (2004)
- Lal, A., Guo, H., Li, H.: The Engines Of Lilliput, *The Economist* (March 2002)
- Lal, A., Bilbao, R.M., Leon, Y., Guo, H., Li, H., Santanam, S., Yao, R., Blanchard, J., Henderson, D.: A Nuclear Microbattery For Mems Devices in Energy. Wisconsin (2005)
- Lanning, D.D., Painter, C.L.: Frapcon-3: Modifications To Fuel Rod Material Properties And Performance Models For High-Burnup Application. In: Nuclear Regulatory Commission, Nureg/Cr-6534 U.S. 1(PNNL-11513), pp. 1–124 (1997)
- Lewis, A.R., Meyer, K., Smith, G.A., Howe, S.D.: Aimstar: Antimatter Initiated Microfusion For Pre-Cursor Interstellar Missions, Laboratory For Elementary Particle Science, Department Of Physics, The Pennsylvania State University, NASA STTR Phase I (1999)

- Linder, E.G.: Method And Means For Collecting Electric Energy Of The Nuclear Reactions in US Patent 2517120 (1946)
- Lipinski, R.J., Wright, S.A., Lenard, R.X., Harms, G.A.: A Gas-Cooled Reactor Surface Power System. In: E1-Genk, M.S. (ed.) AIP Conference Proceedings of Space Technology and Applications International Forum 1999 (STAIF 1999), vol. (458), pp. 1470–1475 (1999)
- Little, R.C., Fisher, H.M., Alam, B., Block, R.C., Harris, D.R., Slovaceck, R.E.: Monte Carlo Modeling of The Neutronics of A Lead Slowing Down Time Assay Device. Transactions of the American Nuclear Society 43, 119–121 (1999)
- Loaiza, D.J., Stratton, W.: Criticality Data For Spherical 235U, 239Pu, and 237Np Systems Reflector-Moderated By Low Capturing-Moderator Materials. Nuclear Technology 146(2), 143–154 (2004)
- Mason, L.S.: A Comparison Of Brayton And Stirling Space Nuclear Power Systems For Power Levels From 1 Kilowatt To 10 Megawatts. In: STAIF, Albuquerque, NM, p. 4 (February 2001)
- Maue, H.W., Meinzer, R., Paradiso, T.J., Feldmeier, M.: Self-Powered Wireless Switch in US Patent 6700310 (2004)
- Mcginnis, S.J.: Nuclear Power Systems For Manned Mission To Mars in Astronautical Engineering, Monterey, California, United States Naval Postgraduate School, 105 (2004)
- Metcalf, K.J.: Power Conditioning System Modelling For Nuclear Electric Propulsion in NASA Jet Propulsion Laboratory Canoga Park, Ca(Cr-191136) (1993)
- NASA, Project Prometheus (2008),
<http://Exploration.Nasa.Gov/Programs/Prometheus.html>
- Newkirk, D., Salama, S.A., Stewart, S.L.: Sp-100 Scaleup To 40 MWe (General Electric Company American Institute Of Physics) report (1991)
- Polansky, G.: Direct Energy Conversion Fission Reactor in Proposal No. 99-0199 (1999)
- Popa-Simil, L.: Multi-Thermocouple Structure For Measuring The Target's Temperature Field During Beam Power Deposition in Internal Report. Bucharest, NIPNE-HH, Accelerator Division, p. 2 (1986)
- Popa-Simil, L.: Hetero-Multi-Layer Nuclear Fuel For 0.1-1GWe Solid-State Nuclear Reactor Structure. In: Proceedings of MRS 2005 Spring Meeting (2005a)
- Popa-Simil, L.: The Use Of Ultra-Thin Layers Structure To Obtain A Self Cleaning Nuclear Reactor Fuel. In: Proceedings of MRS - Spring Meeting, Heat And Mass Transport At Nanoscale - From Fundamentals To Devices (Symposium: Dd) paper 12679 (2005b)
- Popa-Simil, L.: On The Possibility Of Using A Solid State Nuclear Reactor - Accelerator Propulsion. In: Proceedings Of The Space Nuclear Conference 2005, San Diego, California, paper 1114 (2005c)
- Popa-Simil, L.: Cryogenic Nuclear Reactor Fuel And Blankets For Energy Harvesting. In: Proceedings of MRS 2006, Fall, Boston, MA (2006a)
- Popa-Simil, L.: Long Life Single Load Reactor Fuel. In: Proceedings of ICAPP 2006, vol. (1), pp. 140–148 (2006b)
- Popa-Simil, L.: Space Shuttle Nuclear Antenna For Deep Space Remote Power Transfer. In: Proceedings of ICAPP 2006 NV (2006c)
- Popa-Simil, L.: Liquid Metal Cooled Variable Geometry Reactor Structure. In: Proceedings Of ICAPP 2007, Nice, France, May 13-18, paper 7104 (2007a)
- Popa-Simil, L.: Micro-Hetero Fuel Structure With Minimized Radiation Damage. In: Proceedings Of ICAPP 2007, Nice, France, May 13-18 p. 7102 (2007b)
- Popa-Simil, L.: Micro-Structured Nuclear Fuel And Novel Nuclear Reactor Concepts For Advanced Power Production in Advance. Nuclear Reactor Design INES 3, 12 (2007c)

- Popa-Simil, L.: Nuclear Energy Among Very Few Alternative For Future Clean Reliable Power. In: Proceedings of ICAPP 2007, Nice, France, May 13-18, p. 7101 (2007d)
- Popa-Simil, L.: The Nuclear Power Demand And Limitations For Deep Space Exploration. In: Proceedings of Space Nuclear Conference Boston, MA (2007e)
- Popa-Simil, L.: Intergalactic Nuclear Power Systems. In: Proceedings Of ICAPP 2007, Nice, France, May 13-18, paper 7099 (2007f)
- Popa-Simil, L.: Nanotube Potential Futur. In: Nuclear Power MRS-Spring (2008)
- Popa-Simil, L., Muntele, C.: Nano-Structured Materials As Deflective Radiation Shielding. In: Proceedings of MRS-Fall 2007 Meeting (2007)
- Poston, D.I.: The Heatpipe-Operated Mars Exploration Reactor (Homer) Los Alamos National Laboratory Report LA-UR 5299 (2000)
- Poston, D.I., Voit, S.L., Reid, R.S., Ring, P.J.: The Heatpipe Power System (HPS) For Mars Outpost and Manned Mars Missions. In: El-Genk, M.S. (ed.) Space Technology And Applications International Forum 2000 (STAIF 2000), Aip Conference Proceedings, vol. (504), pp. 1327–1334 (2000)
- Rappaport, P.: Radioactive Batteries in US Patent 3094634 (1953)
- Ritter, J.C.: Radioisotope Photoelectric Generator in US Patent 4178524 (1979)
- Rockwell International, Ultra High Power Space Nuclear Power System Design And Development in NASA Report CR 2001-210767, 11 (2001)
- Romano, A., Hejzlar, P., Toderas, N.E.: Fertile-Free Fast Lead-Cooled Incinerators For Efficient Actinide Burning. Nuclear Technology 147(3), 368–387 (2004)
- Ronchi, C.: Thermal Conductivity Of Uranium Dioxide Up To 2900 K From Simultaneous Measurement Of The Heat Capacity And Thermal Diffusivity. J. Appl. Phys. 85(2), 776–789 (1999)
- Rostoker, N.: The Colliding Beam Fusion Reactor (CBFR). UC Irvine Press Release (2006)
- Saito, J., Morinaga, M., Kano, S., Furui, M., Noda, K.: Corrosion Behavior Of Mo-Re Based Alloys In Liquid Li. J. Nuclear Materials 264, 1–2 (1999)
- Schuyler, M.C.: Radiation Responsive Voltage Sources in US Patent 2847585 (1958)
- Seward, C.: Six-Kilowatt Power Supply Using The Colliding Est Fusion Reactor in Electron Power Systems, White Paper (2002)
- Shanks, H.R.: Nuclear Battery in US Patent 5721462 (1998)
- Shimabukuro, R.L., Russell, S.D.: Micro-Power Source in US Patent 6911711 (2005)
- Slough, J.: Performance Capability And Mission Analysis For A Pulsed High Density FRC Fusion Rocket. In: Proceedings of AIAA 2001- 3674 Salt Lake City, Utah, July 8-11 (2001)
- Smith, G.A.: Antiproton-Catalyzed Microfission/Fusion Propulsion Systems For Exploration Of The Outer Solar System And Beyond. In: Proceedings of JPL Workshop On Advanced Propulsion, Pasadena, CA, May 20-23 (1997)
- Snyder, G.J., Patel, J., Fleurial, J.P.: Extremely-Efficient, Miniaturized, Long-Lived Alpha-Voltaic Power Source Using Liquid Gallium in US pat.6700298 (2004)
- Space.Com, Havo in The Heavens: Soviet Era Satellite's Leaky Reactor's Lethal Legacy (2003), http://www.space.com/news/mystery_monday_040329.html
- Space.com, Falling on A City Near You: Dangerous Space Reentries (2006),
http://www.space.com/news/spacehistory/dangerous_reentries_000602.html
- Spacex Launch Services, Falcon 1 Overview of Space Exploration Technologies (2008),
<http://www.spacex.com/>
- Toffer, H., Roblyer, S.P.: Hanford Site Production Reactor Data Pertinent To Actinide Burning. ANS Transactions 19(1), 455–460 (1992)

- Tsang, T.F.: The Liquid Electronics Advanced Power Systems (Leaps), 1–3 (2004),
http://www.darpa.mil/GTI_News/
- Vermoyal, J.J., Frichet, A., Dessemond, L.: Contribution To The Understanding Of The ZrNb(1%)O(0.13%) Oxidation Mechanism at 500 °C in Dry Air. *J. Nuclear Materials* 328(1), 31–45 (2004)
- VanDine, J.E., Kiss, Z.: Directed Energy Conversion Of Semiconductor Materials in US Patent 4675467 (1987)
- Wallenius, J.: Thermal Conductivity and Transmutation Course (2004),
<http://www.neutron.kth.se/courses/Transmutation.shtml>
- Willson, V.C.: Generator Of Power in US Patent 2728867 (1955)
- Yarygin, V.I.: Thermoelectricity And Thermionic. In: Space Nuclear Power Systems Using Direct Energy Conversion Current Status And Prospects in Space Nuclear Conference Proceedings, SNC 2005 (2005)
- Young, R.D., Hageman, J.P., Light, G.M., Seale Jr., S.W.: Charged-Particle Powered Battery in World Patent WO 9748105A 1 (1997)
- Young, R.D., Hageman, J.P., Light, G.M., Seale Jr., S.W.: Charged-Particle Powered Battery in US Patent 5861701 (1999)

Chapter 10

New Solutions for Nuclear Energy and Flights on Mars

Alexander A. Bolonkin

CR & C, New York, USA

10.1 Introduction

The very rarefied Mars atmosphere and the low gravitation allow implementation of various technologies difficult to accomplish on Earth. Two projects based on such technologies are described in this chapter, related to energy delivery and transportation.

A Mars colony and Mars transportation, mining, life support and operations—especially industrial operations - will need reliable energy sources. But Mars does not have available chemical fuel and atmospheric oxygen. Also, the global dust storms severely penalize the performance of solar installations. Of course, these installations cannot be used during the long polar nights. These comments suggest the utilization of nuclear power; and we shall now consider the thermonuclear reactors. Conventional thermonuclear reactor designs, such as those constructed today, would weigh thousands of tons and are not acceptable for delivery to Mars. Section 10.2 proposes a new lightweight thermonuclear reactor to provide electrical energy on Mars. In Sec. 10.3 a new method for the transportation of payloads on Mars is proposed, by using a spaceport placed on the highest of Mars' mountains.

10.2 Thermonuclear Reflect AB-Reactor on Mars

Most part of this section is based on Bolonkin (2008b).

10.2.1 Brief Information about Thermonuclear Reactors

Fusion power is useful energy generated by nuclear fusion reactions. In this kind of reaction two light atomic nuclei fuse together to form a heavier nucleus and release energy. The largest current experiment, Joint European Torus (JET), has resulted in fusion power production somewhat larger than the power brought into the plasma, maintained for a few seconds. In June 2005, the construction of the

International Thermonuclear Experimental Reactor (ITER), designed to produce several times more fusion power than the power brought into the plasma, for several minutes, was announced. The production of net electrical power from fusion is planned for the next generation experiment after ITER.

Unfortunately, this task is not easy, as scientists thought earlier. Fusion reactions require a very large amount of energy to initiate in order to overcome the so-called *Coulomb barrier* or *fusion barrier energy*. The key to practical fusion power is to select a fuel that requires the minimum amount of energy to start, that is, the lowest barrier energy. The best fuel from this standpoint is a one-to-one mix of deuterium and tritium; both are heavy isotopes of hydrogen. The D-T (Deuterium & Tritium) mix has a low barrier energy. In order to create the required conditions, the fuel must be heated to tens of millions of degrees, and/or compressed to immense pressures.

At present, D-T is used by two main methods of fusion: inertial confinement fusion (ICF) and magnetic confinement fusion (MCF)(for example, the tokamak). In *inertial confinement fusion*, nuclear fusion reactions are initiated by heating and compressing a target. The target is a pellet that most often contains deuterium and tritium (often only micro or milligrams). Intense laser or ion beams are used for compression. The beams explosively detonate the outer layers of the target. That accelerates the underlying target layers inward, sending a shockwave into the center of pellet mass. If the shockwave is powerful enough and if high enough density at the center is achieved some of the fuel will be heated enough to cause fusion reactions. In a target which has been heated and compressed to the point of thermonuclear ignition, energy can then heat surrounding fuel to cause it to fuse as well, potentially releasing tremendous amounts of energy. A few words about *Magnetic confinement fusion*. Since plasmas are very good electrical conductors, magnetic fields can also confine fusion fuel. A variety of magnetic configurations can be used, the basic distinction being between magnetic mirror confinement and toroidal confinement, especially tokamaks and stellarators.

One of the earliest (in the late 1970's and early 1980's) serious attempts at an ICF design was Shiva, a 20-armed neodymium laser system built at the Lawrence Livermore National Laboratory (LLNL) that started operation in 1978. Shiva was a "proof of concept" design, followed by the NOVA design with 10 times the power. Funding for fusion research was severely constrained in the 80's, but NOVA nevertheless successfully gathered enough information for a next generation machine whose goal was ignition. Although net energy can be released even without ignition (the breakeven point), ignition is considered necessary for a *practical* power system. The resulting design, now known as the National Ignition Facility, commenced being constructed at LLNL in 1997. Originally intended to start construction in the early 1990s, the NIF is now six years behind schedule and overbudget by over \$1.4 billion. Nevertheless many of the problems appear to be due to the "big lab" mentality and shifting the focus from pure ICF research to the nuclear stewardship program, LLNLs traditional nuclear weapons-making role. NIF is now scheduled to "burn" in 2010, when the remaining lasers in the 192-beam array are finally installed.

Laser physicists in Europe have put forward plans to build a £500m facility, called HiPER, to study a new approach to laser fusion. A panel of scientists from seven European Union countries believes that a "fast ignition" laser facility could make a significant contribution to fusion research, as well as supporting experiments in other areas of physics. The facility would be designed to achieve high energy gains, providing the critical intermediate step between ignition and a demonstration reactor. It would consist of a long-pulse laser with an energy of 200 kJ to compress the fuel and a short-pulse laser with an energy of 70 kJ to heat it. Confinement refers to all the conditions necessary to keep a plasma dense and hot long enough to undergo fusion, i.e.:

- *Equilibrium.* There must be no net forces on any part of the plasma, otherwise it will rapidly disassemble. The exception, of course, is inertial confinement, where the relevant physics must occur faster than the disassembly time.
- *Stability.* The plasma must be so constructed that small deviations are restored to the initial state, otherwise some unavoidable disturbance will occur and grow exponentially until the plasma is destroyed.
- *Transport.* The loss of particles and heat in all channels must be sufficiently slow. The word "confinement" is often used in the restricted sense of "energy confinement".

To produce self-sustaining fusion, the energy released by the reaction (or at least a fraction of it) must be used to heat new reactant nuclei and keep them hot long enough that they also undergo fusion reactions. Retaining the heat generated is called energy *confinement* and may be accomplished in a number of ways. The hydrogen bomb weapon has no confinement at all. The fuel is simply allowed to fly apart, but it takes a certain length of time to do this, and during this time fusion can occur. This approach is called inertial confinement. If more than about a milligram of fuel is used, the explosion would destroy the machine, so controlled thermonuclear fusion using inertial confinement causes tiny pellets of fuel to explode several times a second. To induce the explosion, the pellet must be compressed to about 30 times solid density with energetic beams. If the beams are focused directly on the pellet, it is called direct drive, which can in principle be very efficient, but in practice it is difficult to obtain the needed uniformity. An alternative approach is indirect drive, in which the beams heat a shell, and the shell radiates x-rays, which then implode the pellet. The beams are commonly laser beams, but heavy and light ion beams and electron beams have all been investigated and tried to one degree or another. They rely on fuel pellets with a "perfect" shape in order to generate a symmetrical inward shock wave to produce the high-density plasma, and in practice these have proven difficult to produce. A recent development in the field of laser-induced ICF is the use of ultra-short pulse multi-petawatt lasers to heat the plasma of an imploding pellet at exactly the moment of greatest density after it is imploded conventionally using terawatt scale lasers. This research will be carried out on the (currently being built) OMEGA EP petawatt and OMEGA lasers at the University of Rochester and at the GEKKO XII laser at the Institute for Laser Engineering in Osaka Japan which, if fruitful, may have the effect of greatly reducing the cost of a laser fusion-based power source.

At the temperatures required for fusion, the fuel is in the form of a plasma with very good electrical conductivity. This opens the possibility to confine the fuel and the energy with magnetic fields, an idea known as magnetic confinement. Much of this progress has been achieved with a particular emphasis on tokamaks.

In fusion research, achieving a fusion energy gain factor $Q = 1$ is called *break-even* and is considered a significant although somewhat artificial milestone. Ignition refers to an infinite Q , that is, a self-sustaining plasma where the losses are made up for by fusion power without any external input. In a practical fusion reactor, some external power will always be required for things like current drive, refueling, profile control, and burn control. A value on the order of $Q = 20$ will be required if the plant is to deliver much more energy than it uses internally. In a fusion power plant, the nuclear island has a plasma chamber with an associated vacuum system, surrounded by a plasma-facing components (first wall and divertor) maintaining the vacuum boundary and absorbing the thermal radiation coming from the plasma, surrounded in turn by a blanket where the neutrons are absorbed to breed tritium and heat a working fluid that transfers the power to the balance of plant. If magnetic confinement is used, a magnet system, using primarily cryogenic superconducting magnets, is needed, and usually systems for heating and refueling the plasma and for driving current. In inertial confinement, a driver (laser or accelerator) and a focusing system are needed, as well as a means for forming and positioning the pellets.

The magnetic fusion energy (MFE) program seeks to establish the conditions to sustain a nuclear fusion reaction in a plasma that is contained by magnetic fields to allow the successful production of fusion power.

In nuclear fusion research, the *Lawson criterion*, first derived by John D. Lawson in 1957, is an important general measure of a system that defines the conditions needed for a fusion reactor to reach ignition, that is, that the heating of the plasma by the products of the fusion reactions is sufficient to maintain the temperature of the plasma against all losses without external power input. As originally formulated the Lawson criterion gives a minimum required value for the product of the plasma (electron) density n_e and the "energy confinement time" τ . Later analyses suggested that a more useful figure of merit is the "triple product" of density, confinement time, and plasma temperature T . The triple product also has a minimum required value, and the name "Lawson criterion" often refers to this inequality. In thirty years, scientists have increased the Lawson criterion of the ICF and tokamak installations by tens of times. Unfortunately, all current and some new installations (ICF and totamak) have a Lawrence criterion that is tens of times lower than is necessary.

Economics. It is far from clear whether nuclear fusion will be economically competitive with other forms of power. The many estimates that have been made of the cost of fusion power cover a wide range, and indirect costs of and subsidies for fusion power and its alternatives make any cost comparison difficult. The low estimates for fusion appear to be competitive with but not drastically lower than other alternatives. The high estimates are several times higher than alternatives.

While fusion power is still in early stages of development, vast sums have been and continue to be invested in research. In the EU almost € 10 billion was spent on

fusion research up to the end of the 1990s, and the new ITER reactor alone is budgeted at € 10 billion. It is estimated that up to the point of possible implementation of electricity generation by nuclear fusion, R&D will need further promotion totalling around € 60-80 billion over a period of 50 years or so (of which € 20-30 billion within the EU). In the last EU research programme (FP6), nuclear fusion research received € 750 million (excluding ITER funding), compared with € 810 million for all non-nuclear energy research combined, putting research into fusion power well ahead of that of any single rivaling technology.

An important aspect of fusion energy in contrast to many other energy sources is that the cost of production is inelastic. The cost of wind energy, for example, goes up as the optimal locations are developed first, while further generators must be sited in less ideal conditions. With fusion energy, the production cost will not increase much, even if large numbers of plants are built.

Despite being technically non-renewable, fusion power has many of the benefits of long-term renewable energy sources (such as being a sustainable energy supply compared to presently-utilized sources and emitting no greenhouse gases) as well as some of the benefits of such much more finite energy sources as hydrocarbons and nuclear fission (without reprocessing). Like these currently dominant energy sources, fusion could provide very high power-generation density and uninterrupted power delivery (due to the fact that they are not dependent on the weather, unlike wind and solar power).

Several fusion reactors have been built, but as yet none has produced more thermal energy than electrical energy consumed. Despite research having started in the 1950s, no commercial fusion reactor is expected before 2050. The ITER project is currently leading the effort to commercialize fusion power.

To summarize, most efforts of scientists are directed today toward very large thermonuclear stations. These stations request gigantic finances, years of development and complex technology. One expects industrial thermonuclear electric stations to enter the market in the next 20 – 40 years and their energy will be more expensive than that provided by a current technology conventional electric station.

This chapter proposes a smaller and cheaper thermonuclear installation, which don't require decades for development and is ideally suited to be used on Mars. Smaller power units are actually more practical to be used on the Red Planet, whose future economy will depend on transportation. The small thermonuclear engines proposed here may be used for a large variety of vehicles operating on Mars surface.

10.2.2 Description and Innovation

10.2.2.1 AB Plasma Reflector

A three net plasma reflector was proposed in previous works (see, for example, Bolonkin 1983b-g). The plasma reflector (Fig. 10.1) has three conductive nets (1, 2, 3) and electric voltage between them. Voltage between nets 1 and 2, respectively, is about 100 kV in our case, when plasma temperature is 50 keV. The electric field intensity is taken so that 1, 2 reflect the positive charged particles

(for example atoms of deuterium D and tritium T). The voltage between the nets 2 and 3 nets is about 200 kV in our case. The electric field intensity is taken so the 2, 3 reflect the negative charged particles (electrons).

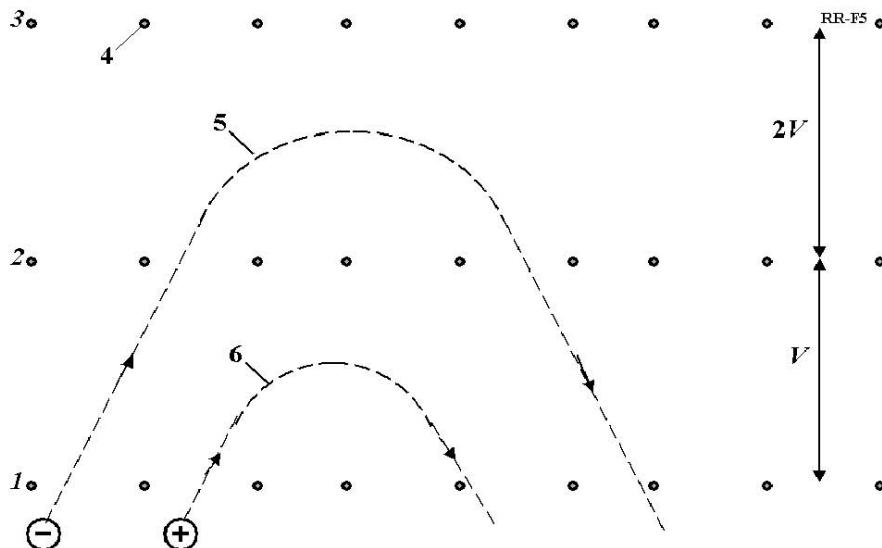


Fig. 10.1 Three net plasma AB-reflector. Notation: 1, 2, 3 – the first, second, and third nets; 4 - wire of net; 5 – trajectory of negative charged particles; 6 – trajectory of positive charged particles. V is voltage between nets 1 and 2 (about 100 kV in our case). The electric field intensity is taken so the 1, 2 reflect the positive charged particles (for example, atoms D and T). $2V$ is the voltage between the nets 2 and 3 (about 200 kV in our case). The electric field intensity is taken so the 2, 3 reflect the negative charged particles (electrons).

The reflector works in the following way. When plasma is inserted in the reflector confinement, the nets 1, 2 reflect the positive charged atoms back to the plasma if the electric voltage is more than plasma temperature in eV. The particle speed has a Maxwell distribution. That way the voltage between nets must be a minimum of two times more than the average temperature of the plasma. In this case, most positive particles will be reflected back to the plasma. Simultaneously, the nets 1,2 accelerate the negative particles (electrons) and increase their energy (temperature) additional up to 100 keV. That way the voltage between nets 2, 3 must be more than V plus the electron temperature of plasma (about $200 \text{ keV} > 100 + 50 \text{ keV}$ in our case).

The plasma contacts with the nets and heats them. This heating significantly depends upon the density of the plasma. Low density of hot plasma impedes significant heat transfer. A few comments follow. The inner Van Allen Belt extends from an altitude of 700–10,000 km above the Earth's surface and contains high concentrations of energetic protons with energies exceeding 100 MeV and

electrons in the range of hundreds of keV, trapped by the strong (relative to the outer belts) magnetic fields in the region. The charged particles in Earth's radiation belts have a temperature of some billions degrees, but a space apparatus orbiting there has a low temperature. It is possible because the density of plasma at these altitudes is very low and a heat flow from the high temperature plasma is small. This effect is used in the offered thermonuclear AB-reactor. In the present project the plasma has a density of a million times less than Earth's air and a hundred times less than a conventional tokomak machine. As a result, the net has an equilibrium temperature (after its own re-radiation of heat) of less than the melting temperature of a refractory conductive material (for example, tungsten has a melting temperature of 3689 K). Also, low density plasma produces low output power.

Development of a small power thermonuclear installation (hundreds and thousands of kW) based on the above principles is useful for vehicles to operate on Mars and will help to solve many future problems of the big thermonuclear stations on a smaller, more realizable scale. Developing small reactors should cost thousands of times less than building ITER or NOVA.

The computations show the main contribution to net heating may not be contact of the plasma to the net, but from interfacing with the products of the thermonuclear reaction – neutrons (14 MeV) and alpha particles (3.5 MeV). The neutrons have full reflection if the angle between trajectory and surface is 10 – 12 degrees. That way the author offers to use a wedge-like form for the cross-section of the net wire (Fig. 10.2b form 2). This form will reflect part of the high-energy particles and has more wire surface for thermal re-radiation than a circular cross section. We can also use the form Fig. 10. 2b, 3 which has an internal channel for cooling liquid. Another heat input into the net (of lesser magnitude, than heat transfer by plasma contact or thermonuclear particle effects) is produced by the Bremsstrahlung, or braking radiation. The contact heating is about 5 – 25% of the full energy.

Note that most heating occurs at the first net, which contacts with the nuclear particles. The second net only has contact with electrons. The mass of electrons and density of electron gas is less by about 2000 - 5000 times, than the mass of nuclear particles D, T. That way one cannot be troubled by heating of the second net. The third net is heated by reaction products (α , n) and by Bremsstrahlung radiation.

One can also use solid refractory material for nets, for example, tungsten, having a melting temperature of 3416 °C. One tungsten atom has an atomic mass of about 184. The deuterium D has an atomic mass of 2 and tritium T has an atomic mass of 3. That means the light atoms impact the heavy atom and passes to the heavy atom only a small part of its kinetic energy. The other important requirement of the net material is good tensile strength at high temperature. The tungsten has such a good property mix. Carbide of tantalum and zirconium has the melt temperature up to 3500 – 3900 °C. The nanotube has good stability up to 2300 °C in vacuum and excellent tensile stress.

Innovations brought by the plasma reflector in comparison with a conventional 'particle mirror' are the following:

1. The AB plasma reflector has three nets and it can reflect the plasma. Conventional electrostatic mirrors have two nets and they can reflect only charged particles of *one* particular sign (i.e. positive *or* negative).

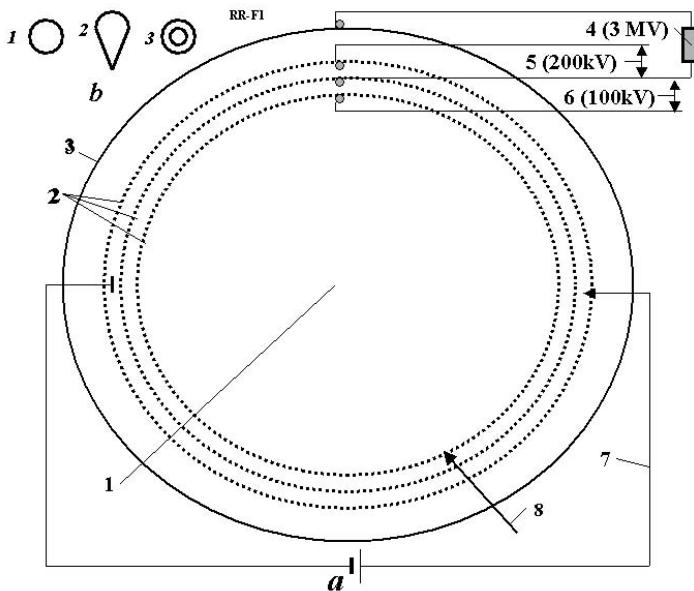


Fig. 10.2 Scheme of thermonuclear reflector AB-reactor, having spherical form. *Notations:* **a)** Reactor: 1 – plasma; 2 – three net AB-reflector; 3 – conductivity cover; 4 – voltage between net and cover; 5 – voltage between the second and third nets; 6 – voltage between the first and the second net; 7 – plasma electric heater; 8 – fuel injector. **b)** Cross section of the net wire. *Notation:* 1 – round; 2 – spherical (wedge) form; 3 – tube.

2. Net wires can have an internal cooling channel and heat can be removed that way (Fig. 10.2b-3).
3. The cross-section of net wires can have a special form (Fig. 10.2b-2) for reflecting the particles.
4. Net wires are made from heat-resistant electro-conductive material having good tensile stress in high temperature.
5. These net wires can be electronically cooled.

10.2.2.2 AB Thermonuclear Reflect Reactor for Isolated Places on Mars

A schematic of the proposed reactor is shown in Fig. 10.2. This implementation uses only charged α -particles and lets out (sets free) the neutrons and Bremstrahlung (X-soft ray) radiation (it doesn't protect from them). That significantly simplifies the reactor and decreases its integral mass (non-built in shielding is cheap to add on Mars surface, if necessary). The reactor has the spherical form and contains: AB spherical reflector enclosed D+T plasma, spherical conductivity cover, two sources of electric voltage (about 100 kV and 200 kV), any power consuming apparatus required (taps to external users), plasma injector, and plasma electric heater.

The Sphere's conducting cover is a thin (0.1 mm) aluminum film, which collects the positive charged high energy (3.5 MeV) (α) particles, created by thermonuclear reaction. When they move between net 3 and reactor cover (3, Fig. 10.2), they are

braked and accumulated by the cover. As a result, the thermonuclear reactor produces electric currents at about 3 MV in voltage. Part of this energy is used for heating the plasma, fuel, and support of the voltage in plasma reflector. The loss of plasma temperature from contact plasma with reflector net and Bremsstrahlung (X-ray) radiation is compensated by a plasma heater (Fig. 10.2, notation 7) by passing the electric current through the plasma (this is one of author inventions).

The reactor works continuously, producing tens to hundreds of kW of electric energy. That output harnesses only 20% of the liberated thermonuclear energy (efficiency may be sacrificed for the sake of lightness).

10.2.2.3 Thermonuclear Reactor for Martian Populated Regions

The reactor described above is not acceptable for operation in Martian regions with population because that version produces neutrons and X-ray radiation. The neutrons create radioisotopes which damage biological organisms. The reactor described in this section utilizes a special protection for men and environment from neutrons and, as well, Bremsstrahlung radiation. This protection increases the reactor mass by some (4-8) times. But that allows utilizing the full energy of reactor and produces the tritium – the second important component of thermonuclear fuel. The full energy received from the thermonuclear reactor increases by 4 times. About 40% this additional energy can be converted to electric energy (total is 60%) and the rest may be utilized for building and process heating.

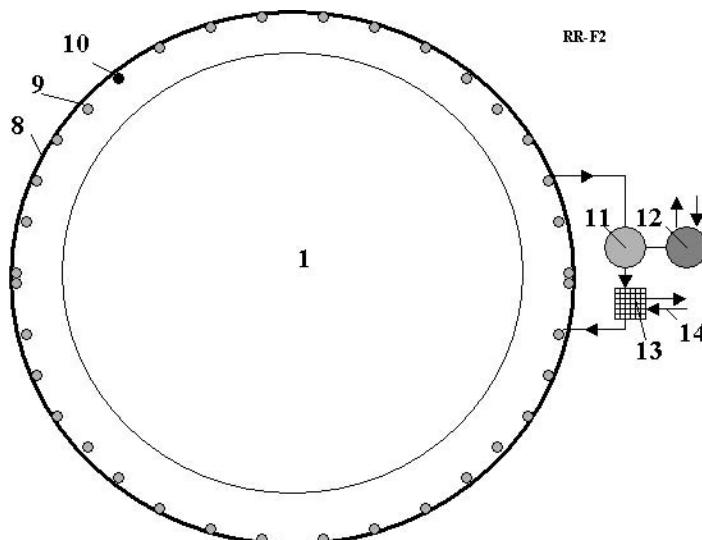


Fig. 10.3 Scheme of thermonuclear reflector AB-reactor to be used in populated regions on Mars. *Notations:* 1 – reactor of Fig. 10.2; 8 – protection against neutrons (neutron moderator); 9 – cooling (turbine) tubes; 10 – blanket for production of tritium; 11 - vapor turbine; 12 – electric generator; 13 – heat exchanger; 14 – hot water (vapor) for home heating (freshwater production, etc.).

This version of the thermonuclear reactor is shown in Fig. 10.3. Reactor contains the installation of Fig. 10.2 inserted into protection cover. Protective cover has cooling tubes 9, which connects to vapor turbine 11. The turbine drives the electric generator 12, producing useful energy. The vapor after turbine flows to the heat exchanger 13 which produces hot water (vapor) 14 for building and process heating. The reactor has lithium blankets 10 for reproduction of tritium. This design with no shielding can be used in case the reactor is located remotely (i.e. far away from populated regions) (Fig. 10.4).

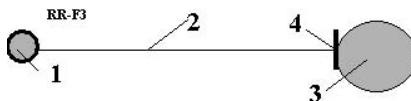


Fig. 10.4 Reflector AB-reactor operating far way from populated regions on Mars. *Notation:* 1 – reactor; 2 – connection to Martian colony; 3 – Artificial environment for a Martian colony; 4 – protection from neutrons and reactor radiation.

10.2.2.4 Thermonuclear Fuel

Deuterium (D), also called heavy hydrogen, is a stable isotope of hydrogen with a natural abundance in the oceans of Earth of approximately one atom in 6500 of hydrogen. The nucleus of deuterium, called a deuteron, contains one proton and one neutron. Deuterium occurs in trace amounts naturally as deuterium gas, written $^2\text{H}_2$ or D₂. Density: 0.180 kg/m³ at STP (0 °C, 101.325 kPa). Data at approximately 18 K for D₂ (triple point): Density: liquid - 162.4 kg/m³, gas - 0.452 kg/m³. Viscosity: 12.6 μPa·s at 300 Kelvin (gas phase). Specific heat capacity at constant pressure c_p : solid 2950 J/(kg·K), gas 5200 J/(kg·K)

Tritium (T or ^3H) is a radioactive isotope of hydrogen. It is a gas (T₂ or $^3\text{H}_2$) at standard temperature and pressure. Tritium combines with oxygen to form a liquid called tritiated water T₂O or partially tritiated THO. Tritium occurs naturally due to cosmic rays interacting with atmospheric gases. Because of tritium's relatively short half-life, however, tritium produced in this manner does not accumulate over geological timescales, and its natural abundance is negligible. Industrially, tritium is produced in nuclear reactors by neutron activation of lithium-6. Tritium is also produced in heavy water-moderated reactors when deuterium captures a neutron; however, this reaction has a much smaller cross section and is only a useful tritium source for a reactor with a very high neutron flux. It can also be produced from boron-10 through neutron capture. Tritium has a half-life time of only 12.32 years. That is why tritium does not exist in natural way on Earth. It may be found in small amounts on Mars, because the rarefied atmosphere there allows cosmic rays to reach the surface and produce tritium through bombardment of lithium isotopes.

Current nuclear fusion research is based on the D + T thermonuclear fusion reaction, which needs a high temperature about 1 - 5 keV:



This reaction can occur in high-temperature deuterium-tritium plasma. Most energy released by the reaction is converted to the kinetic energy of the neutron. Since the neutron is not confined or reflected by a magnetic or electrostatic field it leaves, going outwards to the environment or hits the screen or vessel wall (or blanket) immediately after the reaction. In last instance, the neutron kinetic energy is converted to heat. The heat is taken away from the screen by direct radiation or indirect circulating coolant and can be used to run an electric generator. If we add ${}^6\text{Li}$ inside the blanket, then tritium can be produced by reaction



with it then used subsequently as the fuel. Another reaction product is the alpha (α) particles ${}^4\text{He}$ carrying 3.5 MeV which can be directed or confined by electromagnetic field.

Table 10.1 Aneutronic reactions. Reaction energy in parentheses (MeV)

$\text{D} + {}^3\text{He} \rightarrow {}^4\text{He}(3.6) + \text{p}(14.7)$	$\text{p} + {}^{11}\text{B} \rightarrow {}^3\text{He}(8.7)$
$\text{D} + {}^6\text{Li} \rightarrow {}^2\text{He}(22.4)$	$\text{p} + {}^6\text{Li} \rightarrow {}^4\text{He}(1.7) + {}^3\text{He}(2.3)$
${}^3\text{He} + {}^3\text{He} \rightarrow {}^4\text{He}(4.3) + 2\text{p}(8.6)$	$\text{p} + {}^7\text{Li} \rightarrow {}^2\text{He}(17.2)$
${}^3\text{He} + {}^6\text{Li} \rightarrow {}^2\text{He}(1.9) + \text{p}(16.9)$	

The reaction that produces only charged particles are best for the proposed system. Unfortunately, these reactions are not easily achieved (Table 10.1). All of them require very high temperatures for ignition and have low cross-section of reaction (it is hard to produce large amounts of reactions and energy in an equivalent volume unit of the reactant). The D-T reaction is favored since it has the largest fusion cross-section (~ 5 barns peak) and reaches this maximum cross-section at the lowest energy (~ 65 keV center-of-mass) of any potential fusion fuel.

10.2.3 Theory and Results

The computation of the proposed AB- thermonuclear reflect reactor for Mars are made in this way. First, a reactor operating in a remote site is considered. Assign the radius of the first net $r = 0.5 - 10$ m, plasma density $n = 10^{18} - 10^{20} \text{ m}^{-3}$, and temperature of plasma $T = 1 - 50$ keV. [The plasma density is about $n = 10^{21} \text{ m}^{-3}$ in a conventional big tokomak]. For a numerical example, let us take: $r = 3$ m, $n = 10^{19} \text{ m}^{-3}$, $T = 50$ keV $= 5 \times 10^4 \times 1.16 \times 10^4 = 5.8 \times 10^8$ K. The surface and volume of a sphere are:

$$S = 4\pi r^2, \quad V = (4/3)\pi r^3, \quad S = 113 \text{ m}^2, \quad V = 113 \text{ m}^3 \quad (10.1)$$

Thermonuclear energy of reaction D + T for $T = 50$ keV released in form of charged particles is (in m^{-3}):

$$P_{DT} = 5.6 \cdot 10^{-13} \cdot 0.25 \cdot n^2 (\bar{\sigma v})_{DT} = 1.22 \cdot 10^4 \text{ W/m}^3, \quad (10.2)$$

where n enters in cm^{-3} and $(\bar{\sigma}v)_{DT}$ is taken from Table 10.2, where the additional reactions are:

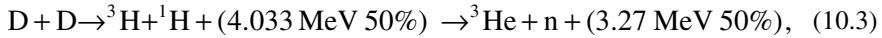


Table 10.2 Reaction rates $(\bar{\sigma}v)_{DT}$ (cm^3/s), averaged over Maxwellian distribution (American Institute of Physics 2003)

Temperature, keV	Reaction D - D	Reaction D - T	Reaction D-He ³
5	1.8×10^{-19}	1.3×10^{-17}	6.7×10^{-21}
10	1.2×10^{-18}	1.1×10^{-16}	2.3×10^{-19}
20	5.2×10^{-18}	4.2×10^{-16}	3.8×10^{-18}
50	2.1×10^{-17}	8.7×10^{-16}	5.4×10^{-17}
100	4.5×10^{-17}	8.5×10^{-16}	1.6×10^{-16}

The power density released in the form of charged particle is:

$$P_{DD} = 3.3 \cdot 10^{-13} n_D^2 (\bar{\sigma}v)_{DD}, \quad (10.4)$$

$$P_{DHe^3} = 2.9 \cdot 10^{-12} \cdot 0.25 \cdot n^2 (\bar{\sigma}v)_{DHe^3} \text{ (W/cm}^3\text{)}$$

where n enters in cm^{-3} and $(\bar{\sigma}v)_{DT}$ is taken from Table 10.2. Full thermonuclear energy is:

$$P_{DT,F} = P_{DT} V = 1.38 \cdot 10^6 \text{ (W)} \quad (10.5)$$

The energy per unit surface area is:

$$P_{DT,S} = r P_{DT} / 3 = 1.22 \cdot 10^4 \text{ (W/m}^2\text{)} \quad (10.6)$$

The number of reactions per unit volume is:

$$N = \frac{P_{DT}}{17.5 \text{ MeV}} = \frac{1.22 \cdot 10^4}{17.5 \cdot 10^6 \cdot 1.6 \cdot 10^{-19}} = 4.36 \cdot 10^{15} \text{ (m}^{-3}\text{)} \quad (10.7)$$

Fuel consumption (D+T) per unit volume in a day is:

$$C_{f,1} = N \cdot m_i = 0.157 \cdot 10^{-2} [\text{g}/(\text{day} \cdot \text{m}^3)] \quad (10.8)$$

Total daily fuel consumption is given by $C_f = C_{f,1} \times V$. The output power for fuel heating is

$$p_f = N V k T = 3.81 \cdot 10^3 \text{ (W)} \quad (10.9)$$

where $k = 1.38 \times 10^{-23}$ J/K is Boltzmann's constant. Computation of reflector net starts with obtaining the plasma pressure, i.e.

$$p = nkT = 8 \cdot 10^4 \text{ (N/m}^2\text{)} \quad (10.10)$$

The plasma pressure in one net is

$$p_1 = (1/3)p = 2.67 \cdot 10^4 \text{ (N/m}^2\text{)} \quad (10.11)$$

The thickness of the net as a continuous cover is

$$\delta = \frac{r p_1}{2\sigma} = 10^{-4} \text{ m} \quad (10.12)$$

where $\sigma = 4 \times 10^8 \text{ N/m}^2$ is safety tensile stress of net wire. The diameter net wire for net cell $L = 100 \times 100 \text{ mm}$ (for a distance between nets of 0.5 m) is

$$D = 2\sqrt{\frac{\delta L}{\pi}} = 3.57 \text{ mm} \quad (10.13)$$

The coefficient of the net transparency is

$$\xi = \frac{2D}{L} = 0.0714 \quad (10.14)$$

This coefficient is a very important quantity when the computation of the reactor efficiency is considered. If net transparency is less, the loss of plasma energy is smaller. If the net transparency is high, the required compensation energy may be more than the generated thermonuclear energy. One can decrease the *transparency* by using a conductive material having a high safety tensile stress at high temperature. We can also increase the size L of the net cell. But this method requires some increase in the size of reactor.

The volume density of Bremsstrahlung radiation is:

$$P_B = 5.34 \cdot 10^{-37} n^2 T^{0.5} = 3.78 \cdot 10^2 \text{ W/m}^3 \quad (10.15)$$

Here n enters in m^{-3} while T enters in keV. The total power associated to Bremsstrahlung radiation is

$$P_{B,T} = P_B V = 4.27 \cdot 10^4 \text{ W} \quad (10.16)$$

The surface density of Bremsstrahlung radiation is

$$P_{B,S} = r \cdot P_B / 3 = 3.78 \cdot 10^2 \text{ W/m}^2 \quad (10.17)$$

One assumes the interior of the reactor enveloping the plasma region consists of air at standard conditions. The contact heat transfer between plasma and air is

$$P_c = \frac{\rho}{\rho_a} k_1 T = 2 \times 10^3 \frac{\text{W}}{\text{m}^2}, \quad \text{where} \quad \rho = n \mu m_p, \quad \rho_a = n_a \mu_a m_p \quad (10.18)$$

Here ρ is the density of plasma (kg/m^3); $\rho_a = 1.225 \text{ kg/m}^3$ is standard density of air; $k_1 = 100 \text{ W/(m}^2\text{K)}$ – heat transfer coefficient from air to a solid well (one is right

for T equals some thousands K, but comparing with other estimations shown that may be applied for T millions K); $\mu = m/m_p \approx 30$ for air and $\mu \approx (2 + 5)/2 = 2.5$ for D + T plasma. For our example, $P_c = 2 \times 10^3 \text{ W/m}^2$. The full energy incident per unit net surface is (sum of Bremsstrahlung radiation and plasma contact transfer):

$$q = P_{B,S} / \pi + 2 \cdot P_c = 4 \cdot 10^3 \text{ W/m}^2 \quad (10.19)$$

The temperature of the first net is

$$T = 100 \sqrt[4]{q/C_S} = 517 \text{ K}, \quad \text{where } C_S = 5.67 \text{ J/(m}^2\text{K}^4\text{)}. \quad (10.20)$$

One notices that the temperature of a tungsten fiber in a light lamp is in some times more. But the tungsten fiber into a good conventional light lamp has a life time about 2400 hours. The tungsten net into reactor will have not less a life time. The good design of reactor allows to change the net in 20-40 minutes.

The power is getting out of the reactor through the electric generator (assumed efficiency $\eta = 0.9$):

$$E_e = \eta \cdot V \cdot P_{DT} = 1.24 \cdot 10^6 \text{ W} \quad (10.21)$$

The power lost for heating the three nets is:

$$E_N = 3\xi \cdot VP_{DT} = 2.95 \cdot 10^5 \text{ W} \quad (10.22)$$

The total power lost for supporting the continuous thermonuclear reaction is (compensation of plasma loss for heating nets, fuel and radiation loss):

$$L_s = P_{B,T} + E_N + p_f = 3.38 \cdot 10^5 \text{ W} \quad (10.23)$$

The useful electric output power is:

$$P = E_e - L_s = 902 \text{ kW} \quad (10.24)$$

The mass of the reactor consists of the mass of three nets, M_1 :

$$M_1 = 3\gamma\delta S = 656 \text{ kg} \quad (10.25)$$

(where $\gamma = 19340 \text{ kg/m}^3$ is the density of tungsten) and the mass M_2 of the aluminum cover having $\delta_2 = 0.1 \text{ mm}$, $\gamma_2 = 2700 \text{ kg/m}^3$:

$$M_2 = \gamma_2 \delta_2 S = 36.6 \text{ kg} \quad (10.26)$$

The radius of the cover is about 4.2 m. Here a coefficient 1.2 may be used to increase the mass. Thus, the total mass of the reactor is about 700 – 800 kg.

The output power of the version of the AB Reactor operating in populated regions of Mars is significantly larger:

$$P = (17.5/3.5)P_{DT,F} - p_f = 6.9 \text{ MW} \quad (10.27)$$

In this version there is only the loss for fuel heating. The neutron, Bremsstrahlung and net radiation do not leave the reactor (because of more massive radiation shielding).

About 55 – 60% of this energy may be utilized as electric energy. The rest may be used as heat energy. The mass of this version of the reactor is about 6 – 8 tons.

Figure 10.5 shows the maximal tensile stress of the tungsten versus temperature.

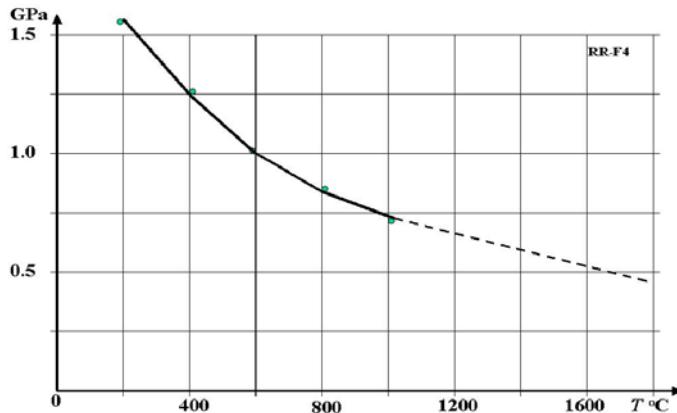


Fig. 10.5 Tensile stress of tungsten versus temperature in °C. Continuous curve is experiment, broken curve is extrapolation to the melting point 3416 °C

Results of computation for various reactor radii and different plasma densities are presented in Figs. 10.6 – 10.8.

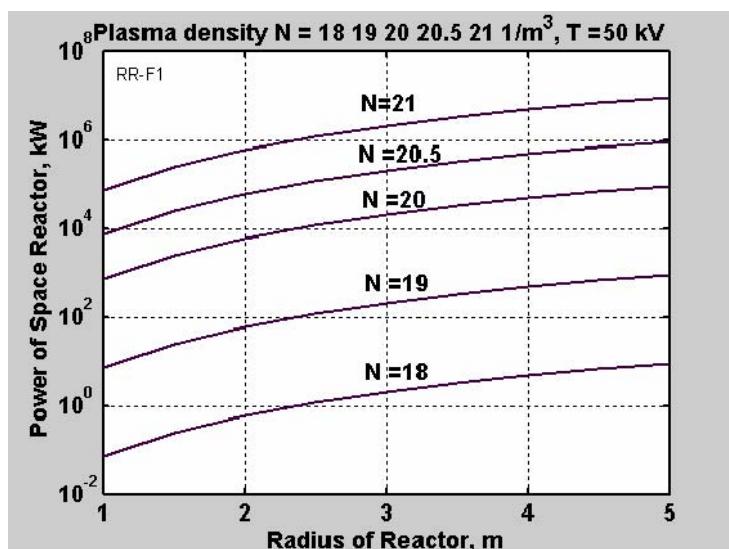


Fig. 10.6 Electric power of thermonuclear reflector AB-reactor (18% from total power) versus radius of reactor and plasma density. Initial data for computation: coefficient of electric generator efficiency is $\eta = 0.9$;

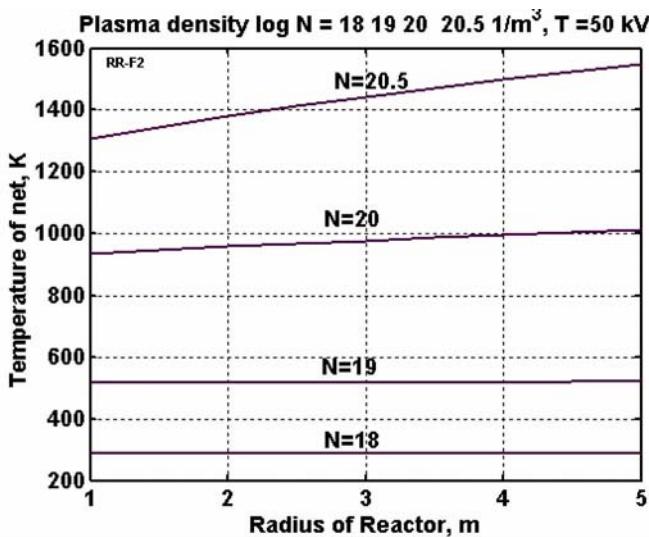


Fig. 10.7 Temperature (K) of the first net versus reactor radius and plasma density. Initial data for computation: plasma temperature 50 keV; net cell size 100 mm, form of net wire is round; specific density of tungsten is 19340 kg/m^3 , safety tensile stress of net for $T = 1500 \text{ K}$ is $\sigma = 2 \times 10^8 \text{ N/m}^2$.

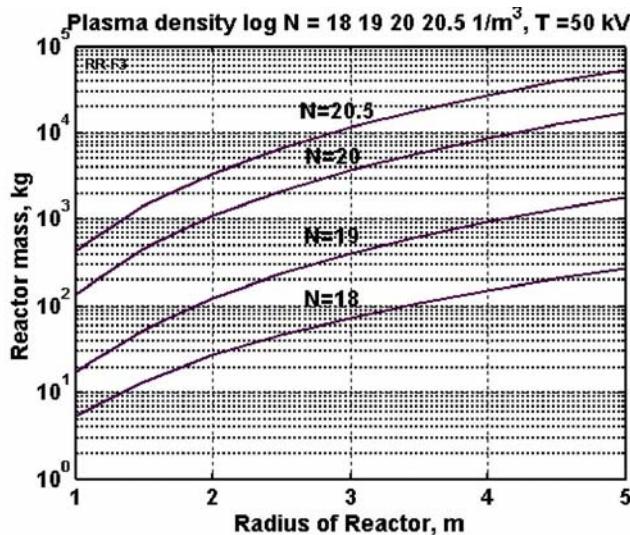


Fig. 10.8 Reactor mass (kg) versus reactor radius and plasma density. Initial data for computation: plasma temperature 50 keV; net cell size $L = 100 \times 100 \text{ mm}$, form wire is round; specific density of tungsten is $\gamma_1 = 19340 \text{ kg/m}^3$, safety tensile stress of net for $T = 1500 \text{ K}$ is $\sigma = 4 \times 10^8 \text{ N/m}^2$; specific density of aluminum is $\gamma_2 = 2700 \text{ kg/m}^3$, thickness of cover is 0.1 mm.

In this section one considered the D+T reaction only. Other reactions (for example, D+D) are needed in special cases.

10.2.4 Application

A sample application of the AB Thermonuclear Reflect Reactor operating in a remote site on Mars is considered now. One uses the following initial data: $r = 3 \text{ m}$, $n = 10^{19} \text{ m}^{-3}$, $T = 50 \text{ keV} = 5.8 \times 10^8 \text{ K}$. One gets as a result useful electric power of 902 kW and a reactor mass of about 400 kg. Full diameter of reactor is about 8.5 m. Distance between nets is about 0.5 m.

Selected reactor parameters are not the best possible neither fully optimized. That is merely an example computation. The usage of materials having more stress capability (safety factor) at high temperatures significantly decreases the loss from the nets and increases the useful electric power. The increasing of plasma density significantly increases the reactor power and decreases the reactor size (diameter) and mass.

The same notes apply to the AB-reactor operating on Mars in populated areas. That is more complex, has more mass, but it produces significantly more energy.

10.2.5 Discussion

The low-density plasma radically decreases the heat flux towards solid surfaces. The high temperature solid surface radiates intensively. Thus, the temperature of the solid conductive material may be lower than its melting temperature. Consequently, the entire magnetic bottle difficulty is avoided.

The confinement nets contact to a rarefied thermonuclear plasma and absorb a part of the plasma energy. This energy compensates (as X-radiation and heating of fuel) an electric energy produced by reactor. This total loss may be 5 – 30% of the electric energy created by the charged α particle. If one compensates the loss of plasma energy by the electric (or high frequency) heating, one can obtain a stable plasma which will permanently produce thermonuclear energy.

The disadvantage of the proposed project is the large size of the reactor (from 1 through 8 m of diameter). That is consequence of using a low-density plasma (low produced energy of rarefied plasma per volume) and required for a given engine power. But, there are a lot of potential reactor users who will be fully satisfied by the useful energy output of 100–2000 kW coming from a functioning reactor that is 1–8 m in diameter.

In the offered project, as in any innovation, the obstacles may appear daunting as, for example, in the heating of the plasma. But there are some not yet fully developed research ideas for solution of even these macro-problems.

10.3 AB-Flight Propulsion on Mars

10.3.1 Short Description of Problem

Most part of this section is based on Bolonkin (2008c).

10.3.1.1 Rockets

A rocket is a vehicle, missile or aircraft which obtains thrust by the reaction to the ejection of fast moving fluid from within a rocket engine. Chemical rockets

operate due to hot exhaust gas made from "propellant" acting against the inside of an expansion nozzle. This generates forces that both accelerate the gas to extremely high speed, as well as, since every action has an equal and opposite reaction, generating a large thrust on the rocket. Rockets are used for fireworks and weaponry, as launch vehicles for artificial satellites, human spaceflight and exploration of other planets. While they are inefficient for low speed use, they are, compared to other propulsion systems, very lightweight, enormously powerful and can achieve extremely high speeds.

Chemical rockets contain a large amount of energy in an easily liberated form, and can be very dangerous, although careful design, testing, construction and use can minimise the risks. A *rocket engine* is a jet engine that takes all its reaction mass (*propellant*) from within tankage and forms it into a high speed jet, thereby obtaining thrust in accordance with Newton's third law. Rocket engines can be used for spacecraft propulsion as well as terrestrial uses, such as missiles. Most rocket engines are internal combustion engines, although non combusting forms also exist.

10.3.1.2 Transfer of Electricity Outside Planetary Atmospheres

The production, storage, and transfer of large amounts of electric energy raise difficult problems for humanity. The energy transfer outside planetary atmospheres is one of these challenging problems. There are no revolutionary ideas in this field except for those in reference (Bolonkin 1983b). Solving the problem of transferring energy outside the planetary atmospheres opens opportunities for new flight methods, including those to be used on planets with rarefied atmospheres, such as Mars. For example, flight-ships can move long distances by using efficient electric engines, orbiting satellites can operate for unlimited time periods without falling prey to orbital decay and premature re-entry to planet atmosphere, communication satellites can transfer a strong signal directly to customers. In the future, Mars outposts can better explore the celestial bodies on which they are placed at considerable expense (Bolonkin 2008a).

Another important problem is the efficient transfer of electric energy for long distances. Nowadays, a lot of loss occurs from such energy transportation. The consumption of electric energy strongly depends on time and season. But an electric station can operate most efficiently in a permanent base-load generation regime. One needs to transfer the energy long distance to any region that requires a supply in any given moment or to special pumped storage stations. One solution is efficient energy storage (Bolonkin 2007a). The storage of a considerable amount of electric energy can help to solve the problem of cheap flight-ship launch. New linear electrostatic engine may be used to accelerate the ship, as suggested in (Bolonkin 2006c). The cheap cable space launch offered by author in reference (Bolonkin 2006b) requires the use of gigantic amounts of energy within a short time period. (It is inevitable for any launch method because we must accelerate big masses to the very high speeds of 8 - 11 km/s). However, in most practical cases it is impossible to use the whole output of a power plant to a single consumer. The offered electric energy storage can help solving this difficult problem (Bolonkin 1982-2008).

10.3.1.3 Railgun

The scientists used a *railgun* for high acceleration of small conducting body. A railgun is a form of gun that converts electrical energy (rather than the more conventional chemical energy from an explosive propellant) into projectile kinetic energy. It is not to be confused with a *coilgun* (*Gauss gun*). Rail guns use magnetic force to drive a projectile. Unlike gas pressure guns, railguns are not limited by the speed of sound in a compressed gas, so they are capable of accelerating projectiles to extremely high speeds (many kilometers per second).

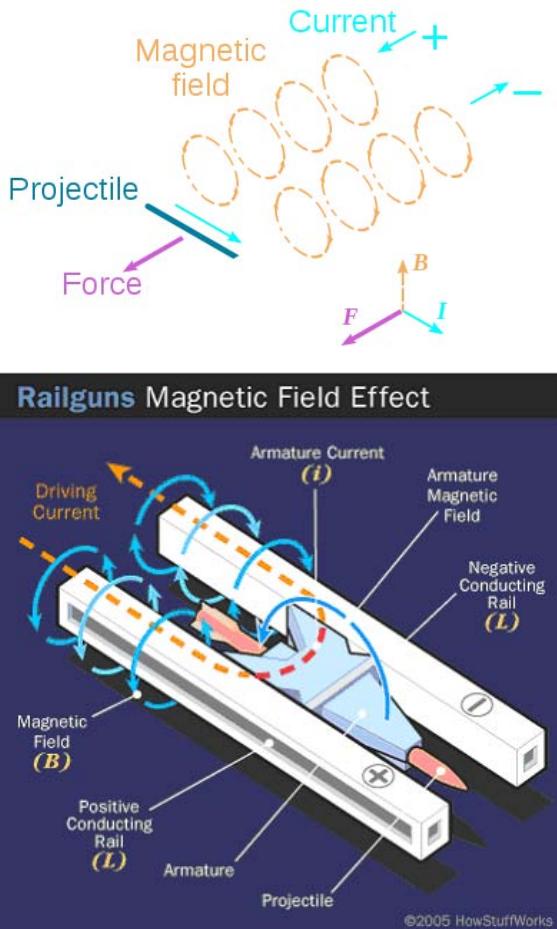


Fig. 10.9 Schematic diagrams of a railgun

A wire carrying an electrical current, when placed in a magnetic field, experiences a force perpendicular to the direction of the current and the direction of the magnetic field. In an electric motor, fixed magnets create a magnetic field,

and a coil of wire is carried upon a shaft that is free to rotate. An electrical current flows through the coil causing it to experience a force due to the magnetic field. The wires of the coil are arranged such that all the forces on the wires make the shaft rotate, and so the motor runs.

A railgun consists of two parallel metal rails (hence the name) connected to an electrical power supply. When a conductive projectile is inserted between the rails (from the end connected to the power supply), it completes the circuit. Electrical current runs from the positive terminal of the power supply up the positive rail, across the projectile, and down the negative rail, back to the power supply (Fig. 10.9). This flow of current makes the railgun act like an electromagnet, creating a powerful magnetic field in the region of the rails up to the position of the projectile. In accordance with the right-hand rule, the created magnetic field circulates around each conductor. Since the current flows in opposite direction along each rail, the net magnetic field between the rails (**B**) is directed vertically. In combination with the electric current (**I**) flowing across the projectile, this produces a Lorentz force which accelerates the projectile along the rails. The projectile slides up the rails away from the end with the power supply.

If a very large power supply providing a million Ampere or so of current is used, then the force on the projectile will be tremendous, and by the time it leaves the ends of the rails it can be travelling at many kilometres per second. 20 kilometers per second has been achieved with small projectiles explosively injected into the railgun. Although these speeds are theoretically possible, the heat generated from the propulsion of the object is enough to rapidly erode the rails. Such a railgun would require frequent replacement of the rails, or use a heat resistant material that would be conductive enough to produce the same effect.

The need for strong conductive materials with which to build the rails and projectiles; the rails need to survive the violence of an accelerating projectile, and heating due to the large currents and friction involved. The force exerted on the rails consists of a recoil force - equal and opposite to the force propelling the projectile, but along the length of the rails (which is their strongest axis) - and a sideways force caused by the rails being pushed by the magnetic field, just as the projectile is. The rails need to survive this without bending, and thus must be very securely mounted.

The power supply must be able to deliver large currents, with both capacitors and compulsators being common. The rails need to withstand enormous repulsive forces during firing, and these forces will tend to push them apart and away from the projectile. As rail/projectile clearances increase, arcing develops, which causes rapid vaporization and extensive damage to the rail surfaces and the insulator surfaces. This limits most research railguns to one shot per service interval.

Massive amounts of heat are created by the electricity flowing through the rails, as well as the friction of the projectile leaving the device. This leads to two main problems: melting of equipment and safety of personnel. As briefly discussed above, the stresses involved in firing this sort of device require an extremely heat-resistant material. The immense heat released in firing a railgun could potentially injure or even kill bystanders. All of these problems can be solved by the invention of an effective cooling method.

Full-scale railgun models have been built and fired, including a very successful 90 mm bore, 9 MJ kinetic energy gun developed by Defense Advanced Research Project Agency (DARPA), but they all suffer from extreme rail damage and need to be serviced after every shot. Rail and insulator ablation issues still need to be addressed. Probably the most successful system was built by the UK's Defence Research Agency at Dundrennan Range in Kirkcudbright, Scotland. This system has now been operational for over 10 years at an associated flight range for internal, intermediate, external and terminal ballistics, and is the holder of several mass and velocity records. At the University of Texas at Austin Institute for Advanced Technology, military railguns capable of delivering tungsten armor piercing bullets with kinetic energies of 9 MJ, have been developed; this is enough energy to deliver 2 kg of projectile at 3 km/s. The United States Naval Surface Warfare Center Dahlgren Division demonstrated an 8 MJ rail gun firing 3.2 kg projectiles in October of 2006 as a prototype of a 64 MJ weapon to be deployed aboard Navy warships. Due to the very high muzzle velocity that can be attained with railguns, there is interest in using them to shoot down high-speed missiles.

10.3.2 Proposed Design Solutions and Brief Description

10.3.2.1 Transfer of Electricity by Plasma Cable

This section proposes a series of innovative solutions for thrust production and transferring energy generated on Mars.

1. Transfer of electrical energy outside Martian atmosphere using a conductive cord from plasma. The author has solved the main problem - how to keep the plasma cord from dissipation, and in compressed form. He has developed the theory of electric transference, made computations that show the possibility of realization for these ideas with existing technology. The electric energy may be transferred from Mars for hundreds millions of kilometers (including towards the Earth) (Bolonkin 2007a).
2. Method of construction of electric lines and electric devices.
3. Method of utilization and tapping of the plasma cable electric energy.
4. Two methods of converting the electric energy to impulse (thrust) motion of a spaceship (these two means are utilization of the magnetic field and of the kinetic energy of ions and electrons of the electric current).
5. Design of a triple electrostatic mirror (plasma reflector), which can reflect the plasma flow (Bolonkin 2007a).

Below are brief descriptions of some projects based on these revolutionary ideas.

10.3.2.2 Transferring Electric Energy in Outside Martian Atmosphere

The electric source (generator, station) is connected to the distant location in space by two artificially generated rarefied plasma cables (Fig. 10.10a). These cables can be created by a plasma beam (Bolonkin 2007a, 2006d) sent from Mars mounted super high tower. If the plasma beam is sent remotely from Mars, a local reflector station is required at the target site or at a third location to turn the circuit back toward its' starting point and closure.

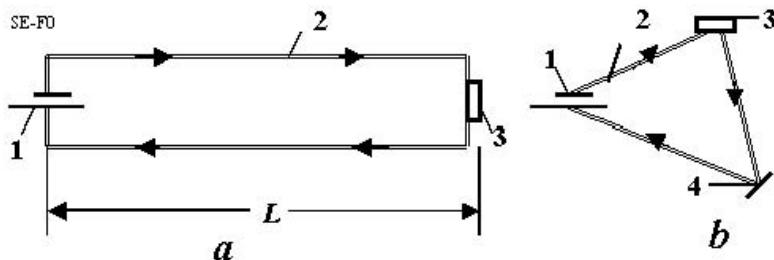


Fig. 10.10 Long distance plasma transfer electric energy and thrust. (a) - plasma transfer with parallel plasma cable, (b) - plasma transfer with triangular (three-wire) plasma cable. *Notations:* 1 - current source (generator), 2 - plasma wire (cable), 3 - spaceship, orbital station or other energy destinations, 4 - plasma reflector located at planet, asteroid or space station.

The plasma cable can be made of ultra-cold (in radial direction) plasma. The plasma cable is self-supported in cable form by the magnetic field created by the electric current going through the plasma cable. The axial electric current produces a contracting magnetic pressure opposed to an expansive gas dynamic plasma pressure (the well-known theta-pinch effect)(Fig. 10.11). The plasma has a good conductivity (equal to that of silver and more) and the plasma cable can have a very big cross-section area (up to thousands of square meters cross-section). The plasma conductivity does not depend on its density. That way the plasma cable has no large resistance although the length of plasma cable may be hundreds of millions of kilometers. The needed minimum electric current is derived from parameters of a plasma cable researched in the theoretical section of this section.

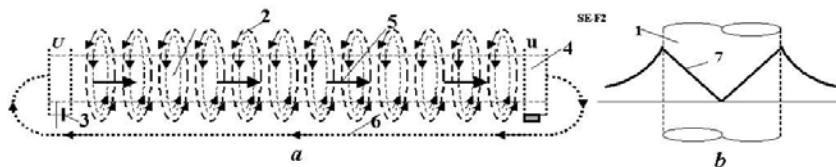


Fig. 10.11 (a) A plasma cable supported by its own magnetic field, (b) Magnetic intensity into and out of plasma cable. *Notations:* 1 - plasma cable, 2 - compressing magnetic field, 3 - electric source, 4 - electric receiver, 5 - electric current, 6 - back plasma line; 7 - magnetic intensity into and out of plasma cable.

The parallel cables having opposed currents repel each other by magnetic force (Fig. 10.10a). This force may be balanced by an attractive electric force if one charges the cables by electric carriers (see theoretical section). They also can be separated by a special plasma reflector as it shown in Fig. 10.10b. The electric line can be created and exist independently. The spaceship connects to this line at a suitable point. By altering the diameter and direction of the plasma cable one can supply energy to a spacecraft. Though one must supply energy to accelerate the spacecraft one can also regenerate energy by braking it. At any time the spaceship

can disconnect from the line and can exist without line support (propulsion, electricity, etc). The apparatus can hook up to or disconnect from the plasma cable at will. But breaking (loss of continuity) of the plasma cable itself destroys the plasma cable line to the remote location. One must have additional (parallel) plasma lines and apparatus must disconnect from a damaged or occulted (for example on the far side of a remote planet) plasma line and connect to another line to keep the connection in existence. The same situation is true in a conventional electric net. The apparatus can also restore the damaged part of plasma line by its' own injected plasma, but the repairs literally take time for the connection to be re-established (tens of minutes or some hours). The original station can also resend the plasma beam, which connects the ends of a damaged part of the line or create a new space electric lines (Bolonkin 2006d).

The electric tension (voltage) in a plasma cable is between the two ends (for example, as cathode- anode) of the conductor in the issuing electric station (electric generator) (Bolonkin 2007a, 2006d,n). The plasma cable current has two flows, i.e. electron (negative) flow and opposed ions (positive) flow within the same cable. These flows create an electric current. (In a metal there are only electron flow, in a liquid electrolytes there are ions flowing). The author offers methods (for extraction and inserting) of energy from the plasma electric cable (Fig. 10.12) by customer (spacecraft, other energy destination or end user). The double net can accelerate the charged particles and insert energy into the plasma cable (Fig. 10.12a) or brake charged particles and extract energy by an electric current from the plasma cable (Fig. 10.12b). In the first case the two nets create the straight electrostatic field, in the second case the two nets create the opposed electrostatic field in plasma cable (resistance in the electric cable (Bolonkin 2007a, 2006d,n) (Figs. 10.10, 10.12c). This apparatus resistance utilizes the electric energy for the spaceship or space station. In the second case the charged particles

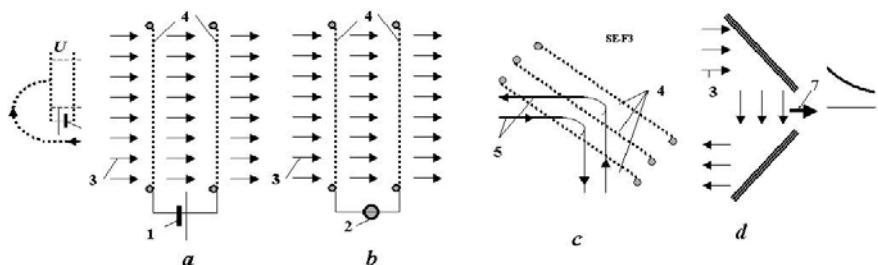


Fig. 10.12 Inserting energy into and (getting energy out of) plasma cable and redirection by plasma reflector of the plasma cable., (a) – inserting electric energy into plasma cable by means of two thin conducting nets or films; (b) - getting the energy from plasma cable by means of two thin conducting nets or films; (c) – offered triple net plasma reflector; (d) – double triple net plasma reflectors - the simplest AB thruster. *Notations:* 1 - spaceship or space station, 2 – receiver of energy, 3 - plasma cable, 4 - electrostatic nets, 5 – two opposed flows of charged plasma particles (negative and positive: electron and ions), 7 – thrust of AB-Space Propulsion.

may be collected into a set of thin films and emit (after utilization in apparatus) back into the continued plasma cable (see (Bolonkin 2007a, 2006d,n)).

Figure 10.13c presents the plasma beam reflector (Bolonkin 2007a, 2006d,n) which has three charged nets. The first and second nets reflect (for example) positive particles, the second and third nets reflect the particles having an opposed (negative) charge. Figure 10.13 shows the different designs of the plasma cable in space.

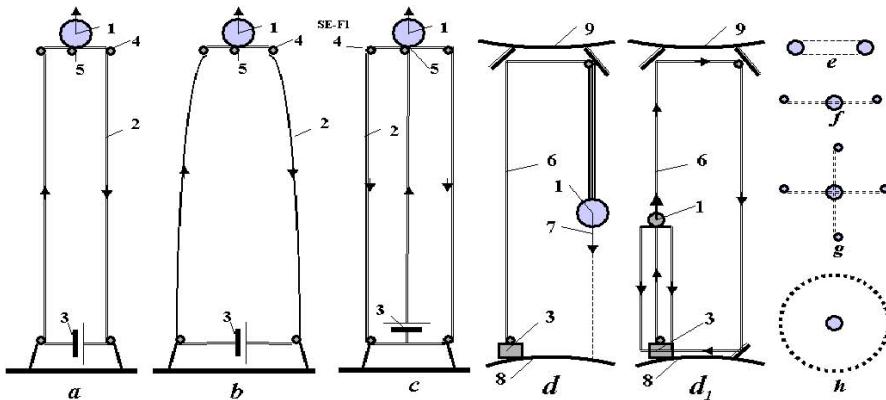


Fig. 10.13 Transfer electricity and thrust by AB-Space Propulsion: (a) Two plasma parallel cables; (b) Curved cable; (c) Plasma multi-cables; (d) Transfer of the back thrust (reaction) into the ‘reaction sink’ of a planet or asteroid; **d₁**. Using of ready plasma line; (e) – (h). Forms of straight and back plasma cables (cross-sections of cables). *Notations:* 1 – Space ship; 2 – plasma cable; 3 – source of electricity; 4 - plasma injector; 5 – user of energy; 6 – double plasma line; 7 - thrust; 8 – Mars; 9 – planet, asteroid, satellite or space ship.

Figure 10.13a shows two plasma parallel cables. Figure 10.13b shows two plasma parallel cables in a curved configuration. Figure 10.13c presents three plasma parallel cables, one to space ship and two for back (return) current. Figure 10.13d shows the transfer of the reverse impulse (or braking) thrust to spaceship through planet or asteroid. Figures 10.13e-h show the different forms of the straight (forward) and back plasma cables (cross-section of cables).

10.3.3 AB-Space Propulsion

The simplest offered AB-Space Propulsion is shown in Fig. 10.12d and more details are shown in Fig. 10.14a. That includes two new triple electrostatic reflectors 2 which turn the plasma cables (flow), (electric current 3) in the return (back) direction. The engine may also contain (optional) the plasma injectors 5 and electric generator (user) 4. Hydrogen gas may be used as feed material for the plasma cable, a ‘plasma reflector’ employing triple conductive nets connected to voltage sources acting to change the direction of the plasma flow (‘reflecting’ plasma flow and redirecting it analogous to a mirror reflecting light). The double conductive

nets may be located in the plasma flow and connected to voltage sources. Voltage applied to these double nets can accelerate the charged particles or brake them. Thus they may increase the speed of a plasma flow (work as generator) or decrease the speed of plasma flow (work to tap that energy for some use).

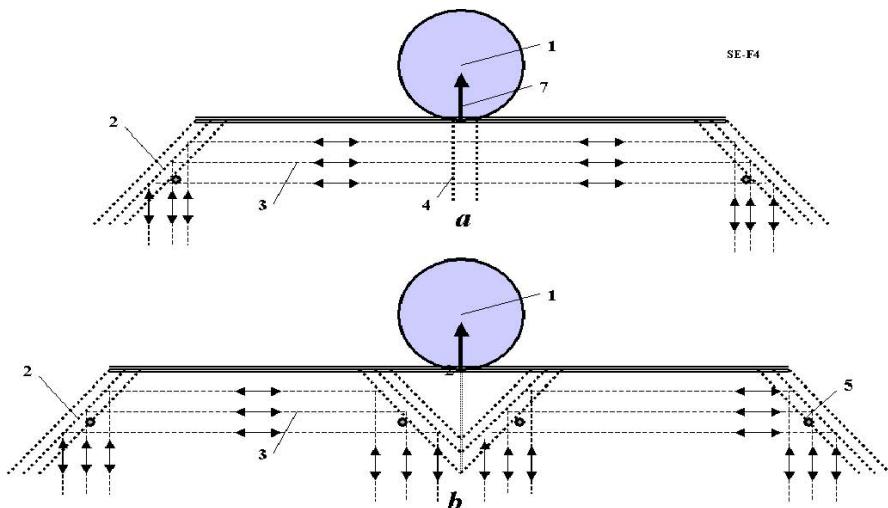


Fig. 10.14 Some versions of AB-Space Propulsion (thruster). (a) two cable AB-Space Propulsion; (b) Three cable AB-Space Propulsion. Notations; 1 – space ship; 2 – offered special (three nets) electrostatic reflector; 3 – plasma cable; 4 – receiver or source of energy; 5 – injector of plasma, 7 – thrust.

The other design of AB-Space Propulsion is shown in Fig. 10.14b. Here the central plasma flow divides in two side flows, which go back to the electric station. The AB-Space Propulsion works as follows. The electric current (voltage) produced by an electric station (which may be located far from the user of AB-Space Propulsion, for example, in orbit around the Mars or mounted on Phobos or another space body) is transferred by plasma cable to the AB-Space Propulsion unit. The power of the electric current in the plasma produces the power via plasma flow of electrons and ions. The engine turns back the plasma flow (electric current) and returns it to the source electric station by the other plasma cable. The magnetic and centrifugal forces appear at the point of turning from outgoing to incoming plasma paths place and create the thrust which can be used for movement (acceleration, braking) the space apparatus (or conventional vehicle or projectile).

The AB-Space Propulsion doesn't violate Newton's third law of action and reaction. The AB-Space Propulsion reacts against the (planet or station mounted) electric station which may be located hundreds of millions of kilometers away from Mars. No other engine has the same capability.

The following differences exist between a railgun and an AB-Space Propulsion device:

1. The railgun uses *solid* physical rails for delivery the electric current to conductivity projectile. These are easily damaged by the required huge electric current. The AB-Space Propulsion uses flexible plasma cables which can self-repair.
2. The railgun uses the rails which are of fixed construction (unalterable) and a spacecraft so launched can move solely in the rail direction. The AB-Space Propulsion creates the plasma cable in the course of apparatus movement and can select and change the apparatus' future direction.
3. The largest (theoretical) railgun may have a length of about 11,000 kilometers (as would any solid construction). The plasma electric line (used by AB-Space Propulsion) may have a length (an acceleration path) of millions of kilometers.

10.3.4 Theory of AB-Propulsion and Results

10.3.4.1 General Theory

The magnetic intensity and magnetic pressure of an electric current reaches a maximum upon the surface of a plasma cable. Let us attempt to equate plasma gas pressure to a magnetic pressure and find the requested equilibrium electric current for a given (same) temperature of electrons and ions

$$\begin{aligned} P_g &= 2nkT_k, \quad P_m = \frac{\mu_0 H^2}{2}, \quad H = \frac{I}{2\pi r}, \\ P_m &= P_g, \quad I = 4\pi r \left(\frac{k n T_r}{\mu_0} \right)^{0.5}, \quad T_k = \frac{m_e u_r^2}{2k}, \end{aligned} \quad (10.28)$$

where P_g is plasma pressure (N/m^2); P_m is magnetic pressure (N/m^2); n is plasma density (number of electron equals number of ions: $n = n_e = n_i$) (m^{-3}), $k = 1.38 \times 10^{-23}$ J/K is Boltzmann constant, $\mu_0 = 4\pi 10^{-7}$ H/m is magnetic constant; H is magnetic intensity (A/m), I is electric current (A), r is radius of plasma cable (m), T_r is plasma temperature in radial direction of plasma cable (K), $m_e = 9.11 \times 10^{-31}$ kg is electron mass, u_r is average electron speed in radial direction of plasma cable (m/s).

Minimum electric current. From Eq. (10.28) one derives a relation between a minimum electric current I_{min} , gas density n and the radial temperature of electrons:

$$\begin{aligned} I_m &= 4\pi r \left(\frac{k n T_r}{\mu_0} \right)^{0.5} \approx 4.16 \times 10^{-8} r \sqrt{n T_r}, \\ j_m &= \frac{I}{\pi r^2} = 4 \left(\frac{k}{\mu_0} \right)^{0.5} \frac{\sqrt{n T_r}}{r} \approx 1.33 \cdot 10^{-8} \frac{\sqrt{n T_r}}{r}, \end{aligned} \quad (10.29)$$

where I_m is minimum electric current intensity (A); j_m is density of electric current (A/m^2), $\pi r^2 = S$ is the cross-section area of plasma cable (m^2).

Assume the temperature (energy) of electrons equals temperature (energy) of ions. Let us write well-known relations

$$j = en(u_i + u_e), \quad \frac{m_i u_i^2}{2} = \frac{m_e u_e^2}{2} \quad (10.30)$$

where $e = 1.6 \times 10^{-19}$ C is charge of electron, $m_e = 9.11 \times 10^{-31}$ kg is the electron mass; m_i is mass of ion (kg) (for H₂ $m_i = 2 \times 1.67 \times 10^{-27}$ kg); u_i, u_e is the speed (m/s) of ions and electrons, respectively, along cable axis produced by electric intensity (electric generator). The computation of j by Eq. (10.29) is presented in Fig. 10.15.

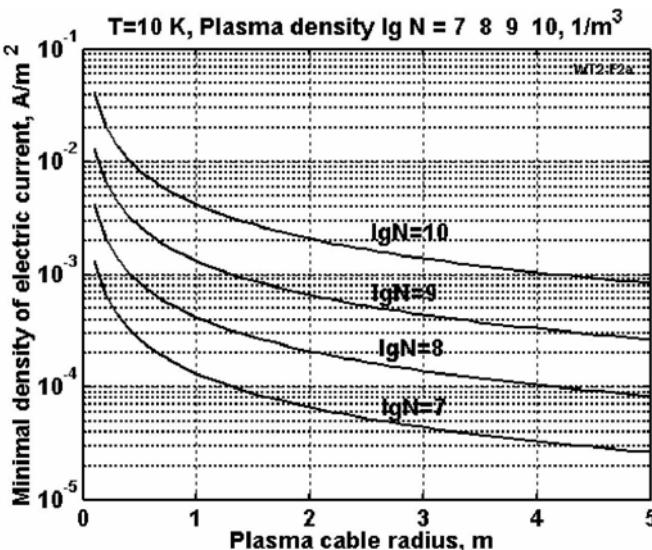


Fig. 10.15 Minimal density of electric current in plasma cable for radial plasma temperature 10 K

From Eq. (10.30) one derives the axial speeds of ions and electrons produced by electric intensity (electric generator)

$$u_e = \frac{j}{en(1 + \sqrt{m_e/m_i})}, \quad u_i = \frac{j}{en(1 + \sqrt{m_i/m_e})} \quad (10.31)$$

or

$$u_e \approx 6.15 \cdot 10^{18} j/n; \quad \text{for H}_2 \quad u_i \approx 10^{17} j/n, \quad (10.31')$$

$$u_e \gg u_i, \quad u = u_e + u_i \approx u_e \quad \text{or} \quad u \approx j/en.$$

Under electric intensity the electrons and ions have opposed speeds along cable axis. The computation of electron speed produced by minimum electric current is presented in Fig. 10.16.

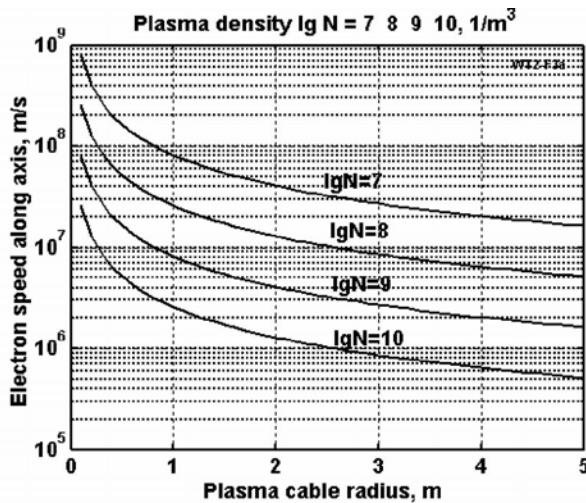


Fig. 10.16 The electron speed produced by the electric current of the minimum current density versus plasma cable radius. The ions (H_2) speed is 61.5 times less than the velocity of the electrons and of the opposite direction of flow.

The temperature along plasma cable axis induces by minimum electric voltage is

$$T_k = \frac{m_e u^2}{2k} \approx 3.3 \cdot 10^{-8} u^2 \quad [K], \quad T = \frac{k}{e} T_k \approx 2.71 \cdot 10^{-12} u^2 \quad [eV] \quad (10.32)$$

where T_k is induced temperature (K); T is this temperature along cable axis (eV). Computations are shown in Fig. 10.17.

Specific Spitzer plasma resistance (the so-called Spitzer Conductivity) and typical resistance of a plasma cable can be computed by:

$$\rho = \eta_{\perp} = 1.03 \times 10^{-4} Z \ln \Lambda T^{-3/2} \quad \Omega \cdot m, \quad R = \rho L / S, \quad (10.33)$$

where ρ is specific plasma resistance ($\Omega \cdot m$); Z is ion charge state, $\ln \Lambda \approx 5 \div 15 \approx 10$ is the Coulomb logarithm; $T = T_k k/e = 0.87 \times 10^{-4}$, T_k is plasma temperature along cable axis (eV); $e = 1.6 \times 10^{-19}$ C is electron charge; R is electric resistance of plasma cable (Ω); L is plasma cable length (m); S is the cross-section area of the plasma cable (m^2). The computation of the specific resistance of a plasma cable for minimum electric current is presented in Fig. 10.18.

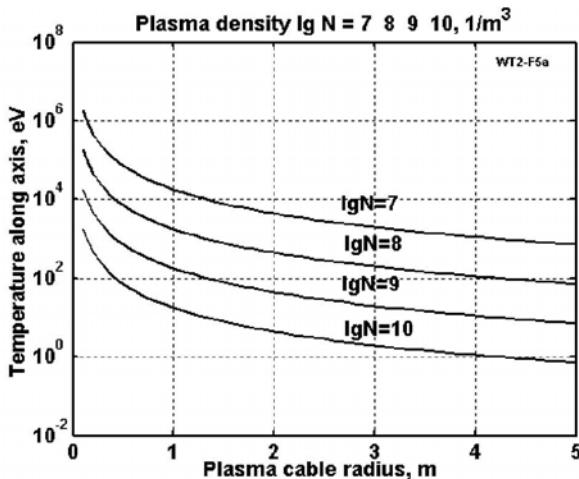


Fig. 10.17 The temperature of electron and ions (H_2) (in eV) produced by the electric current in the minimum current density versus the plasma cable radius and different plasma density. It is assumed that the ions (H_2) temperature equals the electron temperature.

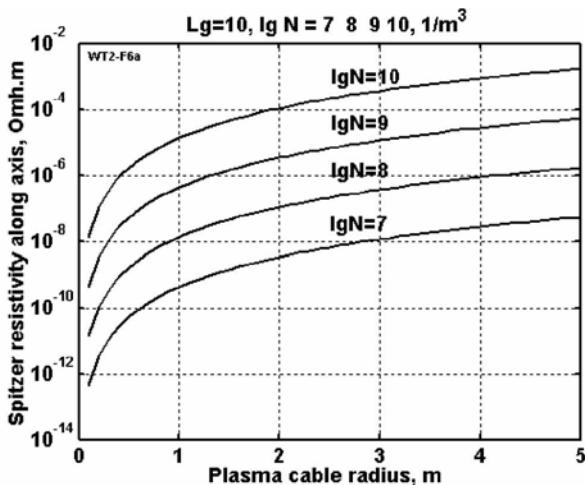


Fig. 10.18 Specific (Spitzer) plasma resistance $\Omega \cdot m$ of equilibrium plasma cable for the minimum electric current versus cable radius and different plasma density. Coulomb logarithm equals 10.

The requested minimum voltage, power, transmitter power and electric efficiency are:

$$\begin{aligned} U_m &= IR, \quad W_m = IU_m, \quad U = U_m + \Delta U, \\ W &= IU, \quad \eta = 1 - W_m / W = 1 - U_m / U \end{aligned} \quad (10.34)$$

where U_m , W_m are requested minimum voltage (V), and power (W), respectively; U is used voltage (V); ΔU is electric voltage over minimum voltage (V); W is used electric power (W); η is the efficiency of the electric line. If $\Delta U \gg U_m$, the efficiency is close to 1. Computation of lost voltage and power into plasma cable having length of 100 million km is in Figs. 10.19-10.21.

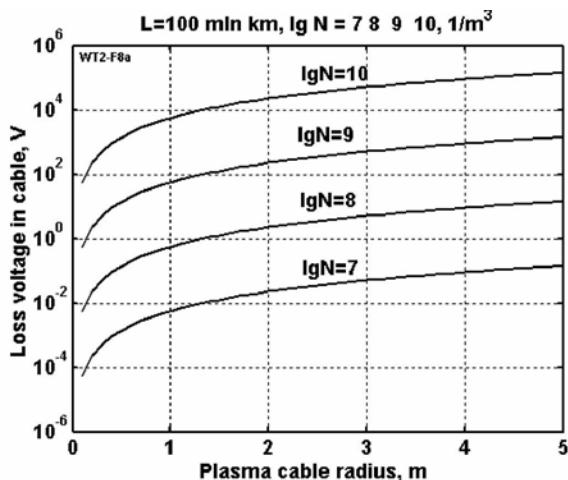


Fig. 10.19 Lost voltage in plasma cable of 100 millions km length via cable radius for the minimum electric current and different plasma densities

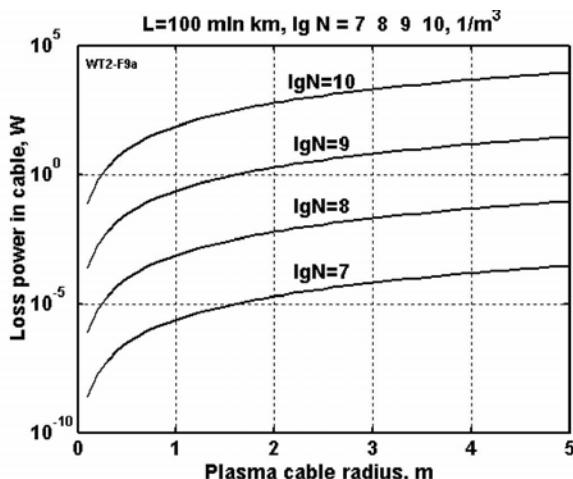


Fig. 10.20 Lost power in plasma cable of 100 millions km length via cable radius for the minimum electric current and different plasma densities

The mass M (kg) of one plasma cable is

$$M = \pi r^2 n m_i L, \quad (10.35)$$

where m_i is ion mass of plasma (kg); L is length of plasma cable (m). The mass of plasma cable is very small, a few grams for 100 millions km length, for $n < 10^{10}$ particles/m³. The mass of a plasma cable is close to zero for any practical case when $R < 5$ m.

The force acting on one particle (proton) moved in electric and magnetic fields may be computed by the equations:

$$\bar{F}_1 = \frac{m_i v^2}{r}, \quad \bar{F}_2 = e \bar{v} \bar{B}, \quad \bar{F}_3 = \frac{e Q_0}{4\pi \epsilon_0 R^2}, \quad \bar{F}_4 = \gamma \frac{m_1 m_2}{R^2}, \quad \bar{F}_4 = g m_i \quad (10.36)$$

where F_1, F_2, F_3, F_4 are centrifugal, Lorenz, electrostatic, and gravitational forces, respectively (all vectors) (N); $m_p = 1.67 \times 10^{-27}$ kg is the mass of proton (or ion m_i); v - speed of particle (m/s); e - electron (proton) charge; B - total magnetic induction (magnetic field strength) (T); Q_0 - charge of central body, C; $\epsilon_0 = 8.85 \times 10^{-12}$ F/m - electric constant; m_1, m_2 are mass of bodies (central and particle) (kg); γ - gravitational constant, (for Mars $g = 3.72$ m/s²); r - radius curve (m); R - distance between charges (gravitational bodies) (m).

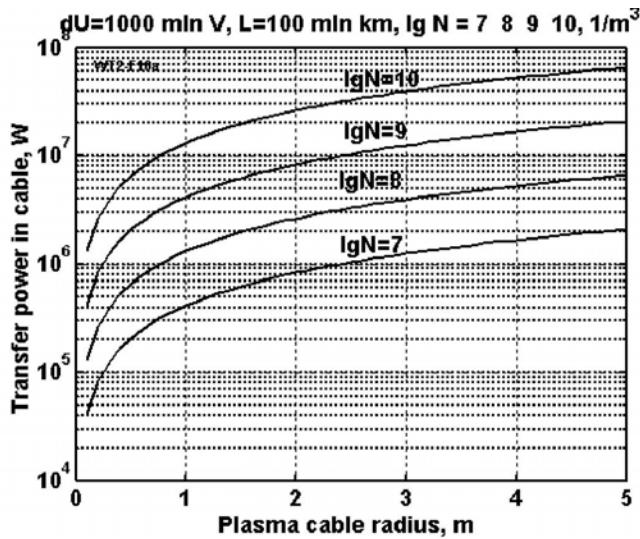


Fig. 10.21 Electric power transfers by plasma cable of 100 millions km length via cable radius for the minimum electric current, over voltage 10^9 V and different plasma densities. The efficiency is about 0.9999.

The mechanical equilibrium condition is:

$$\sum_i F_i = 0$$

Magnetic Pressure (magnetic thrust) from the Plasma Cable. The plasma exerts a pressure within the plasma cable. This pressure is small, but the cable can have a large diameter (up to 200 m or more) and this pressure acting over a long time can accelerate or brake a space apparatus with no reaction mass. This magnetic pressure P (N/m^2) from only one cable can be computed by equations:

$$P_m = \frac{\mu_0 H^2}{2}, \quad H = \frac{I}{2\pi r}, \quad P = \frac{1}{2} 2P_m S = \frac{\mu_0}{4\pi} I^2 \quad (10.37)$$

Estimation. For $I = 10^4$ A, the magnetic pressure equals 10 N; for $I = 10^5$ A, it equals 1000 N; for $I = 10^6$ A, the thrust of one cable is $P = 10^5$ N. That is the magnetic thrust from one cable. The AB-Space Propulsion has two cables (incoming and out coming), that means the magnetic thrust from two cable will be (at minimum) two times more. If one computes the horizontal part of plasma cable, which is pressed by outer plasma magnetic field, the full thrust is:

$$P = \frac{\mu_0}{2\pi} I^2 \ln \left| \left(\frac{d - r}{r} \right)^2 \right| \quad (10.38)$$

where d is the distance between the centers of the incoming and outcoming plasma cables.

Electric (kinetic) Pressure from the Plasma Cable. The high speed electrons and ions of electric current within plasma cable have kinetic energy. This energy produces kinetic (electric) pressure when the spaceship or final station uses the electric energy. Let us estimate the electric pressure. The specific (kinetic) energy of electric current into plasma cable is

$$E = 0.5n(m_e u_e^2 + m_i u_i^2) [\text{W/m}^3]. \quad (10.39)$$

Substituting Eqs. (10.31) in (10.39) one finds

$$\begin{aligned} E &= \frac{j^2}{2ne^2} \left[\frac{m_e}{(1 + \sqrt{m_e/m_i})^2} + \frac{m_i}{(1 + \sqrt{m_i/m_e})^2} \right] \\ &\approx \frac{j^2}{2ne^2} [m_e + m_i] = \frac{m_e}{e^2} \frac{j^2}{n} \end{aligned} \quad (10.40)$$

But the energy per unit volume (units: $\text{Nm/m}^3 = \text{N/m}^2$) equals the pressure $P_e = E_s$ (N/m^2) and

$$P_e = \frac{m_e}{e^2} \frac{j^2}{n} \approx 3.36 \times 10^7 \frac{j^2}{n} \quad (10.41)$$

Estimation. For $j = 100 \text{ A/m}^2$ and $n = 10^{10} \text{ m}^{-3}$ one gets $P_e = 35.6 \text{ N/m}^2$. The total kinetic energy of charged particles the plasma cable is $E = P_e s L$ (J), where s is the cross-section area of plasma cable (m^2) and L is length of plasma cable (m).

Additional Power from a Space Apparatus Motion. This power is:

$$W = PV \quad (10.42)$$

where V is apparatus speed (m/s). *Estimation.* For $V = 11 \text{ km/s}$, $P = 10^{-3} \text{ N}$, this power equals 11 W, for $P = 1 \text{ N}$ the power equals 11000 W. One spends this power when the space apparatus moves away from the energy source ('launch point') and one receives it when apparatus approaches to the energy station ('landing site').

Track Length of Plasma Electrons and Ions. The track length L and the track time τ of particles is

$$L = v_T / \nu, \quad \tau = 1/\nu \quad (10.43)$$

where v_T is particle velocity (cm/s); ν is particle collision rate (s^{-1}). The electron, ion, and electron-ion *thermal* collision rate are, respectively:

$$\nu_e = 2.91 \times 10^{-6} n_e \ln \Lambda T_e^{-3/2} \quad (\text{s}^{-1}) \quad (10.44)$$

$$\nu_i = 4.80 \times 10^{-8} Z^4 \mu^{-1/2} n_i \ln \Lambda T_i^{-3/2} \quad (\text{s}^{-1})$$

$$\nu_{ei} = 4.4 \times 10^{-6} n_i \lg \Lambda T^{-3/2}.$$

where Z is ion charge state, $\ln \Lambda \approx 5 \div 15 \approx 10$ is Coulomb logarithm, $\mu = m_i/m_p$ is relative mass of ion; $m_p = 1.67 \times 10^{-27}$ is mass of proton (kg); n is density of electrons and ions respectively (cm^{-3}); T is temperature of electron and ion, respectively (eV). Electron and ion *thermal* velocity are respectively:

$$v_{Te} = (kT_e / m_e)^{1/2} = 4.19 \times 10^7 T_e^{1/2} \quad (\text{cm/s}) \quad (10.45)$$

$$v_{Ti} = (kT_i / m_i)^{1/2} = 9.79 \times 10^7 \mu^{-1/2} T_i^{1/2} \quad (\text{cm/s})$$

Substitute Eqs. (10.39)-(10.40) in Eq. (10.38) yields the length of electron and ion tracks:

$$L_e = 1.44 \times 10^{13} T_e^2 / n_e \ln \Lambda \quad (\text{cm}), \quad (10.46)$$

$$L_i = 2.04 \times 10^{13} T_e^2 / Z^4 n_e \ln \Lambda \quad (\text{cm}),$$

$$L_{ei} = 0.95 \times 10^{13} T_e^2 / n_e \ln \Lambda \quad (\text{cm}).$$

Estimation. For electron having $n = 10^5 \text{ cm}^{-3}$, $T = 100 \text{ eV}$, $\ln \Lambda \approx 10$ one gets $L = 2 \times 10^6 \text{ km}$, $\tau \approx 300 \text{ s}$. That means the plasma electrons have very few collisions, small dispersion, (in present case) and they have different average *electron* (relative to ion) temperature, along the cable axis and perpendicular to cable axis. It is not a surprise because the plasma can have different average temperatures of electron and ions. That also means that the assumption about the terminal and current

electron velocities being the same is very limited and the parameters of a plasma electric system might be more advantageous than in present computation. The *plasma may be very cool in a radial direction and simultaneously very hot in the axial direction*. That decreases the electric current needed for plasma compression and allows a transfer of the plasma beam, energy, and thrust to a great distance.

Magnetic force between two parallel cables. This force is

$$F_m = -\mu_0 \frac{i_1 i_2}{2\pi d} L \quad (10.47)$$

where F_m is magnetic force (N) (the force is repeal when currents are opposed, and attractive when currents have same directions); $\mu_0 = 4\pi 10^{-7}$ is permeability constant (H/m); i is electric current in the 1-st and 2-nd cable, respectively (A); d is distance between center of cables (m); L is length of cables (m). This force for two cable line (Fig. 10.13e) having current $I = 10^5$ A and distance $d = 1000$ km equals $F_m = 2$ N/km. The force decreases if one uses multi-cable system: for three cables (Fig. 10.13f) it is $3/8 F_m$; for 5 cables (Fig. 10.13g) it is $5/32 F_m$; for multi-cables (Fig. 10.13h), $F_m = 0$.

Electrostatic force between two parallel cables. This force is

$$F_e = k \frac{2\tau_1 \tau_2}{d} L \quad (10.48)$$

where F_e is electrostatic force (N) (the force is attractive when charges is different and repeat when charges are same); $k = 1/4\pi\epsilon_0 = 9 \times 10^9$ electrostatic constant (Nm^2/C^2); τ is linear charge of the 1-st and 2-nd cable (C/m); d is distance between cables (m); L is length of cables (m). Electrostatic force is an attractive force for opposed charges. This force may be used for balancing the electromagnetic force. From $F_m = F_e$, and using Eqs. (10.47) and (10.48) one gets for two line cable system:

$$\tau = \frac{1}{2} \sqrt{\frac{\mu_0}{\pi k}} I, \quad \Delta U = 2k \tau \int_r^d \frac{dR}{R} = \sqrt{\frac{\mu_0 k}{\pi}} I \ln\left(\frac{d}{r}\right) = 60I \ln\left(\frac{d}{r}\right), \quad (10.49)$$

$$\Delta P = \Delta U \cdot I$$

where r is plasma cable radius (m). Example: for $I = 10^4$ A, $d/r = 10$ one have $\Delta U = 1.38 \times 10^6$ V. The linear charge appears on a cable when there is voltage between cables. The other way of balance is the cable design in Fig. 10.13h.

Electric capacity of two parallel cables is

$$C = \frac{\pi \epsilon_0}{\ln(d/r)} L \quad (10.50)$$

where C is electric capacity (F); $\epsilon_0 = 8.85 \times 10^{-12}$ is electrostatic constant, (F/m); r is cable radius (m);

Energy of two parallel cables as electric condenser is

$$E = \frac{1}{2} C U^2 = \frac{1}{2} q U = \frac{1}{2} \frac{q^2}{C} \quad (10.50')$$

where E is energy in condenser (J); U is electric voltage (V); q is electric charge (C). Example: for $d = 100$ km, $r = 10$ m, the electric capacity is $C = 0.05$ F/one million km, the energy is $E = 2.5 \times 10^6$ J/one million km.

Inductance of two parallel cables is

$$L_i = \frac{\mu_0}{\pi} \left(\frac{1}{2} + \ln \frac{d}{r} \right) L \quad (10.51)$$

where L_i is inductance, H.

Inductance energy of two parallel cables is

$$E = L_i \frac{I^2}{2} \quad (10.52)$$

where E is energy in a closed-loop contour (J). Example: for $d = 100$ km, $r = 10$ m, the inductance is $L_i = 3.9 \times 10^3$ H/one million km, the energy is $E = 1.94 \times 10^{15}$ J/one million km. This energy is high and the starting station (where the plasma cables originate) spends a lot of energy for creating the magnetic field.

Change electric current in closed-loop contour is

$$I = I_0 \exp(-t/T), \quad \text{where } T = \frac{L_i}{R} \quad (10.53)$$

where R is electric resistance of closed-loop electric contour (Ohm), T is time decreasing current by factor $e = 2.71$ times. Example: for two lines cable, the length $L = 100$ millions of km and the electric resistance $R = 10^3$ Ohm (see project below), $T \approx 3.9 \times 10^6$ sec = 45 days. That means our electric line is a large storage reservoir of energy.

'Virtual' Specific Impulse of AB-Space Propulsion. The specific impulse of rocket engine is the ratio of an engine thrust to fuel consumption per second. It is difficult to speak about specific impulse of AB-Propulsion because AB-Propulsion doesn't spend fuel to generate the thrust, but it does expend matter (for example hydrogen) for creating the plasma cables. That way one can take the ratio of the thrust to mass expenditure per second for produce of new cable. This 'virtual' specific impulse of the AB-Space Propulsion is

$$I_y = \frac{P}{m_s}, \quad P = \frac{\mu_0}{2\pi} I^2, \quad m_s = 2snm_i V, \quad I_y = \frac{\mu_0}{3\pi m_i} \frac{I^2}{snV}. \quad (10.54)$$

$$\text{For } H_2 \quad I_y = 6 \cdot 10^{19} \frac{I^2}{snV}$$

where I_y is specific impulse of AB-Space Propulsion (m/s); P is magnetic thrust (N); $s = \pi r^2$ is cross section area of plasma cable (m^2); $m_i = 2 \times 1.67 \times 10^{-27}$ is mass of one molecule of hydrogen (kg); V is speed of apparatus (m/s). *Estimation.* Let

us take $I = 10^6$ A, $s = 10^2$ m², $n = 10^{14}$ /m³, $V = 10^4$ m/s. We get $I_y = 6 \times 10^{11}$ m/s. That is a very large specific impulse. No present *rocket engine* has such a specific impulse and a rival is unlikely in the future. For comparison, the specific impulse are: conventional liquid-propellant rocket engine had maximum $I_y = 4200$ m/s; hydrogen rocket engine - $I_y = 5180$ m/s; thermonuclear rocket engine (H₂+H₃)(for 100% efficiency) has $I_y = 26 \times 10^6$ m/s; ideal laser engine has $I_y = 3 \times 10^8$ m/s; and the most power – annihilation rocket engine (for 100% efficiency) has theoretical impulse $I_y = 4.24 \times 10^8$ m/s.

The AB-Space Propulsion has very high specific impulse. This system spends less mass for producing the plasma cable than any rocket engine spends for producing thrust. Another advantage is that it gets the energy from a planet-mounted station, i.e. the power source needn't travel with it and weigh it down. The AB-Space Propulsion is very light, simple, safety, and reliable with comparison to any likely nuclear engine. In most cases at least part of the cable mass can be injected from the planet-mounted energy station.

Efficiency of AB-Space Propulsion. The following definition is proposed for the efficiency of AB-Space Propulsions: the ratio of energy (power) getting by apparatus to energy (power) spending by station:

$$\eta = \frac{PV}{N} \quad (10.55)$$

where η is the efficiency; P is full thrust getting by apparatus (N); V is apparatus speed (m/s); N is electric station power (W). The formulas above allow us to compute it, but it is a variable value.

Another efficiency is the ratio of the apparatus thrust to the spending power of electric station (N/W):

$$\eta_N = \frac{P}{N}, \quad P = \frac{\mu_0}{2\pi} I^2, \quad N = I^2 R, \quad R = \rho \frac{L}{s}, \quad (10.56)$$

$$\rho = 1.03 \cdot 10^{-4} \ln \Lambda \cdot T^{-3/2}, \quad T = 2.71 \cdot 10^{-12} \left(\frac{j}{en} \right)^2$$

Substitute all Eqs. (10.56) into the first Eq. (10.56) one gets

$$\eta_N = 1.06 \cdot 10^{36} \frac{s}{L \cdot \ln \Lambda} \left(\frac{j}{n} \right)^3 \quad (10.57)$$

Estimation. For $n = 10^{14}$ /m³, $j = 10^3$ A/m², $s = 10^2$ m², $L = 10^5$ m, $\ln \Lambda = 10$ one finds $\eta_N = 0.1$ N/W. However, for high L , the efficiency is very small. That is because the electric station spends a lot of energy for producing magnetic field of closed-loop cables. The AB space electric system is a very large storage of electric energy and part of this energy will get back when one decreases (spaceship returns back) or one removes the AB space electric line.

Suitability of beam to penetrate Mars Atmosphere. A spaceport may be placed on the highest point on Mars surface (i.e. Olympus Mons), where the atmospheric

pressure is obviously lower than the average value on the Mars surface (6-9 hPa). Indeed, at the top of Olympus Mons the ambient pressure is about 1 hPa. This is around 0.001 of the Earth sea level atmospheric pressure and use of plasma cable propulsion should be possible.

Computational example. To give perspective for the present project a few numerical results are presented now. They will prove the feasibility of a terrestrial application. Of course, the results are covering the case of Martian applications, because the atmosphere is more rarefied on Mars than on Earth.

The energy loss of the beam in the Earth atmosphere may be estimated in the following way:

$$\tau = \frac{100H_0\rho_0\bar{\rho}(h)\bar{p}(h)}{R_t(U)}, \quad R_t = \frac{m}{m_p} R_t \left(\frac{m_p}{m} U \right) \quad (10.58)$$

where $H_0 = P_a/\rho = 10^4/1.225 = 8163$ m is thickness (height) of Earth atmosphere having constant density $\rho = 1.225 \text{ kg/m}^3$, $P_a = 10^5 \text{ N/m}^2$ is the atmospheric pressure; $\bar{\rho}(h)$ is relative atmosphere density; $\bar{p}(h)$ is relative atmosphere pressure; R_t is particle track in atmosphere (cm); m is mass of particle (kg); h is altitude (m); U is beam energy (MeV); $\rho_0 = 0.001225 \text{ g/cm}^3$ is atmosphere density; 100 is conversion coefficient meter into cm. Values of R_t are given in Table 10.3 while values of $\bar{\rho}(h)$ and $\bar{p}(h)$ for protons are given in Table 10.4.

Table 10.3 Value of R_t (g/cm^2) versus energy of proton (Kikoin 1976)

U (MeV)	R_t (g/cm^2)	U (MeV)	R_t (g/cm^2)
0.1	1×10^{-4}	600	176
1	1.09×10^{-2}	700	222
10	0.99×10^{-1}	800	270
50	2.56	1000	370
100	8.835	2000	910
200	29.64	3000	1363
300	58.08	5000	2543
400	93.73	7000	3583
500	133.3	10000	5081

Table 10.4 Standard Earth atmosphere

h (km)	$\bar{\rho}(h)$	$\bar{p}(h)$
0	1	1
5	0.661	0.533
10	0.338	0.261
20	0.072	0.054
40	3.27×10^{-3}	2.92×10^{-3}
60	2.71×10^{-4}	8.35×10^{-4}
100	4.41×10^{-7}	3.20×10^{-7}

Results of computation using Eq. (10.58) are shown in Figs. 10.22 and 10.23.

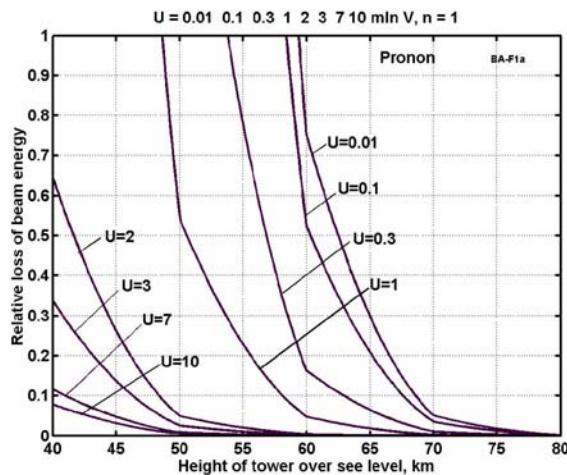


Fig. 10.22 Relative energy loss of the proton particle beam as a function of tower altitude in Earth atmosphere. Acceleration voltage U is in MV. Curves with abrupt slope changes resulted from the linearization data of Tables 10.3 and 10.4.

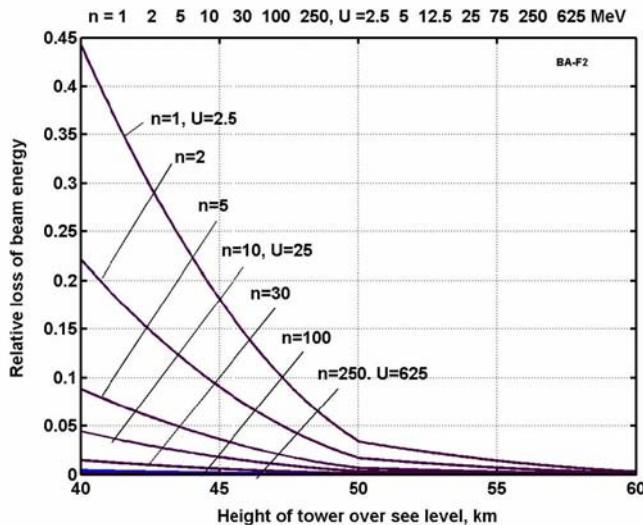


Fig. 10.23 Relative energy loss of the particle beam as a function of tower altitude in Earth atmosphere and number n of nucleons in nucleus. Acceleration voltage U in MV.

Obviously, high energy particle beams break up in the Earth atmosphere at lower altitudes, where atmospheric density is larger. A better situation happens in case the particle beam is generated from a rarefied atmosphere, for instance from the top of space tower of 40 ÷ 80 km height. The results in Figs. 10.22 and 10.23 corresponding to a terrestrial tower of 50 km height may be associated to a spaceport place at the top of Olympus Mons on Mars.

The last formula in Eq. (10.58) allows recalculation by the particle track for any atom. For example, one wants to calculate the particle track for oxidizer particle with $m = 16m_p$ and energy 8,000 MeV. One takes the value of R_t from Table 10.4 for $U = 8,000/16 = 500$ MeV and one multiplies by 16. The result is $R_t = 133 \times 16 = 2128$. The particle track $T_r = R_t/\rho_0 = 17.4$ km in the air having density 1.225 kg/m^3 . That is enough to break the equivalent (constant density) Earth's atmosphere of 8.163 km thickness, but the loss of energy is $\tau = 8.163/17.4 = 0.47$. Of course, the energy loss may be diminished in case the beam station is located on a mountain or on a tower with height about 40 ÷ 60 km (for Earth) or on high mountains on Mars.

10.3.5 Project

As a example, one estimates the parameters of the AB-Space Propulsion having the thrust from one plasma cable with $P = 10^5 \text{ N}$. The next design is not optimal. Most of the computations are made for a single cable. For a real (two cable) engine all values should be doubled. Let us take the following initial data: electric current $I = 10^6 \text{ A}$, thrust of one cable $P = 10^5 \text{ N}$, radius of plasma cable $r = 10 \text{ m}$, plasma density $n = 10^{14} \text{ particles/m}^3$. $I > I_{\min} = 13 \text{ A}$. One gets the following results. Density of electric current and electron speed are:

$$j = \frac{I}{\pi r^2} = 3180 \text{ A/m}^2, \quad u \approx \frac{j}{en} = 2 \cdot 10^8 \text{ m/s},$$

The temperature electrons and ions along cable axis and Spitzer electric resistance (for $\ln \Lambda = 10$) are

$$T = 2.71 \cdot 10^{-12} u^2 = 1.08 \cdot 10^5 \text{ eV}, \quad \rho = 1.03 \cdot 10^{-4} \ln \Lambda \cdot T^{-3/2} = 0.92 \cdot 10^{-10.5} \Omega \cdot \text{m}$$

The electric resistance, requested voltage and electric power for one plasma cable and its length $L = 10^{12} \text{ m} = 1000$ millions of km (one reminds that the minimum distance from Mars to Earth is about 78 millions km):

$$R = \rho \frac{L}{s} = 9.5 \cdot 10^{-2} \Omega, \quad U = IR = 9.5 \cdot 10^4 \text{ V}, \quad N = IU = 9.5 \cdot 10^{10} \text{ W}.$$

The resulted power (950 GW) is about 20% of the peak power of all kinds of energy produced in the USA in 2005. If the thrust is decreased to 1,000 – 10,000 N, the requested electric power decreases by hundreds- to thousands of times. The apparatus can reach a needed speed by increasing the acceleration distance.

The mass of one cable is

$$M = sLnm_i = 105 \text{ kg}$$

The mass of two (forward and backward) cables is 210 kg. This is a small mass, for a cable having cross-section $s = 314 \text{ m}^2$ (diameter 10 m) and length $L = 1000$ millions of km. The mass may be significantly smaller if one uses plasma of smaller density. Part of this mass (about half) may be ejected from the start electric station.

The magnetic thrust of cables (Fig. 10.13g,h) at distance $d = 110 \text{ m}$ near the apparatus is

$$P = \frac{\mu_0 I^2}{2\pi} \ln \left| \left(\frac{d-r}{r} \right)^2 \right| = 9.2 \cdot 10^5 \text{ N}$$

That thrust is different from the initial thrust of the two single plasma cables (20,000 N). The kinetic thrust of charged particles for one plasma cable is

$$P_e = 3.36 \cdot 10^7 \frac{j^2}{n} s = 1.07 \cdot 10^3 \text{ N}$$

As one can see, the kinetic (electric) thrust is small (in comparison with magnetic thrust). One can neglect it. The total kinetic energy (charged particles) for one cable of length $L = 1000$ millions km is

$$E = P_e L = 1.07 \cdot 10^{15} \text{ J}$$

The additional voltage between the two cable line, for balance magnetic (repeal) and electric (attracted) forces, and the power which must be used for the spaceship or electric station are

$$\Delta U = 60 \cdot I \cdot \ln \left(\frac{d}{r} \right) = 1.38 \cdot 10^8 \text{ V}, \quad \Delta P = \Delta U \cdot I = 1.38 \cdot 10^{14} \text{ W}$$

That is a large power. It might be necessary to decrease the current and thrust or use the fig.13h design.

Estimation of flight possibilities for a spaceship having mass $M_s = 92$ tons and flight time 10 days $= t = 8.64 \times 10^5$ sec follows. The acceleration of the spaceship, speed and distance traveled are:

$$a = \frac{P}{M_s} = 10 \text{ m/s.} \quad V = at = 8.64 \cdot 10^3 \text{ km/s,} \quad L_s = \frac{at^2}{2} = 3.73 \cdot 10^{12} \text{ m.}$$

This is in 50 times more than minimal distance from Mars to Earth. Note that the value of the speed V is 2.88% of the light speed.

10.3.6 Discussion

Advantages of AB-Space Propulsion are:

- It is very light, simple, safe, and reliable in comparison with any likely nuclear engine.
- It has a gigantic ‘virtual specific impulse’, being more capable of realistic operation in a near-future environment, than virtually any proposed means of thermonuclear or light-propulsion scheme the author is aware of.
- It can accelerate a near-term space apparatus to very high speed (approaching light speed). At present time this is the single real method to be able to approach this ‘ultimate’ velocity.
- At least part of the needed injected plasma cable mass and nearly all of the energy needed (and the cooling facilities needed to maintain that energy supply) can be obtained from the planet-bound energy-supplying station, further improving the on-board ship ‘mass ratio’.
- It can use far cheaper energy from a planet-bound electric station.

Further research and testing are necessary. These tests are not expensive. For example, the plasma cable may exhibit new forms of instability. There is historical precedent for this: The instability of a plasma cable was found in tokomak R&D, but was successfully solved. The same method (rotation of plasma cable) can be applied in the present case. Another problem is production of the plasma cable. This problem may be solved if the electric station of the plasma cables’ origin is located on the Mars mountains (Bolonkin 2006d) or into AB electronic tube (Bolonkin 2008e).

10.4 Conclusions

This chapter proposes a new confinement for thermonuclear low density plasma (not magnetic, neither inertial). This reflect-reactor may be used for transportation and small power stations on Mars. Section 10.2 presents details of the design for this novel type of reactor.

A new revolutionary idea –the AB-Space Propulsion and wireless transferring of electric energy from Mars in space – is proposed in Sec. 10.3. It is based on a rarefied plasma power cord which operated like an electric cable. It is shown that a certain minimum electric current creates a compression force that supports and maintains the plasma cable in its compacted form. Large amounts of energy can be transferred hundreds of millions of kilometers from Mars surface by this method. The requisite mass of plasma cable is merely hundreds of grams. A sample macro-project is computed. It is also shown that electric current in plasma cord can accelerate or slow various kinds of outer space apparatus.

Acknowledgements

The author thanks Joseph Friedlander (Shave Shomron, Israel) for language editing and useful technical advices.

References

- American Institute of Physics, Physics Desk Reference, 3rd edn., p. 644. Springer, New York (2003)
- Bolonkin, A.A.: Installation for Open Electrostatic Field, Russian patent application #3467270/21-116676, July 9, 1982 (in Russian), Russian PTO (Patent Office) (1982a)
- Bolonkin, A.A.: Radioisotope Propulsion. Russian patent application #3467762/25-116952, July 9, 1982 (in Russian), Russian PTO (1982b)
- Bolonkin, A.A.: Radioisotope Electric Generator. Russian patent application #3469511/25-116927, July 9, 1982 (in Russian), Russian PTO (1982c)
- Bolonkin, A.A.: Space Propulsion Using Solar Wing and Installation for It. Russian patent application #3635955/23-126453, August 19, 1983 (in Russian), Russian PTO (1983a)
- Bolonkin, A.A.: Getting of Electric Energy from Space and Installation for It. Russian patent application #3638699/25-126303, August 19, 1983 (in Russian), Russian PTO (1983b)
- Bolonkin, A.A.: Protection from Charged Particles in Space and Installation for It. Russian patent application #3644168-136270, September 23, 1983 (in Russian), Russian PTO (1983c)
- Bolonkin, A.A.: Method of Transformation of Plasma Energy in Electric Current and Installation for It. Russian patent application #3647344-136681, July 27, 1983 (in Russian), Russian PTO (1983d)
- Bolonkin, A.A.: Method of Propulsion using Radioisotope Energy and Installation for It. Russian patent application #3601164/25-086973, June 6, 1983 (in Russian), Russian PTO (1983e)
- Bolonkin, A.A.: Transformation of Energy of Rarefaction Plasma in Electric Current and Installation for it. Russian patent application #3663911/25-159775, November 23, 1983 (in Russian), Russian PTO (1983f)
- Bolonkin, A.A.: Method of a Keeping of a Neutral Plasma and Installation for it. Russian patent application #3600272/25-086993, June 6, 1983 (in Russian), Russian PTO (1983g)
- Bolonkin, A.A.: Radioisotope Electric Generator. Russian patent application #3620051/25-108943, July 13, 1983 (in Russian), Russian PTO (1983h)
- Bolonkin, A.A.: Method of Energy Transformation of Radioisotope Matter in Electricity and Installation for it. Russian patent application #3647343/25-136692, July 27, 1983 (in Russian), Russian PTO (1983i)
- Bolonkin, A.A.: Method of stretching of thin film. Russian patent application #3646689/10-138085, September 28, 1983 (in Russian), Russian PTO (1983j)
- Bolonkin, A.A.: Light Pressure Engine, Patent (Author Certificate) No. 11833421, 1985 USSR (priority on 5 January 1983) (1983k)
- Bolonkin, A.A.: New Way of Thrust and Generation of Electrical Energy in Space. Report ESTI (Soviet Classified Projects) (1987)
- Bolonkin, A.A.: Aviation, Motor and Space Designs. Collection Emerging Technology in the Soviet Union, pp. 32–80. Delphic Ass. Inc., Washington (1990)
- Bolonkin, A.A.: The Development of Soviet Rocket Engines, p. 122. Delphic Ass. Inc., Washington (1991)
- Bolonkin, A.A.: A Space Motor Using Solar Wind Energy (Magnetic Particle Sail). In: The World Space Congress, Washington, DC, USA, IAF-0615, 28 August- 5 September (1992a)

- Bolonkin, A.A.: Space Electric Generator, run by Solar Wing. The World Space Congress, Washington, DC, USA, IAF-92-0604, 28 August - 5 September (1992b)
- Bolonkin, A.A.: Simple Space Nuclear Reactor Motors and Electric Generators Running on Radioactive Substances. The World Space Congress, Washington, DC, USA,, IAF-92-0573, 28 August – 5 September (1992)
- Bolonkin, A.A.: The Simplest Space Electric Generator and Motor with Control Energy and Thrust.The 45th International Astronautical Congress, Jerusalem, Israel, October 9-14,1994, IAF-94-R.1.368 (1994a)
- Bolonkin, A.A.: Light Multi-reflex Engine. J. British Interplanetary Soc. 57, 353–359 (2004b)
- Bolonkin, A.A.: Utilization of Wind Energy at High Altitude. Presented in International Energy Conversion Engineering Conference at Providence, RI, August 16-19, 2004, AIAA 2004-5705, AIAA-2004-5756, USA (2004c), <http://arxiv.org>
- Bolonkin, A.A.: Design of Optimal Regulator (2004d), <http://arxiv.org>
- Bolonkin, A.A.: Multi-reflex Space Propulsion. J. British Interplanetary Soc. 57, 379–390 (2004e)
- Bolonkin, A.A.: Non-Rocket Space Launch and Flight, p. 488. Elsevier, London (2006a)
- Bolonkin, A.A.: New Thermonuclear Reactor. In: AIAA-2006-7225, Conference Space 2006, USA (2006b)
- Bolonkin, A.A.: Electrostatic AB-Ramjet Space Propulsion. In: AIAA 2006-6173. AEAT, vol. 79(3) (2006c), <http://arxiv.org>
- Bolonkin, A.A.: Beam Space Propulsion. AIAA-2006-7492. Int. J. Aircraft Engng. Aerospace Technol. 80(2) (2008d), <http://arxiv.org>
- Bolonkin, A.A.: Electrostatic Linear Engine. AIAA-2006-4806. Int. J. Aircraft Engng. Aerospace Technol. 78, 502–508 (2006e), <http://arxiv.org>
- Bolonkin, A.A.: Optimal Solid Space Tower. AIAA-2006-7717. In: ATIO Conference, Wichita, Kansas, USA, September 25-27 (2006f), <http://arxiv.org>
- Bolonkin, A.A.: Theory of Space MagSail Some Common Mistakes and Electrostatic MagSail. (2006g), <http://arxiv.org>
- Bolonkin, A.A.: High Speed AB-Solar Sail. AEEE Transaction on Aerospace and Electronic System 2008 44(4) (2006h), <http://arxiv.org>
- Bolonkin, A.A.: New Method of Atmospheric Reentry of Space Ships (2006i), <http://arxiv.org>
- Bolonkin, A.A.: Simplest AB-Thermonuclear Space Propulsion and Electricity Generator (2006l), <http://arxiv.org>
- Bolonkin, A.A.: Wireless Transfer of Electricity in Outer Space. Presented as paper AIAA-2007-0590 to 45th AIAA Aerospace Science Meeting, January 8-11, 2007, Reno, Nevada, USA, vol. 78(3), pp. 273–282 (2007a), <http://arxiv.org>
- Bolonkin, A.A.: New Concepts, Ideas, Innovations in Aerospace and Technology, 480 p. Nova Science, New York (2008a)
- Bolonkin, A.A.: Thermonuclear Reflect AB-Reactor on Mars. Paper AIAA-2008-5150 to 44th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Hartford, CT, July 21-23, 2008 (2008b), <http://arxiv.org>
- Bolonkin, A.A.: AB-Space Propulsion for Mars (2008c), <http://arxiv.org>
- Bolonkin, A.A.: AB Electronic Tubes and Quasi-Superconductivity at Room Temperature (2008d), <http://arxiv.org> (presented on 8 April, 2008)
- Bolonkin, A.A., Cathcart, R.B.: Macro-Projects: Environment and Technologe, 470 p. Nova Science, New York (2008e)
- Kikoin, I.K.: Tables of Physical Values, p. 953. Atom, Moscow (1976) (in Russian)

Chapter 11

Geothermal Energy on Mars

Paul Morgan

Northern Arizona University, Flagstaff, USA

11.1 Introduction

This contribution will concentrate on the implications of data from new studies of Mars during the past decade or so in terms of martian geothermal resources, and the potential differences in exploiting geothermal resources on Mars from our terrestrial experiences. An excellent discussion of the utility of geothermal energy on Mars was given by Fogg (1996), who makes strong arguments for the practicality of using geothermal energy on Mars with respect to other potential energy sources: these arguments will not be repeated here.

Exploitation of terrestrial geothermal resources grew significantly in the fifteen years after 1990 and there is no indication that the increasing rate of use of geothermal energy is decreasing. From 1990 to 2005, geothermal electrical generation capacity was added globally at an average rate of 217 MW a^{-1} , and between 1995 and 2000 direct heat use grew at an average thermal energy rate of 1308 MW a^{-1} (Fig. 11.1). Although there is evidence of direct use of geothermal energy for cooking, therapeutic and other purposes since prehistoric times, and there has been continuous geothermal electricity generation since 1904, it is only rises in prices of fossil fuels and concerns about global warming that have resulted in significant recent increases in the use of this resource. With increased use of the resource technological advances have occurred, including the development of binary geothermal electricity generation power plants, in which geothermal hot water is used to heat a secondary low-boiling-point fluid, which in turn drives a turbine, allowing lower-temperature geothermal resources to be used to generate electricity (e.g., DiPippo 2008), and enhanced or engineered geothermal systems (EGS), in which the permeability of a geothermal reservoir is artificially stimulated and an artificial heat extraction system developed (e.g., Tester et al. 2006). As discussed below in Sect. 11.6, both of these technologies may have application to Mars.

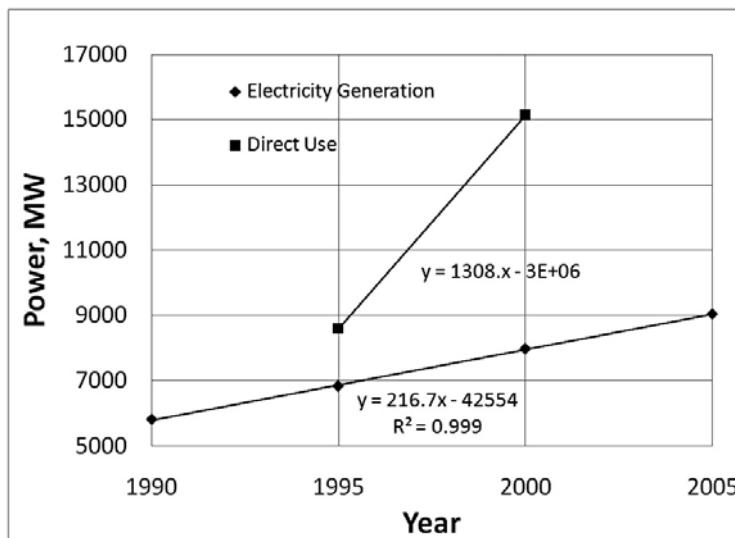


Fig. 11.1 Growth in use of geothermal energy between the years 1995 and 2005. Data source: International Geothermal Energy Association (<http://www.geothermal-energy.org/index.php>, last accessed 2009-4-29)

New remote sensing data from Mars, primarily orbital imagery, have been interpreted to support the concept that some areas of the surface of Mars have young volcanism. Exactly how young is young, is still in question: hundreds or tens of millions, or even hundreds or tens of thousands or years, but they indicate that Mars is probably not a dead, but a dormant planet, with a significant probability of near-surface heat and even the possibility of future volcanic eruptions. This question will not be resolved until either we observe a volcanic eruption on Mars, or, more likely, we are able to make measurements of the crustal thermal gradient in areas of “young” tectonothermal activity. The present discussion is based on the speculation that areas exist on Mars at which the temperatures are hot enough at shallow enough depths to be exploited as a geothermal resource.

Since two Viking spacecraft landed on the surface of Mars in 1976, data have shown how different are the atmospheric conditions on the surface of Mars from those on Earth (Hess et al. 1977). The mean surface temperature on Mars is about -63°C (range -140 to $+20^{\circ}\text{C}$) and the mean pressure is about 7 mb. In contrast, the mean surface temperature on Earth is about 15°C (range -90 to $+55^{\circ}\text{C}$) and the mean pressure is about 1013 mb. Both the low temperature and low pressure are significant in terms of the exploitation of geothermal resources on Mars. Seven mb is very close to the triple point of water (6.11 mb; 0.01°C) which has implications for the phase of water if geothermal fluids are reduced to martian atmospheric pressure, as discussed below. The low temperature of the martian atmosphere is an advantage for cooling in energy extraction from geothermal fluids, but it prevents liquid water from being stable at the surface.

11.2 Geothermal Resources on Mars

In its terrestrial definition, a *geothermal resource* is that part of the Earth's heat that can, or could, be recovered and exploited for human use. This definition may be adapted for Mars by substituting Mars' heat for Earth's heat. Also often included with terrestrial geothermal resources are *geothermal heat pumps*, also known as *ground-source* or *geoexchange heat pumps*. These systems place heat-exchange coils at relatively shallow depths within the Earth (typically 2-100 m) and either pump heat out of or into the shallow crust, most commonly for space heating and cooling. These systems do not strictly meet the definition of a geothermal resource as they do not use part of the Earth's heat and simply use the shallow crust as a thermal reservoir. They work most efficiently when the shallow crust is close in temperature to the desired temperature for space heating or cooling. They are unlikely to be practical for Mars, as used on Earth, where the mean annual surface temperature is -63°C, significantly different from anticipated temperatures for space heating or cooling. Should temperatures in this range be required for any purposes, however, geothermal heat pumps should be considered as they are very efficient in this function.

For the terrestrial planets, their current internal heat budget long after formation is primarily dependent on their size, the larger the planet the greater their internal heat reserve. Thus, Earth, as the largest of the terrestrial planets, has the highest current internal heat budget. This heat budget includes remaining primordial heat associated with formation and differentiation of the planet and heat generated by the decay of relatively abundant unstable isotopes with long half-lives, ²³⁵U, ²³⁸U, ²³²Th, and ⁴⁰K. Even with the highest internal heat budget and surface heat flow (the flow of heat out of the planet), relatively few areas of Earth have shallow high temperatures sufficient for the generation of electricity with current technology. Mars is approximately half the radius of Earth and 10% its mass. It is therefore expected to have a much smaller internal heat budget, but a much smaller surface area over which this heat is lost. By analogy, geothermal resources on Mars would be expected to be restricted to areas where the subsurface is significantly hotter than normal. On Earth, high subsurface temperatures are generally associated with zones of geologically young tectonic and volcanic activity (50 Ma or less). No direct measurements currently exist of subsurface temperatures or heat flow on Mars. However, before proceeding with a discussion of geothermal energy on Mars, a brief summary of evidence suggesting that there is a reasonable probability that these resources exist on Mars is required.

11.3 Evidence for Viable Geothermal Resources on Mars

Evidence for viable geothermal resources on Mars may be drawn from two sources, measures of average heat that is escaping from the surface of Mars, and indications of anomalous potentially high heat loss associated with young tectonism and volcanism. Detailed discussions of these subjects are not given here, but sufficient information and references are given to show that they support the potential for viable geothermal resources on Mars.

11.3.1 Mars Global Heat Flow

A rough idea of the global heat loss from a planet may be obtained from consideration of the internal heat budget of the planet. The primary components of this budget are heat remaining from primordial heat, active internal heat generation, and surface heat loss. For measurement purposes, the surface heat flow, q , is usually defined as the vertical conductive heat flow out of the surface of a planet, and is given by Fourier's Law of conduction:

$$q = -K \frac{\partial T}{\partial z} \quad (11.1)$$

where T is temperature, z is depth, and K is the thermal conductivity of the regolith or rocks in which the temperatures are measured. The negative sign indicates that heat flows down thermal gradient, i.e., out of the planets, and is commonly omitted. Corrections are necessary for conditions that create deviations from this simple assumption of steady-state, one-dimensional, conductive heat flow (e.g., Beardsmore and Cull 2001). Heat is also lost by convection of heat to the surface by magma in volcanism or water in hot springs and geysers. These are typically transient phenomena, and also transfer heat in the subsurface. Subsurface temperatures may be estimated from Eq. (11.1) if the surface heat flow and thermal conductivity are known, assuming conductive conditions.

A minimum estimate of Mars heat flow may be made by estimating the heat produced by radiogenic heat producing elements in the planet and dividing by its surface area. The lowest probable heat production rate for Mars may be based on the assumption that Mars has the composition of a C1 carbonaceous chondrite meteorite (7.4 ppb U, 29 ppb Th, .055% K; McDonough & Sun, 1995). This composition gives a heat production rate of $3.36 \times 10^{-6} \mu\text{W kg}^{-1}$ and a mean surface heat flow of 15 mW m^{-2} . Wänke and Dreibus (1994) gave a bulk composition based on Mars meteorites with 16 ppb U, 56 ppb Th, and 0.0305%K, which gives a bulk heat production rate for Mars of $4.02 \times 10^{-6} \mu\text{W kg}^{-1}$: the mean surface heat flow with this value is 18 mW m^{-2} . Using different assumption for Mars as the parent body of the same meteorites, Lodders and Fegley (1997) calculated heat production concentration values of 12.6 ppb U, 44 ppb Th, and 0.073% K. These values yield a bulk heat production rate of $4.87 \times 10^{-6} \mu\text{W kg}^{-1}$, and mean surface heat flow of 22 mW m^{-2} .

Radiogenic heat production gives a minimum estimate of current heat loss because it assumes that this heat generation is the only heat source in the planet and that this heat is lost at the same rate that it is generated. In practice, additional heat is lost through secular cooling of the planet, and there is a delay in the transport of heat from the interior of the planet to its surface, so heat lost at its surface represents an integration of heat generated over previous time periods when radiogenic heat production was higher. The ratio of internal heat production to surface heat loss is given by a parameter called the Urey ratio (see Korenaga, 2008, for more information). The Urey ratio for Mars has been estimated to be 0.6 (Breuer and Moore, 2007), so the mean surface heat loss for Mars based on internal heat production should be increased from a range of 15 to 22 mW m^{-2} to 25 to 37 mW m^{-2} . This latter estimated range for the mean heat flow of Mars is consistent with the

two heat flow measurements from the Moon (Apollo 15, 22 mW m^{-2} ; Apollo 17, 16 mW m^{-2} ; Langseth et al., 1976), and Earth's mean heat flow of $87 \pm 2 \text{ mW m}^{-2}$ (\pm standard error, $n = 20,201$; Pollack et al. 1993; Stein 1995), when the relative sizes of these three planetary bodies are considered.

Global heat flow may also be estimated from models of mantle convection that include both the internal heat production and thermal evolutionary components of surface heat loss. For example, Hauck and Phillips (2002) developed a model that included generation of the Mars' crust and enrichment of this crust in the heat producing radioisotopes. Surface heat flow, heat flow from the convecting mantle, and radioisotope enrichment of the crust as a function of time predicted by this model are shown in Fig. 11.2. A chondritic concentration of heat production was used in this model, and the modern surface heat flow for Mars predicted by the models is about 25 mW m^{-2} . Breuer and Moore (2007) and Guest and Smrekar (2007) considered three models, a stagnant lid (one plate), early plate tectonics, and lithospheric delamination or convective overturn. They show that these different styles have significant implications for the early thermal history of Mars, but become less significant with time. Convection models by Weizmann et al. (2001) suggest that recent (post 1 – 2.5 Ga) volcanism on Mars should have been related to plume activity rather than large-scale convective upwellings, and Zanarak and Manga (2007) have suggested that these plumes could have originated in a compositionally stratified layer deep in the Mars mantle. These are but a few of the model studies of convection in Mars: few of these studies produce predictive values of mean surface heat flow, but those that do give values in general agreement with the estimates based on radiogenic heat production corrected by the Urey ratio.

A third method of estimating Mars heat flow is based on lithospheric flexural mechanics of local features (Solomon and Head, 1990). McGovern et al. (2002, 2004) have compared gravity and topography admittance functions in the spectral domain with lithospheric flexural models to estimate the elastic thickness of the lithosphere for a variety of areas on Mars. From these elastic thicknesses and rock strength parameters they have estimated geothermal gradient and heat flow. Their results are summarized in Fig. 11.3. The geothermal gradient and heat flow values calculated by this method are applicable to the time at which the major feature under study stabilized in terms of its gravity and topography, and thus by plotting the thermal estimates as a function of age of the features a version of the thermal evolution of Mars is produced. Heat flow estimates by this method generally decrease with time. There are relatively large allowable ranges in the elastic thicknesses of the lithosphere. In particular, a wide range of heat flow values, $13\text{--}140 \text{ mW m}^{-2}$ is allowable by the data for the Amazonian volcano, Ascraeus Mons (McGovern et al., 2004). Continuing studies of the thermal structure and evolution of the lithosphere associated with lithospheric structure have illustrated the significance of the rheological models used in these studies and the potential spatial heterogeneity of the flexural rigidity in defining the elastic thickness and estimating heat flow (e.g., Dombard and Phillips 2005; Grott and Breuer 2008; Phillips et al. 2008).

Other authors have used fault spacing to estimate lithospheric thickness and thermal parameters. Using wrinkle ridges to measure fault spacing, Montési and Zuber (2003) estimated a heat flow of $37 \pm 10 \text{ mW m}^{-2}$ for the northern lowlands

during the Hesperian (3.7 – 3.0 Ga) and 27 – 47 mW m⁻² for Solis Planum for the same period. This range is slightly higher than the range of 18–35 mW m⁻² calculated by McGovern et al. (2002, 2004) for the same area using admittance functions. When compared with the thermal evolution model of Hauck and Phillips (2002, and Fig. 11.3) there is in general agreement.

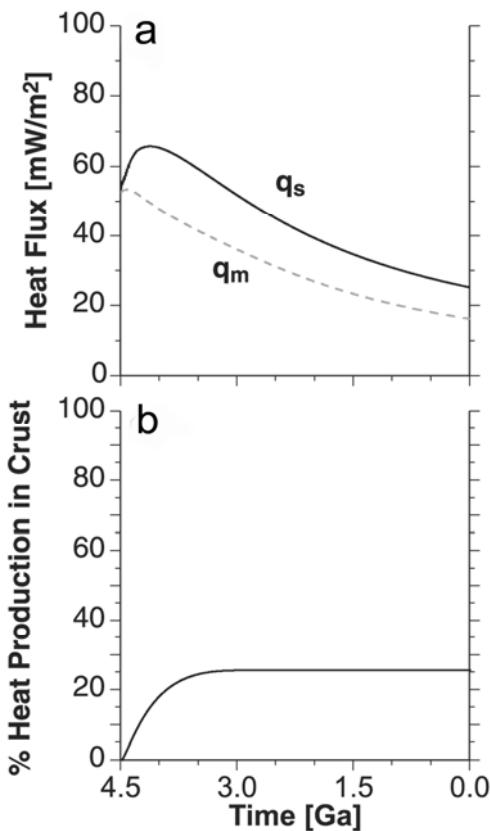


Fig. 11.2 a. Mean surface heat flow (q_s) and mean mantle heat flow (q_m), and b. % of heat production fractionated into the crust as a function of time before present calculated from a model that included generation of the Mars' crust and enrichment of this crust in the heat producing radioisotopes presented by Hauck and Phillips (2002). [Figure modified from Hauck and Phillips, 2002]

In summary, although there are currently no direct thermal gradient or heat-flow measurements on Mars, several lines of evidence indicate that the present mean heat flow is probably in the range of 25–35 mW m⁻². Estimates derived from elastic lithosphere calculations are somewhat lower. However, elastic lithosphere calculations are based on integrated temperatures in the portion of the lithosphere that controls elastic strength, and do not account for the concentration of the heat producing elements into the crust, which significantly increases surface heat flow (e.g., Blackwell, 1971; Morgan, 1984). The terrestrial mean heat flow is 87 ± 2 (standard deviation, $n = 20201$), which breaks down to 65 ± 1.6 ($n = 10337$) for continental regions and 101 ± 2.2 ($n = 9864$) for oceanic regions (Pollack et al., 1993). There is a large variation in terrestrial heat flow, generally low in stable areas and high in regions of young tectonism and volcanism. The hottest geotherms

are buffered by the water boiling point curve and crust and upper mantle melting point curves (solidi). Thus, although Mars mean heat flow is low relative to the terrestrial mean, Mars heat flow may be expected to vary in different regions, with the highest heat flow in areas of young tectonism and volcanism.

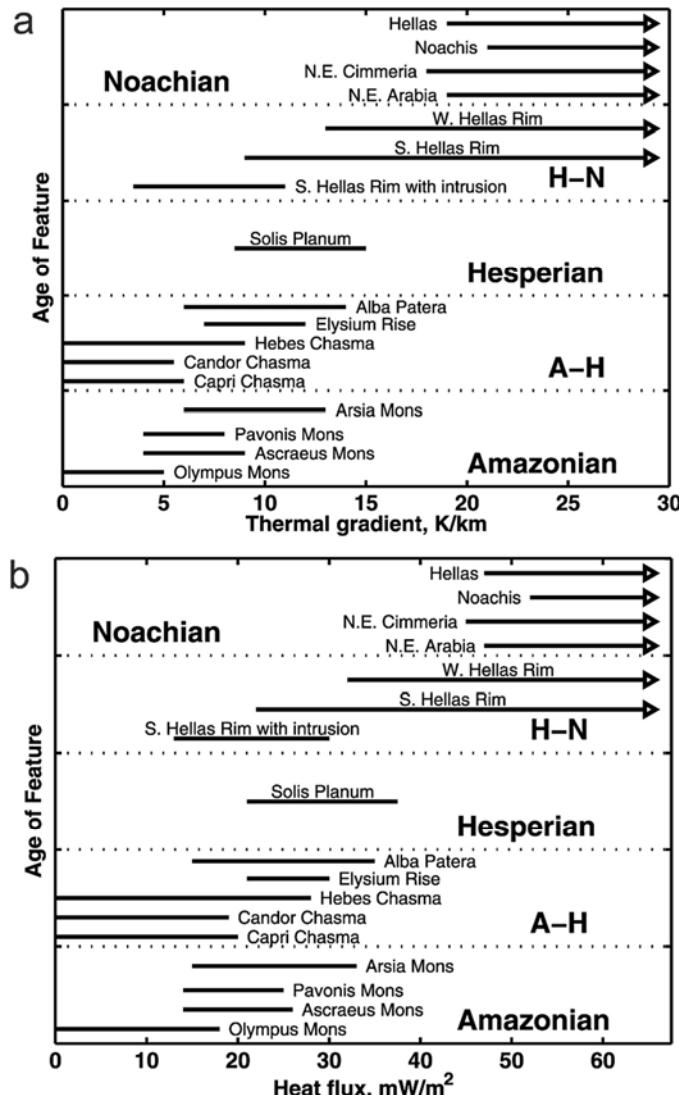


Fig. 11.3 a. Thermal gradient, and b. heat flow, as a function of age of the features calculated by McGovern et al (2002, 2004) in their analysis of gravity/topography admittance functions and elastic lithospheric thicknesses. [From McGovern et al, 2004]

11.3.2 Tectonism

Tectonism is associated with terrestrial geothermal resources on regional and local scales. On a regional scale, areas of geologically young tectonics (faulting, folding and uplift) are usually associated with high heat flow (Morgan 1984). On a local scale, active, or recently active faults act as conduits for geothermal fluids, and even free hydrothermal convection (Benoit 1999; Blackwell et al. 1999, 2002).

The ideal test of active tectonism on Mars would be data from an array of long-term seismic stations, but no seismometers have been successfully installed to record marsquakes. However, image data have been interpreted to indicate that tectonic activity in the form of faults has been occurring throughout Mars geologic history to the present. Anderson et al. (2001) report that fault structures radial to Tharsis, the largest uplift on Mars, have been active from the Noachian to the present, and Neukum et al. (2004) and Basilevsky et al. (2006) report evidence from the eastern flank of Olympus Mons, a massive shield volcano, for very recent ($\leq 25\text{--}40$ Ma) tectonic (wrinkle ridge and troughs) and volcanic (emplACEMENT of lava flows and dikes). Thus, there is good evidence that Mars remains tectonically active at least in limited areas.

11.3.3 Volcanism

The only direct ages of volcanism on Mars come from meteorites that are believed to have originated on Mars. These meteorites are all basalts or basaltic cumulates in composition, and range in age from 4.5 Ga to 180 Ma (Treiman et al., 2000). Most of the Mars meteorites have ages less than a few hundred million years, however, and their age distribution is shown in Fig. 11.4. At 180 Ma, the youngest martian meteorite is just 3.9% of the age of Mars, and as there is no reason to believe that the youngest meteorite represents the youngest volcanism on Mars, this young age suggests that younger volcanic rocks remain to be found, as indicated in crater ages discussed below. In addition, the observation that the ages of the meteorites cluster toward younger ages suggests that either the surface of Mars is biased toward younger ages, or the sites from which the meteorites are ejected are biased toward younger ages. High-resolution computer simulation of meteorites ejected by impacts indicate that meteorites are likely to be biased toward younger terrains mantled by thin regolith at these sites and away from older terrains mantled by thick regolith (Head et al. 2002). However, although the sample may be biased toward younger ages, the meteorites provide direct evidence of geologically young crystallization ages on Mars, young enough to allow speculation that volcanic activity could continue to the present.

Plescia (1990) suggested that large-scale eruptions of low-viscosity lava were still possible on Mars, and during the past decade or so several authors have presented interpretations from spacecraft images supporting volcanism during the past 1% of Mars history (e.g., Fuller and Head 2002). Reports of recent activity have been given for a number of areas based on crater counts, including Olympus Mons (Basilevsky et al. 2006) and Arsia Mons (Hartman et al., 1999) in Tharsis, and Cerebus Plains (Vaucher et al. 2006). Malin et al. (2006) reported that gully

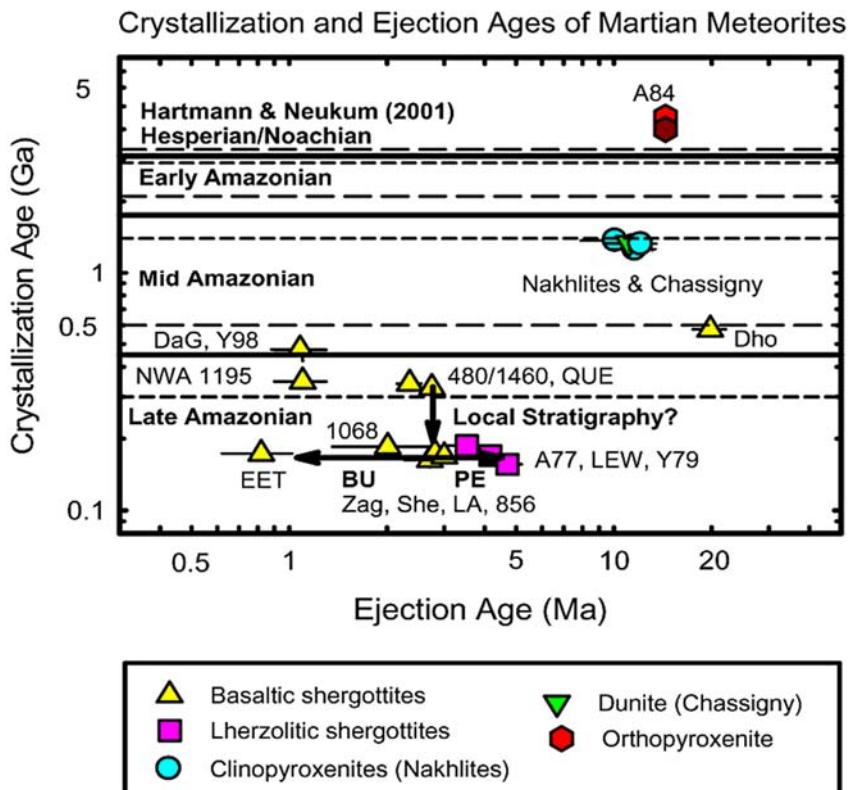


Fig. 11.4 Crystallization and ejection ages of martian meteorites. Note that ages scales are logarithmic which emphasizes the young ages. See Treiman et al. (2000) for details of individual points on the plot. [From Treiman et al., 2000]

forming processes are currently active. Work continues on refinement of crater age-dating for very young sites, but these studies indicate that volcanism has occurred on the surface of Mars with ages <100 Ma (Hartman, 2005, 2007; Malin et al. 2006; Hartmann et al. 2008). Terrestrial measurements indicate that areas of active or recently active (< 15 – 60 Ma) volcanism have significantly elevated heat flow (Morgan, 1984), commonly exceeding $200 - 300 \text{ mW m}^{-2}$. These high values are probably not a linear function of the age of a planet, but are buffered by the cooling regime of magmas in the near surface, and hydrothermal activity if water is present. When the global heat loss of a planet is higher, more areas of high heat flow will occur, but maximum values are expected to be of similar magnitude but less frequent as long as volcanism continues as global heat loss decreases. Thus, the evidence for continuing volcanic activity on Mars strongly suggests that local areas of high heat flow will persist to the present.

11.4 Martian Geothermal Resources

Analyses of the global thermal budget of Mars using a variety of technique suggest that the present mean heat flow from Mars is in the range of 25 to 35 mW m⁻². On Earth a wide scatter in heat flow is observed about the mean: low heat flow is measured above the fore-arc regions of subduction zones and high heat flow is measured on mid-oceanic ridges and in areas of recent volcanism and tectonics. Evidence of plate tectonics on Mars is cryptic, definitely in the recent surface geologic record, and thus extensive areas of low heat flow associated with subduction zones are not expected. However, uniform heat flow over the surface of Mars is unlikely. Assuming a modest surface heat flow of 25 mW m⁻² and a thermal conductivity of 2.0 W m⁻¹ K⁻¹ for the crust and 4.0 W m⁻¹ K⁻¹ for the mantle, with a 40 km thick crust with heat production typical of basalts, temperatures in the range of 1200 to 1600°C are predicted to be reached at depths of >120 km, commensurate with the lithospheric thickness calculations of the mechanical models. At these depths the mode of heat transport must change from conduction to convection, perhaps with a component of radiation, or super-solidus temperatures are predicted. Convection is unlikely to result in uniform input of heat into the base of the Mars lithosphere, especially if some component of the convection involves plumes: rising convecting material probably still generates magma by pressure release melting at present in Mars (even if restricted to plumes; Weizmann et al, 2001), and this magma may result occasionally in volcanism.

Intrusion of magma into the upper crust of Mars should give rise to viable geothermal resources. Head and Wilson (2002, 2007) extensively discussed potential magma-water interactions on Mars including subsurface hydrothermal systems. Many of the major surface features on Mars, including the chaotic terrain, have been interpreted as the result of catastrophic melting of ground ice by magmatic intrusions, and it is likely that magma continues to interact with water and other volatiles in the upper crust of Mars. No specific site has yet been targeted as a definite geothermal resource on Mars, but, as exploration continues, more than one viable resource is likely to be found, probably in one or more of the areas that have already been identified to have very recent volcanism.

11.5 Exploitation of Geothermal Resources on Mars

A primary consideration for the exploitation of geothermal resources on Mars will be that use will need to be relatively close to the resources. Steam and hot water can be transported a few tens of kilometers in insulated pipes, but these pipes are resource intensive and transport should be minimized if possible. Similarly, if the resource is used to generate electricity, as with any electrical energy source, a transmission system will be needed to take the electrical energy to the user, and this should be minimized on a planet where no infrastructure exists. Depending on the temperature and flow of the resource, typical electrical power production from a single well is a fraction to a few MW.

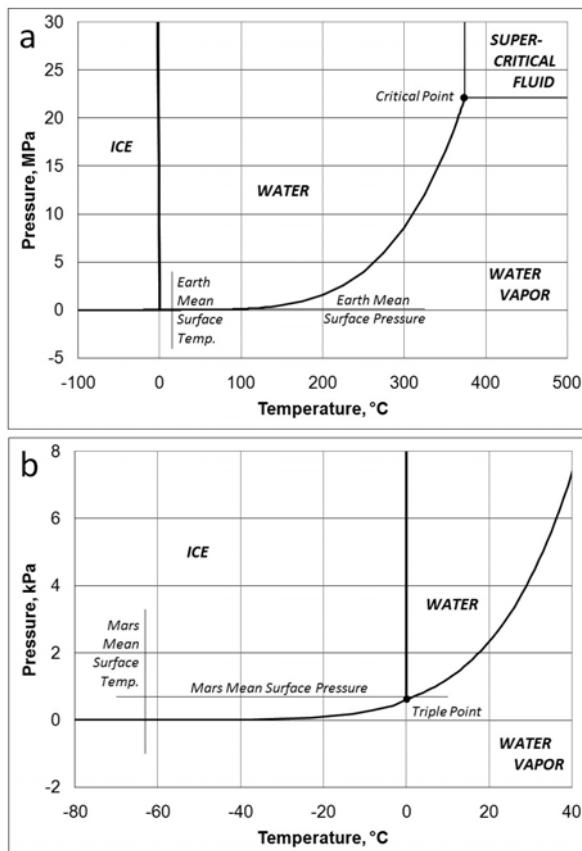


Fig. 11.5 Phase diagram for water. **a.** Portion of phase diagram relevant to geothermal fluid production. Above the critical point, 374.11°C , 22.106 MPa : water and water vapor do not exist as separate states. **b.** Enlarged portion of phase diagram around the triple point of water, 0.01°C , 611.73 Pa , the temperature and pressure at which all three phases of water can co-exist. Note different vertical scales on a. and b.

As the mean annual surface temperature of Mars is -63°C , direct use of low-temperature geothermal fluids ($\sim 30\text{--}50^{\circ}\text{C}$) for space heating and other low-temperature process heat would be a very efficient use of this resource. Energy is lost whenever energy is converted into a different form, and if the end product is heat, the most efficient use of geothermal energy is direct use. A simple direct heat system would consist of one or more production wells and one or more injection wells and the water would need to be maintained under sufficient pressure to prevent boiling in the tenuous martian atmosphere before reinjection (Fig. 11.5).

Three basic types of geothermal generating power plant are in use at present, as shown diagrammatically in Fig. 11.6. In order of increasing temperature of geothermal fluid they are: a) binary power plant; b) single-flash or wet-steam power plant; and c) triple-expansion or dry-steam power plant. The binary power plant

exchanges the heat from the geothermal fluid with a lower boiling-point fluid (usually an organic fluid), which is in turn used to drive a turbine. The lower-boiling-point fluid is then cooled, and recycled to be reheated. The geothermal fluid is not exposed to the atmosphere, but is reinjected to replenish the source.

In a single-flash power plant, wet steam is first separated into steam and water, and then the steam is used to drive a turbine before pumping into a cooling unit. In older systems the waste water and used steam were sometimes vented into the environment. But today the steam is fully condensed, added to the waste water, and reinjected to replenish the source.

In a triple-expansion power plant, dry steam is fed directly into a turbine, and then, when it has lost some of its energy, it is fed into a high-pressure separator, steam is removed to drive a high-pressure turbine, and then the residual water is fed into a low-pressure separator yielding low-pressure steam that is fed into a coupled low-pressure turbine. The residual steam is cooled and condensed. As with the single-flash units, fluids were once released to the atmosphere, but are now reinjected to replenish the source.

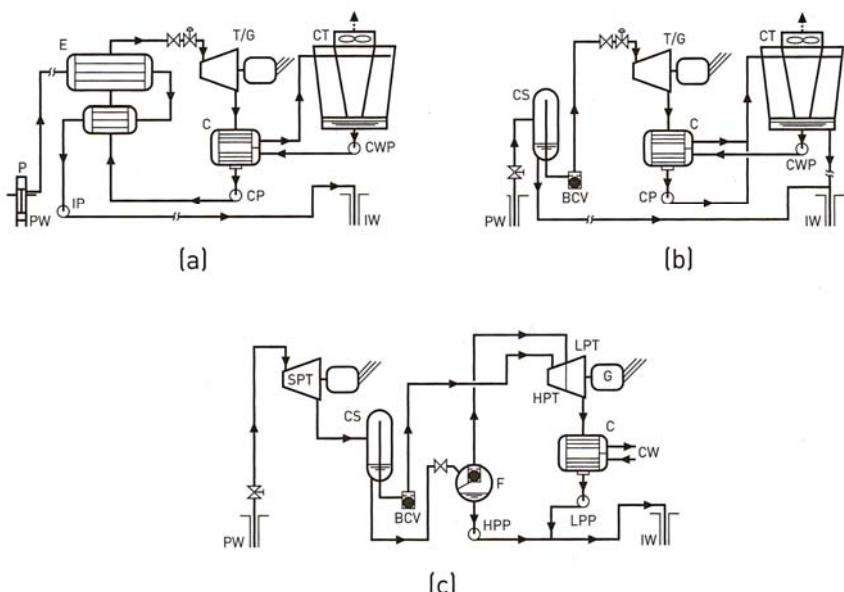


Fig. 11.6 Diagrammatic representations of different types of geothermal power plants. **(a)** Binary power plant. **(b)** single-flash power plant. **(c)** triple-expansion power plant. Key: PW production well; P pump; IP injection pump; E heat exchanger; T/G turbine/generator; C condenser; CP compressor pump; CWP cold-water pump; CT cooling tower; IW injection well; CS compression separator; BCV back compression valve; SPT steam pressure turbine; F low-pressure separator; HPT high-pressure turbine; LPT low-pressure turbine; G generator; HPP high-pressure pump; LPP low-pressure pump; CW cold water. [From Tester et al., 2006].

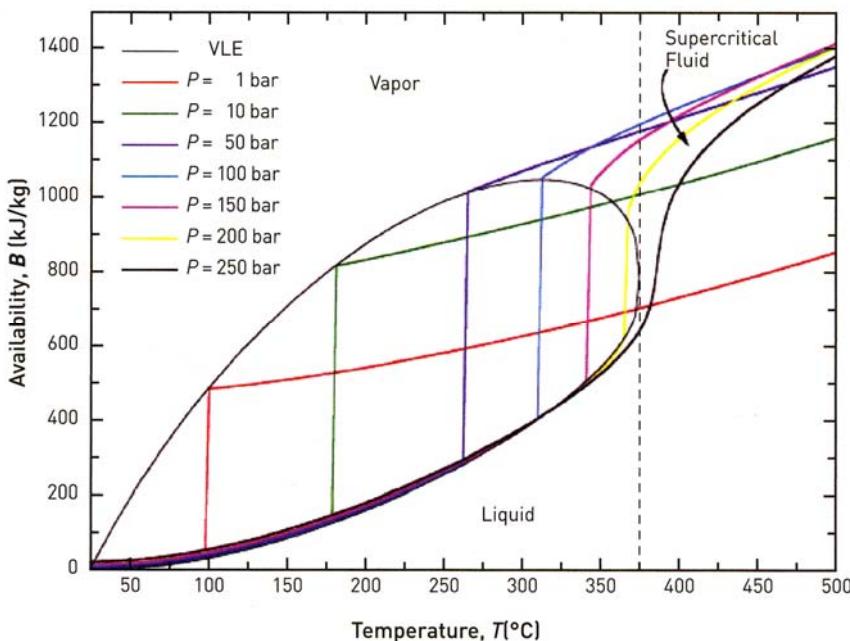


Fig. 11.7 Water availability diagram. This diagram shows the energy released from water as it cools at different pressures. VLE is the vapor-liquid envelope. [From Tester et al. 2006, Fig. 1.10].

Without knowledge of the temperatures and depths of potential Mars geothermal resources, their use for electricity production must be examined in terms of theoretical energy cycles. Making the simplifying assumption that geothermal waters on Mars are relatively pure, we may use the water boiling-point curve as a function of pressure, shown in Fig. 11.5a, to estimate the flash temperature of these fluids. Assuming a density of 1.0 Mg m^{-3} for water and the acceleration due to gravity for Mars to be 3.71 m s^{-2} , pressure may be converted to hydrostatic depth using the conversion factors 0.27 km MPa^{-1} (Fig. 11.5a) and 27 m kPa^{-1} (Fig. 11.5b). Thus, a water boiling point of 100°C would be reached at a hydrostatic depth of 27.3 m , 200°C at 419 m , 300°C at 2.32 km , and the critical temperature of 374.11°C would be reached at 5.96 km . A geotherm falling on this temperature depth curve would indicate temperatures moderated by the boiling point curve for Mars. A cooler geotherm would indicate geothermal fluids that would not flash to steam: a hotter geotherm would indicate fluids that would be capable of flashing to steam and that would be at pressures above hydrostatic. At temperatures greater than 374.11°C water would exist as a supercritical fluid.

Figure 11.7 indicates the energy that may be extracted from a geothermal fluid, again using the assumption that the fluid has the properties of pure water. Curves on Fig. 11.7 show paths for isobaric (constant pressure) cooling, but some pressure drop between the production and rejection of the geothermal fluid is to be

expected. From Fig. 11.5a and 11.5b, the phase of any geothermal fluid may be determined from its pressure and temperature. A geothermal fluid may be defined by conditions at any point in the fluid regions of Fig. 11.5a, including the super-critical fluid region, but in general, fluids have more energy if they have higher temperatures and higher pressures. However, once the temperature and pressure of a fluid is defined, its energy per unit mass is defined, as shown in Fig. 11.7. What may be changed is where that energy is released.

Rock is a poor conductor of heat and as a geothermal fluid rises to the surface the rock surrounding the well is quickly heated to the temperature of the geothermal fluid and conductive heat loss from the fluid is minimal. Therefore temperature change in the geothermal fluid during ascent is primarily associated with pressure change, typically pressure release. Significant pressure release during ascent is generally undesirable as it involves energy release, and may result in phase change from liquid to vapor in the well-bore, and therefore the geothermal fluid is usually maintained under pressure until it enters a power plant. These conditions are unlikely to be significantly different on Mars from terrestrial geothermal power plants. The significant differences on Mars would be on the cooling side of the system.

For maximum efficiency, the cooling side of the fluid running through the generating turbine should be as cold and at as low pressure as possible with reasonable engineering and economic constraints. The lowest practical pressure in terrestrial systems is atmospheric pressure, and typically the minimum temperature is controlled by the availability of cooling water. These conditions will never cool the geothermal fluid below 0°C (Fig. 11.5a), and if the fluid is kept under reasonable pressure it can be maintained or returned to a liquid form for reinjection. Ideally, systems can be designed to operate with only small net loss of geothermal fluid. On Mars, however, atmospheric temperatures and pressure are significantly lower than on Earth, the stable phase for water ice (Fig. 11.5b), and no surface water is available for cooling. There is good evidence for reservoirs of water in the shallow subsurface of Mars (e.g., Carr and Wänke 1992). However, if these waters are pumped to the surface for cooling, in the event of a mechanical breakdown, these waters will be subject to freezing. Thus, “air” cooling may be less subject to environmental problems than water cooling on Mars.

Air cooling has been successfully used for hydrothermal and binary terrestrial power plants (e.g., Kanoğlu and Çengal, 1999). The only disadvantage of this method of cooling is that plant efficiency changes seasonally as the air temperature, and therefore the cooling temperature changes. On Mars, the mean annual surface temperature of -63°C would be an advantage in terms of air cooling. However, cooling efficiency also depends on air density, and as the Mars atmosphere is only about 1% that of Earth’s atmosphere, its cooling efficiency per unit volume would be only about 1% that of Earth’s atmosphere for the same temperature difference. Fan power required to move air scales inversely with density, thus, if there were no energy losses, the same fan power could produce the same cooling efficient on Mars as on Earth, but the fans would be required to move approximately one hundred times the volume of the low density Mars atmosphere for the same cooling with the same temperature difference (Comair Rotron, 2009). As the

Mars atmospheric temperature is significantly colder than Earth's atmospheric temperature, less than one hundred times the volume of Mars air would be required for the same cooling effect, and thus the power to drive the Mars fans would be less than for an equivalent terrestrial system, but they would be required to move a larger volume of low-density air.

11.6 Geothermal Development on Mars

The next stage in serious consideration of using geothermal energy on Mars should be continued efforts to locate and confirm the areas of young volcanism. These efforts include remote sensing, but require visiting the sites by one or more landers to perform quantitative age dating on the volcanic rocks and to collect samples for sample return. Special attention should be given to the search for evidence of young hydrothermal activity.

The only way to test a geothermal resource is through drilling. Shallow thermal gradient/heat flow measurements should be made to locate the best site for deeper drilling, and to estimate the depth needed to drill for a suitable resource. Typical terrestrial geothermal wells are of the order of 1000 m deep. When a deep test well is drilled, it should be suitable not only for temperature and other geophysical logs, but also for pump tests to be made to investigate the permeability of the rocks where temperatures are high enough for exploitation. Once the resource has been evaluated, decisions may be made on future paths for development. Geothermal binary power plants are commonly constructed by aggregating a number of small power plants (e.g., Kanoğlu, 2002), and unit sizes currently in production range from about 500 kW to 2 MW. These units are self-contained and require only a feed from the geothermal source and an outlet to the cooling system. As such, they would be the preferred current technology for use on Mars.

At this stage I do not believe that further speculation is appropriate. From the information available at present, the probability of finding geothermal resources on Mars is good in terms of temperature. The other essential component of a geothermal resource is adequate fluid and permeability. Basaltic lava flows on Earth make good aquifers, but these are in hydraulic regimes that have significant groundwater flow and continual tectonic readjustment. Flow is generally best between flows, in fractures, or in faults, as natural porosity in basalt tends to be primarily in the form of non-interconnected vesicles. Impact activity on Mars may have increased fracture permeability, or decreased permeability at depth by compaction. Evidence of chemical deposition has been seen on images from the Mars rovers: will porosity be decreased by deposition at depth, or increased by solution? Flow rates to produce a few MW are of the order of 100 kg s^{-1} . The availability of adequate fluids in martian geothermal resources is a significant question.

Research and development is currently in progress to stimulate artificial geothermal reservoirs in what are called engineered or enhanced geothermal systems (EGS). An extensive discussion of these techniques is given in Tester et al. (2006). The primary stimulation technique used in EGS is hydrofracturing in which rocks are fractured at depth by pumping water into a drill hole until the rocks fracture.

This is a mature technology in stimulating production of oil and gas in low permeability formations. Such techniques are potentially applicable to Mars if sufficient permeability is not found, but require large volumes of water for pumping and extensive pumping machinery. Thus, such considerations are probably not suitable for consideration for an early settlement of Mars.

11.7 Outlook for Geothermal Use on Mars

Theoretical studies, meteorite ages, and remote-sensing observations suggest that Mars has a good potential for geothermal resources. This potential is highest in areas that have been identified as having recent volcanism. If age dating of sample from these sites confirms the youth of the volcanism, then the potential for geothermal resources is very high. A program of shallow thermal gradient/heat flow measurements should be made before locating the site for a deep geothermal test.

Geothermal resources will only be confirmed by deep drilling, and this test should include flow tests in addition to temperature and other logs. As basalt lavas form good aquifers on Earth, the martian lavas may host good geothermal reservoirs. Depending on the fluid temperature, the resource may be used for direct use, power generation, or both. Practical considerations restrict the use of the resource to relatively close (a few 10s of km) of its source. Binary power plant technology with air cooling is the current system that is most ready for adaptation for use on Mars.

If the resource is found to have high temperature but low permeability, it could be used for direct heat, but at first is unlikely to be useful for power generation. Methods of stimulating reservoirs are under development that improves their productivity: at present these techniques are probably not practical for application to Mars. However, if new less resource-intensive techniques are developed, or as more equipment intensive projects are feasible on Mars, low-permeability reservoirs may be opened up for power production.

Geothermal resources have a good potential to contribute to the energy needs of settlements on Mars.

Acknowledgement

I am very grateful for very constructive reviews of an early draft of this manuscript by M. Wiezorek and M. Grott. I am also grateful to many authors who provided reprints and preprints.

References

- Anderson, R.C., Dohm, J.M., Golombek, M.P., Haldermann, A.F.C., Franklin, B.J., Tanaka, K.L., Lias, J., Peer, B.: Primary centers and secondary concentrations of tectonic activity through time in the western hemisphere of Mars. *J. Geophys. Res.* 106, 20563–20585 (2001)

- Basilevsky, A.T., Werner, S.C., Neukum, G., Head, J.W., van Gasselt, S., Gwinner, K., Ivanov, B.A.: Geologic recent tectonic, volcanic and fluvial activity on the eastern flank of the Olympus Mons volcano Mars. *Geophys. Res. Lett.* 33, GL026396 (2006)
- Beardsmore, G.R., Cull, J.P.: Crustal Heat Flow: A Guide to Measurement and Modelling, 324 p. Cambridge University Press, Cambridge (2001)
- Benoit, D.: Conceptual models of the Dixie Valley, Nevada geothermal field. *Geothermal Resource Council, Trans.* 23, 505–511 (1999)
- Blackwell, D.D.: The thermal structure of the continental crust. The Structure and Physical Properties of the Earth's Crust. In: Heacock, J.G. (ed.) Am. Geophys. Un., *Geophys. Mono.*, vol. 14, pp. 169–184 (1971)
- Blackwell, D.D., Wisian, K.W., Benoit, D., Gollan, B.: Structure of the Dixie Valley geothermal system, a “typical” Basin and Range geothermal system, from thermal and gravity data. *Geothermal Resource Council, Trans.* 23, 525–531 (1999)
- Blackwell, D.D., Leidig, M., Smith, R.P., Johnson, S.D., Wisian, K.W.: Exploration and development techniques for Basin and range geothermal systems: examples from Dixie Valley, Nevada. *Geothermal Resource Council, Trans.* 26, 513–518 (2002)
- Breuer, D., Moore, W.B.: Dynamics and thermal history of the terrestrial planets, the Moon, and Io. In: Spohn (ed.) *Planets and Moons*, Treatise on Geophysics, vol. 10, pp. 299–340. Elsevier, Amsterdam (2007)
- Carr, M.H., Wänke, H.: Earth and Mars: water inventories as clues to accretional histories. *Icarus* 98, 61–71 (1992)
- Rotron, C.: Establishing cooling requirements: air flow vs pressure (2009), http://www.comairrotron.com/airflow_note.shtml (last accessed, 2009-6-28)
- Dombard, A.J., Phillips, R.J.: Tectonic evidence for crustal underplating of the Tharsis Montes, Mars. *Lunar and Planetary Science Conference XXXVI*, Lunar and Planetary Science Institute, Houston, abstract 1878, 2 p. (2005)
- DiPippo, R.: Binary cycle power plants. In: *Geothermal Power Plants*, 2nd edn., ch. 8, pp. 157–190. Elsevier, Amsterdam (2008)
- Fogg, M.J.: The utility of geothermal energy on Mars. *J. British Interplanetary Soc.* 49, 403–422 (1996)
- Fuller, E.R., Head, J.W.: Amazonis Planitia: the role of geologically recent volcanism and sedimentation in the formation of the smoothest plains on Mars. *J. Geophys. Res.* 107, E01842 (2002)
- Grott, M., Breuer, D.: The evolution of the martian elastic lithosphere and implications for crustal and mantle rheology. *Icarus* 193, 503–515 (2008)
- Guest, A., Smrekar, S.E.: New constraints on the thermal and volatile evolution of Mars. *Phys. Earth Planet. Int.* 164, 161–176 (2007)
- Hartmann, W.K.: Martian cratering 8: isochron refinement and the chronology of Mars. *Icarus* 174, 294–320 (2005)
- Hartmann, W.K.: Martian cratering 9: Toward resolution of the controversy about small craters. *Icarus* 189, 274–278 (2007)
- Hartmann, W.K., Malin, M., McEwan, A., Carr, M., Soderblom, L., Thomas, P., Danielson, E., James, P., Veverka, J.: Evidence for recent volcanism on Mars from crater counts. *Nature* 397, 586–589 (1999)
- Hartmann, W.K., Neukum, G., Werner, S.: Confirmation and utilization of the “production function” size-frequency distributions of martian impact craters. *Geophys. Res. Lett.* 35, GL031557 (2008)
- Hauck, S.A., Phillips, R.J.: Thermal and crustal evolution of Mars. *J. Geophys. Res.* 107, E01801 (2002)

- Head, J.W., Wilson, L.: Mars: a review and synthesis of general environments and geological settings of Magma-H₂O interactions. In: Smellie, J.L., Chapman, M.G. (eds.) *Volcano-Ice Interactions on Earth and Mars*. Geol. Soc. Lond. Spec. Pub., vol. 202, pp. 27–57 (2002)
- Head, J.W., Wilson, L.: Heat transfer in volcano-ice interactions on Mars: synthesis of environments and implications for processes and landforms. *Annals Glaciology* 45, 1–13 (2007)
- Hess, S., Henry, R.M., Leovy, C., Ryan, J., Tillman, J.E.: Meteorological results from the surface of Mars: Viking 1 & 2. *J. Geophys. Res.* 82, 4559–4574 (1977)
- Head, J.N., Melosh, H.J., Ivanov, B.A.: Martian meteorite launch: high-speed ejecta from small craters. *Science* 298, 1752–1756 (2002)
- Kanoğlu, M.: Energy analysis of a dual-level binary geothermal power plant. *Geothermics* 31, 709–724 (2002)
- Kanoğlu, M., Çengel, Y.A.: Improving the performance of an existing air-cooled binary geothermal power plant: a case study. *J. Energy Resources Tech.* 121 (1999) 10.1115/1.2795982
- Korenaga, J.: Energetics of mantle convection and the fate of fossil heat. *Geophys. Res. Lett.* 30, GL016982 (2003)
- Korenaga, J.: Urey ratio and the structure and evolution of Earth's mantle. *Reviews Geophys.* 46, RG000241 (2008)
- Langseth, M.G., Keihm, S.J., Peters, K.: The revised lunar heat flow values. In: *Lunar and Planetary Science Conference VII*. Lunar and Planetary Science Institute, Houston, pp. 474–475 (1976)
- Lodders, K., Fegley, B.: An oxygen isotope model for the composition of Mars. *Icarus* 126, 373–394 (1997)
- Malin, M.C., Edgett, K.S., Posiolova, L.V., McColley, S.M., Debrea, E.Z.N.: Present-day impact cratering rate and contemporary gully activity on Mars. *Science* 314, 1573–1577 (2006)
- McDonough, W.F., Sun, S.-S.: The composition of the Earth. *Chem. Geol.* 120, 223–253 (1995)
- McGovern, P.J., Solomon, S.C., Smith, D.E., Zuber, M.T., Simons, M., Wiezorek, M.A., Phillips, R.J., Neumann, G.A., Aharonson, O., Head, J.W.: Localized gravity/topography admittance and correlation spectra on Mars: implications for regional and global evolution. *J. Geophys. Res.* 107, E01854 (2002)
- McGovern, P.J., Solomon, S.C., Smith, D.E., Zuber, M.T., Simons, M., Wiezorek, M.A., Phillips, R.J., Neumann, G.A., Aharonson, O., Head, J.W.: Correction to Localized gravity/topography admittance and correlation spectra on Mars: implications for regional and global evolution. *J. Geophys. Res.* 109, E02286 (2004)
- Montési, G.J., Zuber, M.T.: Clues to the lithospheric structure of Mars from wrinkle ridge sets and localization instability. *J. Geophys. Res.* 108, E01974 (2003)
- Morgan, P.: The thermal structure and thermal evolution of the continental lithosphere. *Phys. Chem. Earth* 15, 107–193 (1984)
- Neukum, G., Jaumann, R., Hoffmann, H., Hauber, E., Head, J.W., Basilevsky, A.T., Ivanov, B.A., Werner, S.C., van Gasselt, S., Murray, J.B., McCord, T.: The HRSC Co-Investigator Team, Recent and episodic volcanic and glacial activity on Mars revealed by the High Resolution Stereo Camera. *Nature* 432, 971–979 (2004)
- Nyquist, L.E.: Martian meteorite ages and implications for martian cratering history. In: *Planetary Chronology Workshop*, 2 p. Lunar & Planetary Institute, Houston (2006) abstract 6010

- Phillips, R.J., et al.: Mars north polar deposits: Stratigraphy, age, and geodynamical response. *Science* 320, 1182–1185 (2008)
- Plescia, J.B.: Recent flood lavas in the Elysium region of Mars. *Icarus* 88, 465–490 (1990)
- Pollack, H.N., Hurter, S.J., Johnston, J.R.: Heat loss from the earth's interior: analysis of the global data set. *Rev. Geophys.* 31, 267–280 (1993)
- Solomon, S.C., Head, J.W.: Heterogeneities in the thickness of the elastic lithosphere of Mars: constraints on heat flow and internal dynamics. *J. Geophys. Res.* 95, 11073–11083 (1990)
- Stein, C.A.: Heat flow of the Earth. *Global Earth Physics, A Handbook of Physical Constants*. Am. Geophys. Un., AGU Reference Shelf 1, 144–158 (1995)
- Tester, J.W., Anderson, B., Batchelor, A., Blackwell, D.D., DiPippo, R., Drake, E., Garnish, J., Livesay, B., Moore, M.C., Nichols, K., Petty, S., Toksoz, N., Veatch, R., Augustine, C., Baria, R., Murphy, E., Negruaru, P., Richards, M.: The future of geothermal energy: impact of enhanced geothermal systems (EGS) on the United States in the 21st century, 374 p. Massachusetts Institute of Technology, DOE Contract DE-AC07-05ID14517 Final Report (2006)
- Treiman, A.H., Gleason, J.D., Bogard, D.D.: The SNC meteorites are from Mars. *Planet. Space Sci.* 48, 1213–1230 (2000)
- Vaucher, J., Baratoux, D., Pinet, P.C., Mangold, N., Ceuleneer, G., Gregoire, M., Daydou, Y., Chevrel, S., Neukum, G., HRSC Co-Investigator Team: Cerberus Plains, Mars: chronology of volcanic event and evidence of recent tectonic activity. In: *Lunar and Planetary Science Conference XXXVII*, Lunar and Planetary Science Institute, Houston, abstract 1851, p. 2. (2006)
- Wänke, H., Dreibus, G.: Chemistry and accretion history of Mars. *Phil. Trans. R. Soc. Lond. A* 349, 285–293 (1994)
- Weizman, A., Stephenson, D.J., Prialnik, D., Podolak, M.: Modeling the volcanism on Mars. *Icarus* 150, 195–205 (2001)
- Zaranek, S.W., Manga, M.: Mantle plumes and long-lived volcanism on Mars a result of mantle compositional layering. In: *Lunar and Planetary Science Conference XXXVIII*, 2 p. Lunar and Planetary Science Institute, Houston (2007) abstract 2133

Chapter 12

Comparative Analysis of Power System Architectures: The Case of Human Mars Surface Missions

Matthew Silver, Wilfried Hofstetter, Chase Cooper, and Jeffrey Hoffman

IntAct Labs LLC and Massachusetts Institute of Technology, Cambridge, MA

12.1 Introduction

Human missions to the surface of Mars have inspired engineers, scientists, and the wider public for generations. The 20th century has seen these dreams become attainable, with numerous realistic missions proposed, starting with the architecture proposed by Wernher von Braun in his 1953 book “The Mars Project” (von Braun 1961). Many architectures and concepts for carrying out surface missions have since been proposed, culminating in the Mars design reference missions developed by NASA in the 1990s (Hoffman 1997, Drake 1998, Hoffman 2001, Portree 2001, Drake 2007) and associated follow-on studies as part of the US Vision for Space Exploration (NASA 2005, Drake 2007). Two main motivations for carrying out Mars surface missions have been described in these architecture studies: (1) the scientific exploration of Mars, in particular with regard to extraterrestrial life, and (2) the investigation of the habitability of Mars in the context of establishing a long-term human presence there someday in the future.

Power systems are central to the design of Mars surface missions given that almost every subsystem (perhaps with the exception of structures) is completely dependent on power supply. The architecture of these systems can be defined as a network of complimentary technologies operating together to perform a meta-function (see Sect. 12.2). Yet, the majority of design studies carried out in support of the above-mentioned architecture studies focus largely on individual power systems. McKissock and Kohout (1990) investigate the use of a power system for opposition class missions based on non-tracking thin-film roll-out solar arrays and regenerative fuel cells. Kerslake and Kohout (1999) provide a study of a power system for conjunction class missions based on thin-film tent-like solar arrays and

regenerative fuel cells. Withrow and Morales (1994) provide a comparative analysis of solar-electrochemical power system concepts based on tracking and non-tracking photovoltaic arrays as well as a variety of energy storage technologies. Surface nuclear fission system point designs are described in the above design reference mission studies (Hoffman 2001, Drake 2007) together with point designs for auxiliary surface photovoltaic systems. One comprehensive comparative analysis of different Mars surface power system concepts and associated technologies (including fission-based systems as well as RTG-based technologies) is provided in (Littman 1994).

This cursory review illustrates that most surface power analyses assume a fixed architecture. Work on comparative analysis of many power system architectures for the same mission requirements and assumptions is significantly more limited. This chapter focuses on the latter type of analysis and associated investigations into the sensitivity of particular architecture concepts to changes in surface site location and power utilization patterns by the end user.

Section 12.2 provides a short introduction into architecture analysis methodology for complex systems. In Sect. 12.3, a set of technologies that are applicable to Mars surface power systems are described and metrics for architecture evaluation are discussed. Section 12.4 provides results from comprehensive architecture enumeration and evaluation and sensitivity analysis of the Mars surface power system concepts. In Sect. 12.5 future trends with regard to power generation and energy storage technologies are outlined; the discussion includes developments outside the aerospace sector. Section 12.6 summarizes important insights and findings and provides recommendations for future Mars surface power system analyses.

12.2 Architecture Analysis Methodology

This section reviews methods for architecture-level comparison and evaluation. The term “system architecture” refers to the kinds of subsystems used (functions performed) within a complex system and their associated interrelations (Ulrich 1995). The architecture can thus be fully defined by the performance/cost limits of the subsystems together with information pertaining to their interconnection. Though less informative than a detailed design, such a comparison can provide important insights, including regions of the design and performance space inhabited by different combinations of technologies.

A number of methodologies and tools have been developed for the study and comparison of system architectures over the years. Object-Process Methodology represents architectures using two classes of entities – those that performance functions (objects) and the functions themselves (processes) (Dov and Edward 1999). Koo has developed an executable meta-language based on OPN which could theoretically be used to compare power-systems automatically using the Object Process Methodology (Simmons et al. 2005). Methods for automatic architecture comparison have been developed. Graff and de Weck, for example, develop a method similar to Executable OPN, based on the concept of representing a system’s “operand” using a state-vector which is passed among various subsystem

models. (Graff and de Weck 2006). An operand, in this case, is defined as the object or process altered by the system to generate value. Recently, Simmons has developed a methodology for modeling the architecting of complex systems as a set of interrelated decisions (Simmons 2008). His Architecture Decisions Graph (ADG) allows for comprehensive analysis of the space of feasible decision variable assignment and subsequent evaluation of system metrics and global sensitivity to changes in analysis assumptions.

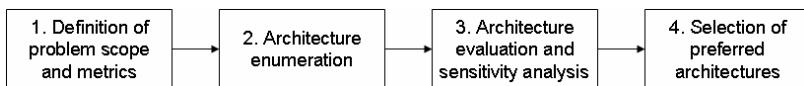


Fig. 12.1 Meta-process for architecture analysis of complex systems

In this chapter, we will use a simplified process for architectural analysis which can incorporate these and other related methodologies. Figure 12.1 shows a meta-process for architecture analysis of complex systems which was developed based on experience gained in the practice of aerospace system architecting (Hofstetter 2006, Buonova 2005). The process consists of four steps: the first step involves clearly bounding the scope of the architecture analysis in terms of functionality considered and interfaces to other systems as well as the metrics to be used for the evaluation of alternative architectures. Step 2 is devoted to enumerating architecture alternatives which are feasible from the perspective of logical, technical, and operational constraints. In Step 3, the feasible architectures are evaluated with regard to the metrics defined in Step 1 and sensitivity analyses with regard to requirements and analysis assumptions are carried out as needed. Step 4 is the selection of preferred architectures for further detailed analysis during following design phases. It should be noted that typically several iterations through the meta-process are required in order to arrive at a set of robust preferred architectures.

This meta-process is an efficient means of identifying a set of robust preferred architectures during the early stages of the design as input to more detailed design phases. In the sections below we follow the application of this meta-process for Mars surface power system architectures.

12.3 Technology Overview and Metric Definition

This section corresponds to the implementation of Step 1 of the above meta-process for Mars surface power systems for human missions. The scope of the analyses includes power generation, energy storage, power management and distribution and any thermal control and structural support associated with power generation and energy storage. Metrics suitable for the evaluation of surface power system alternatives are discussed at the end of this section. The following is an overview of the different power generation and energy storage technologies considered in the architecture-level analysis. Performance assumptions and references are provided where possible. Specific technologies for the architectures

were researched in order to ascertain their level of readiness. This includes a number of RTG technologies currently being developed by the NASA Science Mission Directorate (Shaltens 2007), traditional rigid solar arrays (tracking), and deployable thin film arrays (non-tracking).

12.3.1 Photovoltaic Power Generation

Two technologies were considered here; they included ultra-light thin-film amorphous silicon rollout blanket arrays and high efficiency inflexible tracking arrays. The ultra-light thin-film arrays have efficiencies of 15 % and a mass/area of 0.063kg/m^2 (Hanak 1986). Due to their much higher optical absorption, thin-film amorphous silicon solar cells are less than a micrometer in thickness. These arrays have only been tested as small units so the technology readiness level for a large system that would be required for human surface exploration would be lower than that for already existent inflexible array systems. The high efficiency inflexible tracking arrays are based on ISS array technology; they have 20 % efficiencies and mass/area of 2.5 kg/m^2 . The structural overhead for tracking and structural integrity of these arrays is based on ISS design data. Also, multi-axis tracking was assumed for perpendicular solar flux incidence throughout the day.

An important added consideration for the ultra-light thin-film arrays is how to protect the rolled blanket from high winds. It was found that if the blankets are simply laid on the surface without any additional anchoring, a light wind of only 7.35 m/s would lift the arrays. Therefore a concept was developed to secure the arrays to the ground by adding Kevlar areas equal to 10% or the total array area in which rocks will be placed (see Fig. 12.2). It was found that 9.2 kg/m^2 of rock is needed in the 10 % Kevlar regions to secure the array against the top recorded Mars wind of 25 m/s. The major effect of this consideration is increased deployment time which will be discussed below. Note: this array securing system is only required for the thin-film array systems; the structure associated with the inflexible tracking arrays protects against any dislocation due to wind forces.

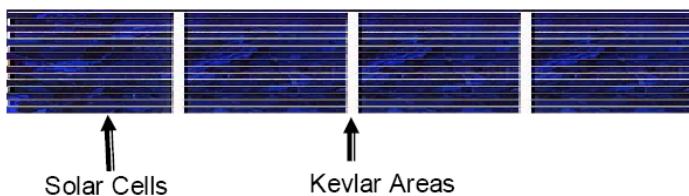


Fig. 12.2 Ultra-light blanket arrays with Kevlar portions for rock placement

12.3.2 Energy Storage Using Secondary Batteries

Batteries are generally employed for both secondary power generation and for energy storage. Li-ion batteries were considered in this study for their high energy density and common use in aerospace systems. To be conservative, current performance numbers were used. The batteries have a mass-specific energy density of

150 Wh/kg and a volume-specific energy density of 270 kWh/m³. It is interesting to note that battery energy densities of 200 Wh/kg or more are expected for the 2020 timeframe and would therefore be available for human Mars missions (Yoder 2007).

12.3.3 Energy Storage Using Regenerative Fuel Cells

Regenerative fuel cells can perform both the tasks of secondary power generation and energy storage (see Chap. 6). NASA has to date used both Alkaline Fuel Cells and Proton Exchange Membrane (PEM) fuel cells. This analysis considers hydrogen/oxygen Proton Exchange Membrane fuel cells as a reference due to their superior performance in comparison to the older alkaline fuel cells. The fuel cells are assumed to have a mass-specific energy density of 700 Wh/kg and volume-specific energy density of 200 kWh/m³ (Burke 2003). It was assumed that the reactants were stored in tanks at 200 atm pressure which allowed for analysis of needed tank masses that went into the above obtained performances.

12.3.4 Nuclear Fission Power Generation

Two designs were considered for nuclear primary power production in this study. Both are nuclear reactors with dynamic conversion. One design uses a Brayton engine for the conversion and the other a Stirling engine. The Brayton based design is adapted from the NASA Prometheus design for a lunar based reactor (Elliot 2006). The radiator was resized for use in the Martian thermal environment thus slightly reducing mass of the overall system. The Brayton design must be located 210 m from base and have a 3.5 m effective regolith shield to mitigate radiation exposure. The Stirling engine based design comes from JSC element/systems database (Johnson Space Center 1991). It is composed of an SP-100 type reactor and 4 Stirling engines. The Stirling design must be located 1 km from the base and the reactor itself must be located below the surface, potentially requiring significant infrastructure for excavation and moving of regolith.

12.3.5 Radio-Isotope-Based Power Generation

Dynamic conversion RTG systems can act as secondary power generation elements as well as provide a redundant constant power source for added safety in the power system. Here we considered a design for modular General Purpose Heat Sources (GPHS) coupled to Stirling conversion engines. This design has a mass specific power of 13.75 W/kg and volume specific power of 27500 W/m³ (Schmitz 2005). These units use PuO₂ for fuel; a 5 kW unit would require 62.5 kg of fuel. A positive feature of this design is that they primarily have alpha-radiation emissions that can be easily shielded against and thus these units could be located close to base and do not require any kind of auxiliary equipment for excavation or regolith moving for radiation shielding.

12.3.6 Metrics for Comparative Evaluation of Power System Architectures

For any systems architecture level analysis, three types of metrics need to be taken into account: (1) performance metrics, capturing how well a system concept delivers its functionality; (2) cost, capturing the non-recurring and recurring resource expenditure required to deliver the system functionality; and (3) risk, capturing the likelihood of successfully developing, producing, and operating the system (note: risks may not only be technical in nature but may also include risks associated with political support). Schedule is often considered as a fourth metric, but is usually closely coupled with cost and therefore not a fully independent metric.

The performance metrics considered for the surface power system architecture analysis are total power systems mass and volume, captured in normalized form as (average power generation over the course of the day) / (total power system mass) with units of [W/kg] or as (average power generation over the course of the day)/³(total power system volume) in units of [W/m]. In addition, the effort of deploying the surface power system prior to routine operations is considered as a secondary metric.

Cost, both in terms of up-front development cost as well as life-cycle cost including development, production, as well as transportation costs are important metrics to be considered. As the fidelity of estimates at this early stage is very limited, and preliminary estimates of development and production cost tend to be based on system mass, we will just use system mass as a substitute for cost, thereby reducing the number of metrics we need to consider. Transportation cost is, of course, also primarily driven by power system mass.

Risks can be associated with the development of specific technologies used as part of the system concept, as well as with the production of system units and their operation during the mission. For Mars surface power systems, the use of radioactive material must be considered one of the most significant architectural distinguishers with regard to development, production, and operations risk due to the potentially significant impact of failures on the program, as well as due to the potential political opposition to use of nuclear technology. We will discuss other risks in a qualitative fashion together with the analysis results.

12.3.7 Mars Surface Power System Modes of Operation

The evaluation of power system mass and volume is critically dependent on the operational mode of the system. Two general modes of operation are conceivable:

- The system provides constant continuous power to the end user over the course of the Martian day, and indeed over the course of the entire surface stay. In this case average power and momentary power generation are identical.
- The power usage pattern is optimized such that given minimum power constraints the power system mass for each system is minimized, i.e. the power is used and generated when it causes the minimum power system mass requirement within operational constraints (manifest in minimum continuous power requirements through the Martian night). In this case the power system still provides the same amount of usable energy over the course of a Martian day as a continuous power system would.

For nuclear-fission-based power systems, there is practically no distinction between the two operating modes because nuclear reactors provide approximately constant power over their life-cycle. For power systems incorporating photovoltaic power generation and secondary battery or RFC energy storage, however, there is a significant difference between the two operational modes: a requirement for equal, continuous power generation / consumption during the Martian day and night would necessitate heavy energy storage systems which would dramatically increase the normalized metrics described above. The ability to move the majority of power generation / consumption to the Martian day when power generation is cheap and no energy storage is required has the potential to radically improve the metrics values for power systems based on photovoltaics, while still providing the same usable energy during the course of a Martian day as a continuous power system would. We will therefore assume the second mode of operations as the basis for the evaluation of power system architecture alternatives for the analysis presented in this chapter.

12.4 Power System Architecture Enumeration, Evaluation and Sensitivity Analysis

The work presented in this section corresponds to the implementation of Step 2 and Step 3 of the above meta-process for Mars surface power system architecture analysis. An enumeration of architectural options was carried out based on three architectural variables: the choice of daytime power generation technology, the choice of eclipse power generation technology, and the energy storage technology (if required); constrained enumeration yields the alternatives shown in Table 12.1.

Table 12.1 Architecture alternatives for Mars surface power production

Architecture	Primary Power	Secondary Power	Energy Storage
1	Nuclear fission - Stirling cycle	N/A	N/A
2	Nuclear fission - Brayton cycle	N/A	N/A
3	Photovoltaics - tracking	N/A	Li-Ion batteries
4	Photovoltaics - tracking	N/A	Regen FC
5	Photovoltaics - tracking	Dynamic RTG	Li-Ion batteries
6	Photovoltaics - tracking	Dynamic RTG	Regen FC
7	Photovoltaics - tracking	Dynamic RTG	N/A
8	Photovoltaics - non-tracking	N/A	Li-Ion batteries
9	Photovoltaics - non-tracking	N/A	Regen FC
10	Photovoltaics - non-tracking	Dynamic RTG	Li-Ion batteries
11	Photovoltaics - non-tracking	Dynamic RTG	Regen FC
12	Photovoltaics - non-tracking	Dynamic RTG	N/A
13	Dynamic RTG	N/A	N/A

Note: for architectures where primary power generation is based on photovoltaic arrays, there is an option for using radioisotope heat sources with thermoelectric or thermodynamic (“dynamic”) power conversion to supply part or all of the nighttime power; these options also may have different characteristics for contingency operations (e.g. during a global Martian dust storm), because RTG-based architectures are to some degree independent of sunlight and the intensity of insolation.

As can be seen in Table 12.1, the work in this chapter is focused on nuclear and photovoltaic power production architectures then with different options for secondary power generation and energy storage.

In order to compare all the architectures seen in Table 12.1 on a quantitative basis, a model was created to asses mass and volume required to proved sufficient power through the Martian day and night. The nuclear options were modeled directly from reference data available in the literature (see above). The solar power options, however, required the creation of a new parametric model. The major requirements driving this model are as follows. The arrays must be sized for end-of-mission power requirements. Arrays must also be sized to provide the required power during the year’s minimum incident solar energy period.

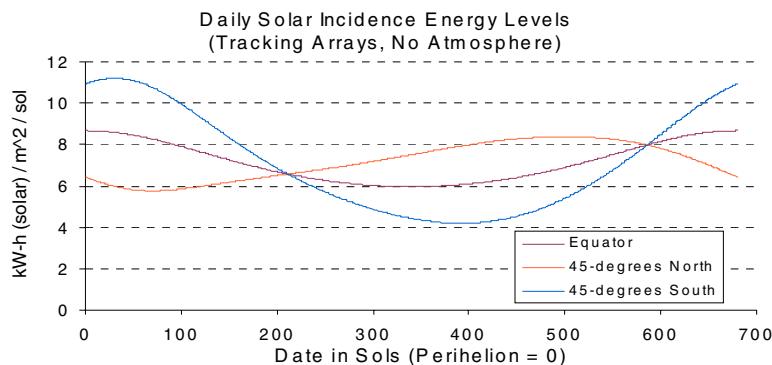


Fig. 12.3 Mars solar incidence energy levels for three latitudes

The model also includes a number of important assumptions. An optical depth of 0.4 is assumed which is equivalent to hazy skies on Mars (the temporary impact of global dust storms which can lead to optical depths as high as 5 on solar-based architecture alternatives are discussed with the analysis results below). Solar arrays are assumed to be cleaned regularly (about once every month) by the crew in order to prevent the accumulation of dust and the associated degradation of solar array efficiency over time (the dust devils observed with the Mars exploration rovers would aid in cleaning the arrays). Tracking arrays are multi-axis and keep incident flux perpendicular to array over the day (for a detailed analysis see Chap. 3). A nighttime power of 12 kW is assumed to be sufficient to sustain 6 crew members. The daytime power requirement is not enforced until the sun is 12 degrees above the horizon. Also, initial analysis for all architectures was done for an equatorial location which is actually not the optimal location for solar power on Mars. Figure 12.3 shows the

daily solar incidence levels over time for three different latitudes. It is seen that some northern latitudes actually have a higher minimum solar incidence over the year. In fact 31 degrees northern latitude has the highest minimum incident energy compared to the rest of Mars (for a more detailed analysis see Chap. 2).

An initial analysis was performed for each architecture for an equatorial location on Mars for a variety of average power levels. For this initial analysis it was assumed that the system would be used for a conjunction class mission without in-situ propellant production to an unprepared site. After this initial analysis, the more feasible architectures were then looked at in the context of performance change as a function of latitude location (sensitivity analysis). The steps taken in the modeling process are outlined below in Table 12.2.

Table 12.2 Procedure for analyzing Mars surface power system architectures

Description	
Step 1	Calculate total energy in Joules that must be produced by the solar arrays in a day based on the day's power requirement
Step 2	Calculate the energy density (J/m^2) produced by the solar array with the sun at the given latitude (Minimim solar energy based on the array's end-of-life characteristics)
Step 3	Integrate to find the total energy/ m^2 produced by an array per day
Step 4	Divide results from Step 1 by results from Step 2 to find required array area
Step 5	Calculate resulting mass and volume for array
Step 6	Caclulate mass and volume of secondary energy components based on night-tine energy requirements

Results show that architectures which include thin film rollout solar arrays and either RFCs or Li-ion batteries can be competitive with nuclear based options. Architectures with RFCs come especially close to matching the mass based performance of nuclear reactors with Stirling engines for dynamic conversion at higher power levels (see Fig. 12.4). This is true at higher power levels because the ultra-light solar arrays begin to dominate the more massive secondary power generation components. Solar-based architectures with regenerative fuel cell energy storage start to dominate fission-based architectures at about 25 kWe of average power generation capability; it is reasonable to assume that any human Mars mission with 4-6 crew members would require at least that much average power generation capability for crew habitation, surface mobility, and communications (Hoffman 1997, Drake 1997).

Looking at volume based performance it is seen that all thin film solar architectures dominate the nuclear options (see Fig. 12.5). All tracking array architectures are non-competitive on both a mass and volume basis. All solar based options were also included in architectures where 5 kW RTGs were included. These architectures see a slight performance boost over their non-RTG counterparts, but the performance increase is small and the major benefit of the RTG is still the added safety that

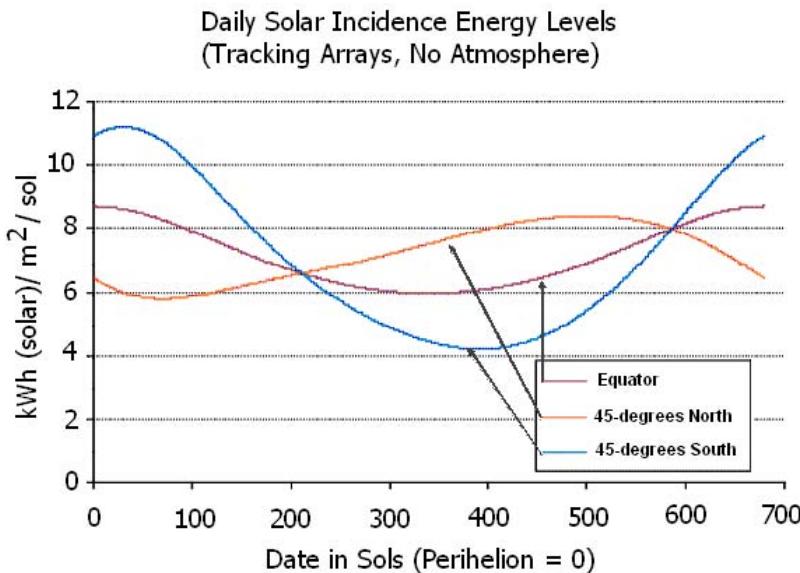


Fig. 12.4 Mass specific power performance versus average power level for all architectures

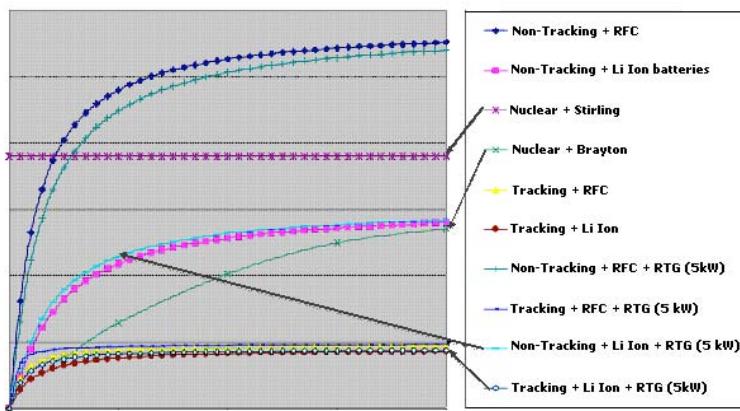


Fig. 12.5 Volume specific power performance versus average power level for all architectures

a continuous power supply imparts. Figure 12.6 gives a 100 kW (power averaged over the Martian day) point design comparison for the competitive architectures.

In addition to the analyses for nominal operating conditions with an optical depth of 0.4, the impact of strongly increased optical depth due to global dust storms on solar power generation was assessed based on solar array efficiency reduction data obtained from the Mars exploration rovers (see also Chap. 3). This

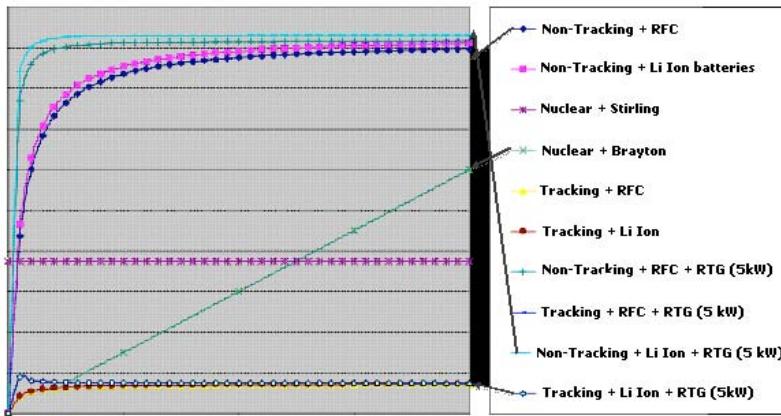


Fig. 12.6 Mass and volume specific power performance for a 100kW average power system for the feasible architecture

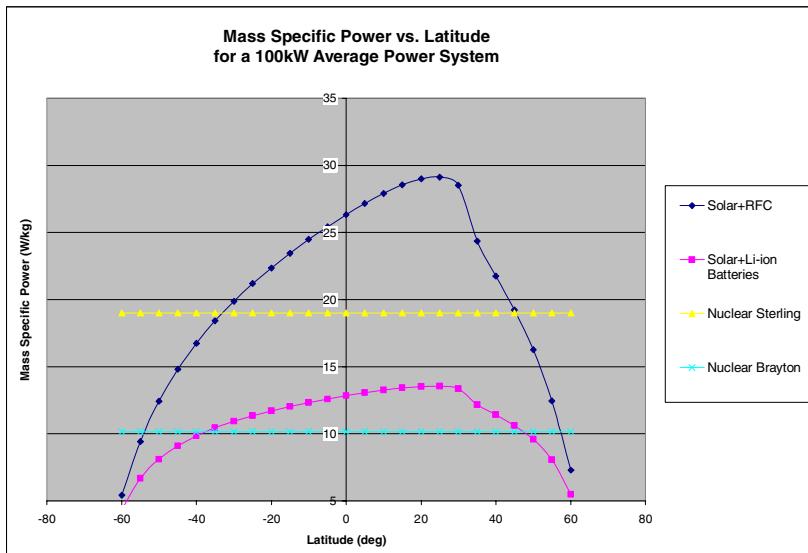


Fig. 12.7 Mass specific power performance for interesting architectures as a function of latitude

assessment indicated that for average power generation capabilities in the vicinity of 100 kWe, solar-based architecture alternatives provide sufficient power generation even under global dust storm conditions to sustain crew activities at a reduced

level. For lower average power generation capabilities (25-50 kWe), radioisotope-based supplementary power sources could be used to support a solar-based system under global dust storm conditions.

Now that thin film solar architectures with RFCs or Li-ion batteries have been singled out as the interesting competitive architectures with nuclear options, it is interesting to look at the effect of latitude location on the power systems' performance. This way, more suitable locations for solar based architectures can be assessed. Taking in the planet's axial tilt and orbital elements about the sun, the minimum solar energy flux based on latitude can be found. Figures 12.7 and 12.8 then present the mass and volume based performance of the power architectures for a range of Mars latitudes. The results show that there is an optimum location for solar architectures around 30 degrees north. The results also show that northern latitudes are always better than their southern counterparts.

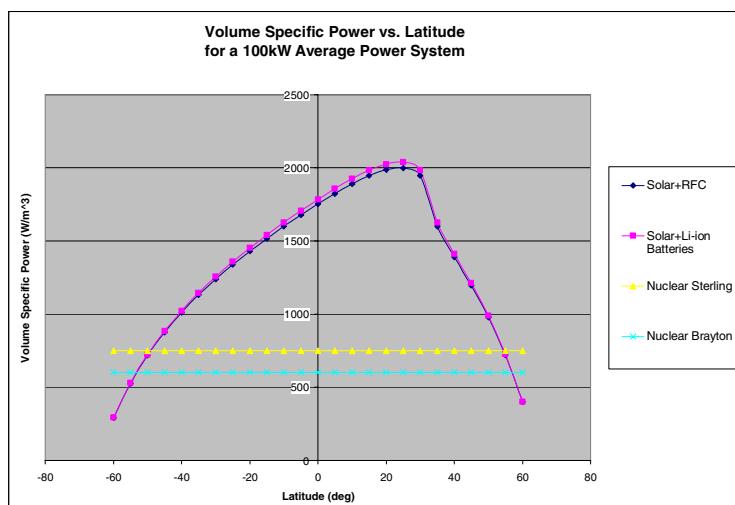


Fig. 12.8 Volume specific power performance for interesting architectures as a function of latitude

Aside from mass and volume based performance, deployment time of these very large arrays ($25,000 \text{ m}^2$ for a 100 kW equatorial system) is very important. Deployment time includes off-loading of the arrays, unrolling the arrays, and finally placing rocks to weigh down the arrays. For this analysis we considered the 100 kW average power system located at the Mars equator in order to get an estimate for deployment time. This requires a $25,000 \text{ m}^2$ rollout array field which includes the addition of the Kevlar areas for wind mitigation. It was assumed that array blankets are 2m wide and weigh ~40 kg for easy storage and handling by two astronauts. With 0.07 kg/m^2 as the expected array density, only 18 blankets are required.

If we assume astronauts can unroll arrays at a walking speed of 1 m/s, the unrolling requires 7 hrs. Time will also be needed for unloading, positioning, and

hookup of arrays. If it is assumed that 1 hr for this for each array is needed, this adds 18 hrs. In addition to this rocks must be placed in the Kevlar areas. Assuming Kevlar areas are 0.3 m in length and the complete 2 m width, 5.6 kg of rock in each area is needed. There are 225 of these Kevlar areas per array so a total of 4050 of these areas. Assuming 2 rocks are needed per area to secure the 2 sides of the array this requires 8100 rocks to be placed. If 30 seconds are needed to pick and place a rock the deployment of rocks of the entire array area will take 33.75 hrs for 2 crew members. All of these results in a total of 66 hrs to deploy the solar array field by two crew members.

Power must also be provided during the deployment process. It is interesting to note, however, that deployment gives 0.76 kW per man hour; therefore we only need 13.2 person-hours to reach a capability of 10 kW which is assumed to be enough for minimal stay-alive power. To be very conservative, we can neglect this and find out what additional fuel cells or batteries are needed to get through the deployment period. If we assume that full deployment and initial usefulness takes 1 week, we need either a 10 kW RTG or fuel cell system to provide 10 kW power over the week. The RTG system would be approximately 1200 kg and 0.6 m³. A RFC system would need 2400 kg system with volume 8.4 m³. This is overly conservative however, and in fact little more than fully charged night-time power generation would be required as 2 crew members could achieve the needed 10 kW in less than 7 hrs.

Alternatively, robotic deployment of the arrays could be utilized; this would also enable the use of a solar-based power architecture for in-situ propellant production prior to Earth departure of the crew. It is interesting to note that a power system based on fission surface power generation would always require significant robotic pre-deployment, both for transmission cables from the reactor to the human-tended surface assets as well as for deployment of radiator systems. The need for robotic pre-deployment of the Mars surface power system components for in-situ resource utilization is therefore not a significant drawback of solar-based systems compared to fission-based systems.

12.5 Future Technologies and Developments

The previous analysis was based on readily available technologies. With the significant current investments into energy technologies, it is likely that new systems will soon disrupt the underlying assumptions of this model. At a high level, it has been observed that energy storage - like many technologies - demonstrate an exponential increase in functionality over time. Figure 12.9 is based on a comprehensive study of functionality of energy technologies over time, taken from Koh and Magee (Koh and Magee 2008). Beyond illustrating that power density, in general, is improving at a low exponential rate, the authors observe that improvements in capacitors and fly-wheels seem likely to outstrip traditional batteries in the coming decade.

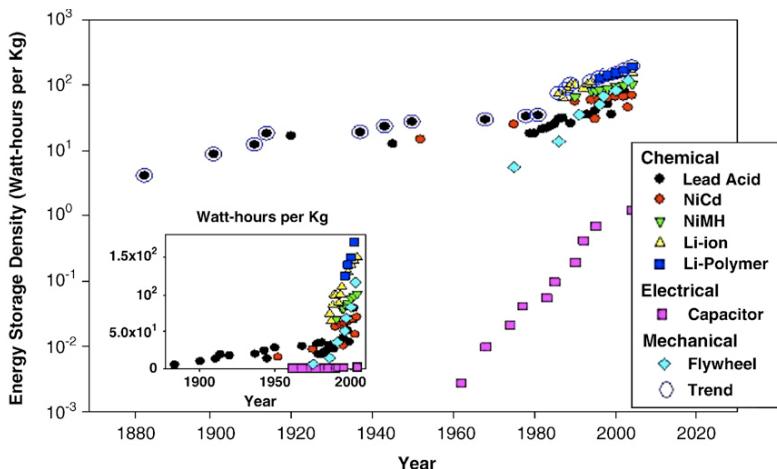


Fig. 12.9 Energy Storage Density through time (source: Koh, 2006)

This prediction seems to be supported by recent developments in the field. In particular, ultra-capacitors based on engineered nano-materials such as carbon nanotubes or graphene show promise. (Science Daily, 2008).

Beyond inorganic nano-materials, a number of very interesting bio-electric technologies may be on the horizon. Such systems, based on inert organic material such as proteins or living systems, might have significant advantages compares to today's energy technologies. This includes self-assembly, self-repair, and synergies with non-energy related base functions such as life-support or structural support. For example, a number of recent advances suggest that solar cells based on photo-active proteins are theoretically feasible. In 2003, Dr. Marc Baldo at MIT proved that by tethering the photo-active protein Photo-system II to a gold substrate using peptide bonds, current could be generated continually for months. More recently, claims have been made that exo-cellular protein life can be extended significantly. Along a different dimension, research is increasing into the potential to generate power using microbial fuel cells. These devices, which utilize newly discovered electron transfer mechanisms in a new class of bacteria known as electricigens, have the potential to generate electricity while treating organic wastes. While power outputs are currently low (theoretical max of approximately 1KW/m^3 of reactor), the use of waste organics as fuel results in significantly higher power/weight ratios in a life-cycle analysis. Early stage research into the application of such technologies to space exploration was initiated by the NASA Institute for Advanced Concepts in 2006 (Tenenbaum 2007).

These forecasted improvements in energy storage capability are paralleled by improvements in the capability of photovoltaic power generation. This indicates that significant improvement of system mass and volume that are expected from multi-junction solar cells and similar developments outside the aerospace sector. Instead of a premature down-selection to a few preferred surface power system architectures, the comparative analysis of a large space of Mars surface power system

architecture should therefore be continued in order to capture these anticipated trends in technology development.

12.6 Conclusions

Most analyses of power systems for Mars surface exploration to date have focused narrowly on individual surface power system architectures and technologies, with some sensitivity to variations in ancillary technology. This chapter outlines a process for the comprehensive analysis of Mars surface power system alternatives for human exploration missions; the process was applied to the comparative quantitative analysis of 13 architecture alternatives.

From this application of the analysis process we can draw conclusions related to the utility of comprehensive analysis as well as to the strengths and weaknesses of the individual architecture and technology alternatives and their conditions of use.

The following are conclusions related to the utility of the analysis process:

- The enumeration, and evaluation of a variety of system architecture alternatives coupled with sensitivity analysis with regard to driving requirements and assumptions is an effective means for identifying preferred Mars surface power system concepts and understanding the conditions under which the alternatives are preferred.
- As our understanding of the requirements for Mars missions becomes more detailed, and specific technologies evolve to greater maturity, it becomes increasingly important to broaden the design analysis beyond point systems and beyond point missions. Comparative architectural analysis as described above is one way to do this rigorously.
- The preceding conclusion becomes even more pertinent if one considers trends in power system technologies. Ultra capacitors, thin-film solar arrays, cutting-edge biological solutions, are all developing at very rapid rates outside of the aerospace sector (and would therefore be available to NASA largely for “free”). The design of a surface power system architectures that could be used ten to twenty years from now and could remain operational for a decade or more should therefore incorporate expectations with regard to the rate of technology improvements. More generally, any architecture designed to last a number of years should be modular enough to allow low-cost switching of time-variable components.

The following conclusions are related to specific system architecture alternatives (drawn within the boundaries of the analysis assumptions):

- Architecture alternatives with thin-film solar arrays are superior in performance to the corresponding architecture alternatives with tracking solar arrays.
- For northerly locations, and for mission architectures that require no in-situ propellant production (i.e. no autonomous deployment of the power system required prior to crew arrival), architecture alternatives based on thin-film solar arrays are mass- and volume competitive with architecture alternatives based on fission

power generation. The higher the power level, the more competitive solar architecture alternatives are because the increased energy can be generated mass-efficiently during the daytime using light-weight thin-film solar arrays.

- Outside the band of northerly locations, fission power generation is superior in performance to solar-based architecture alternatives. This may have significant implications for a human Mars exploration program centered on scientific exploration on a global scale. On the other hand, a human Mars exploration program focused on the build-up of a single surface base could be located in solar-friendly locations and therefore enable a solar-based surface power system.
- Solar-based power systems with average power generation capabilities close to 100 kWe can provide sufficient power for crew survival during global dust storms without cleaning activities; this assessment was based on degradation of solar array efficiencies during global dust storms as observed with the Mars Exploration Rovers.
- Solar-based architecture alternatives were assumed to require manual deployment. During deployment, as well as during any additional pre-deployment surface time required for crew acclimatization to hypogravity, power needs to be provided from sources other than solar arrays; this could be provided by pre-charging the energy storage component of the surface power system during the Earth-Mars cruise as well as by utilizing radioisotope power sources with dynamic conversion for supplementary power generation.
- The development of robotic solar array deployment technologies could greatly enhance the competitiveness of solar-based architecture alternatives, because it would enable the use of solar power generation for in-situ propellant production. It is interesting to note that nuclear fission surface power systems also require robotic pre-deployment technology for connection to the surface assets as well as for large radiator systems; robotic pre-deployment technology may therefore not be a real distinguisher for surface power systems.

These conclusions suggest that future analyses of Mars surface power system architectures for human exploration should follow the above process and consider a variety of architecture alternatives (including ones based on thin-film solar arrays) as well as usage scenarios (single base vs. multiple separate global landing sites) in order to ensure robust down-selection to preferred architecture and technology alternatives are made.

References

- Buonova, G.A., Ahn, J., Hofstetter, W.K., Wooster, P.D., Hassan, R., de Weck, O.L.: Selection and Technology Evaluation of Moon/Mars Transportation Architectures. In: AIAA Space 2005, AIAA-2005-6790 (2005)
- Burke, K.A.: Fuel Cells for Space Science Applications. NASA Technical Reports Server (2003), <http://ntrs.nasa.gov>
- Connolly, J. (ed.): Planet Surface Systems Office JSC, Elements/Systems Database, JSC-45107. NASA Planet Surface Systems Office (1991)

- Dori, D., Edward, F.C.: Object-Process Methodology: A Holistic Systems Paradigm. Springer-Verlag New York Inc., New York (1999)
- Drake, B.: Reference Mission Version 3.0 Addendum to the Human Exploration of Mars: The Reference Mission of the NASA Mars Exploration Study Team. NASA/SP-6107-ADD, Lyndon B. Johnson Space Center (1998)
- Drake, B.: Reducing the Risk of Human Missions to Mars Through Testing, NASA/TM-2007-214761 NASA Technical Reports Server (2007), <http://ntrs.nasa.gov>
- Drake, B. (ed.): Exploration Blueprint - Data Book. NASA/TM-2007-214763. NASA Technical Reports Server (2007), <http://ntrs.nasa.gov>
- Elliot, J.O., Reh, K.: Lunar Fission Surface Power System Design and Implementation Concept, Space Technology and Applications International Forum, Albuquerque, NM, United States, February 12-16 (2006), <http://ntrs.nasa.gov>
- Graff, C.D., de Weck, O.L.: A Modular State-Vector based Modeling Architecture for Diesel Exhaust System Design, Analysis and Optimization. In: 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference Portsmouth, Virginia, AIAA. AIAA 2006-7068 (2006)
- Hanak, J.J., Fulton, C., Myatt, A., Nath, P.: Ultralight Amorphous Silicon Alloy Photovoltaic Modules for Space and Terrestrial Applications, American Chemical Society (3) (1986)
- Hoffman, S.J., Kaplan, D.I. (eds.): The Reference Mission of the NASA Mars Exploration Study Team. NASA Technical Reports Server (1997), <http://ntrs.nasa.gov>
- Hoffman, S.J. (ed.): NASA Mars surface reference mission, NASA/TP—2001-209371 (2001)
- Hoffman, E.J. (ed.): NASA Systems Engineering Handbook. NASA-SP-610S, Washington DC (1995)
- Hofstetter, W.K., Wooster, P.D., Sutherland, T.A., Crawley, E.F.: Architecture and Design Options for NASA's Lunar Surface Access Module (LSAM), IAC-06-D.2.7 / A.3.7.06. In: International Astronautical Congress, Valencia, Spain (2006)
- Kerslake, T.W., Kohout, L.L.: Solar Electric Power System Analyses for Mars Surface Missions. NASA Technical Reports Server (1999), <http://ntrs.nasa.gov>
- Koh, H., Magee, C.L.: A functional approach for studying technological progress: Extension to energy technology. Technological Forecasting and Social Change 75, 735–758 (2008)
- Liker, J.L., Morgan, J.: The Toyota Product Development System: Integrating People, Process And Technology. Productivity Press (2006)
- Littman, F.D.: Mars Power System Concept Definition Study, NASA Contractor Report CR - 19542. NASA Technical Reports Server (1994), <http://ntrs.nasa.gov>
- McKissock, B.I., Kohout, L.L.: A Solar Power System for an Early Mars Expedition, NASA-TM-103219. NASA Technical Reports Server (1990), <http://ntrs.nasa.gov>
- NASA HQ Exploration Systems Architecture Study Team, Exploration Systems Architecture Study - Final Report, NASA-TM-2005-214062, NASA HQ Washington DC (2005)
- Portree, D.S.: Humans to Mars - 50 Years of Mission Planning. NASA SP-2001-4521. NASA Technical Reports Server, ntrs.nasa.gov (2001)
- Schmitz, P.C., Penswick, L.B., Shaltens, R.K.: A Design of a Modular GPHS-Stirling Power System for a Lunar Habitation Module. NASA/TM—2005-213991. NASA Technical Reports Server, ntrs.nasa.gov (2005)

- Schulz, A.P., Clausing, D.P., Negele, H., Fricke, E.: Shifting the View in Systems Development – Technology Development at the Fuzzy Front End as a Key to Success. In: 11th International Conference on Design Theory and Methodology Las Vegas, September 12th-15th (1999)
- Shaltens, R.K., Wong, W.A.: Advanced Stirling Technology Development at NASA Glenn Research Center. In: NASA Science Technology Conference, June 18 (2007)
- Simmons, W.L.: A Framework for Decision Support in Systems Architecting. PhD Thesis, Dept. of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA (2008)
- Simmons, W.L., Koo, B., Crawley, E.: Architecture Generation for Moon-Mars Exploration Using an Executable Meta-Language. AIAA Space 2005, Long Beach, California (2005)
- Tenenbaum, D.: Power in Space: Time for a Biological Solution? Astrobiology Magazine (2007), <http://www.astrobio.net/news/>
- Ulrich, K.: The role of product architecture in the manufacturing firm. *Research Policy* 24, 419–440 (1995)
- Science Daily (eds): Breakthrough In Energy Storage: New Carbon Material Shows Promise of Storing Large Quantities of Renewable Electrical Energy, Science Daily (2008), <http://www.sciencedaily.com/>
- Sobek, D.K., Ward, A.C., Liker, J.K.: Toyota's Principles of Set-Based Concurrent Engineering. *MIT Sloan Management Review* 20, 2 (1999)
- von Braun, W., White, H.: The Mars Project. University of Illinois Press, New York (1961)
- Waiss, R.D.: Cost Reduction on Large Space Systems through Commonality. In: AIAA 25th Aerospace Sciences Meeting. AAA-1987-0585 (1987)
- Withrow, C.A., Morales, N.: Solar-Electrochemical Power System for a Mars Mission. NASA-TM-106606. NASA Technical Reports Server (1994), <http://ntrs.nasa.gov>
- Yoder, G.: Exploration Systems Mission Directorate. NASA Status Briefing, NASA ESMD. In: AIAA Space 2007 Conference Proceedings, Long Beach, CA (2007)

Chapter 13

Economics of Energy on Mars

Joseph E. Palaia IV, Mark S. Homnick, Frank Crossman, Alexander Stimpson,
and John Truett

4Frontiers Corporation. New Port Richey, FL, USA

13.1 Introduction

On Mars, as it has been on Earth throughout recorded history, the prosperity and freedom of action of human population will be dependent, necessarily, on the availability of energy. As such, it is of central importance that the methods by which such energy is collected, generated, stored, converted, transmitted, and ultimately utilized, be thoroughly examined. This book has been written with the intention of beginning such an examination.

All human activity, including that on Mars, must ultimately be considered economic activity. Thus, it is only appropriate that this chapter addresses the subject of the economic consideration of energy in that domain. Such an approach improves our understanding of the context in which energy will play a central role. It also provides a powerful toolset for energy system comparison and long-term strategic planning.

We begin by discussing the significant role that energy has played in the history of human habitation and industry. We observe that there is a direct correlation between standard of living and energy consumption, a correlation which will become even more significant as human beings begin living for long durations in space and on the surface of other heavenly bodies.

Next, we discuss the role which the authors of this chapter perceive energy will play in the development of the Martian frontier. It is expected that demand for energy will vary with each stage of Martian development. Early on, power-rich environments will dominate, driven by the demands of industry, growth, and the requirements for safety. As the Martian economy matures and population expands, energy demand will continue to change, and available energy may very well become a limiting factor in the growth of Martian society.

Having discussed the role of energy at an overview level, we examine how this understanding can help with early decision making and long-term strategic planning. This discussion includes a proposal to establish energy as the basic metric by which economic value is defined. In the specific context of infrastructure planning, energy expenditure can be the common point of reference for evaluating the

costs of infrastructure establishment, maintenance and operation. A worked example, the design of a Martian greenhouse, illustrates this approach. To complete the picture, it is shown that an account must be provided for the energy “cost” of human effort, and an approach for approximating this cost is proposed.

As the Martian economy continues to develop, trade between different Mars surface locales will surely develop, and such trade will be significantly influenced by the energy costs of transportation. This logic extends naturally beyond the Martian surface to trade between Mars and other domains within the inner solar system. The economic advantage of Mars as a source of goods and materials is examined and compared to the energy costs of lifting those materials out of the Earth’s gravity well.

Finally, we present a simple economic model to illustrate the positive impact on the Martian economy of establishing an early Mars industrial center. While serving as an initial beachhead in the Martian frontier, this infrastructure will provide key logistical support for further development and will serve as a focus for trade similar to port cities within the terrestrial biosphere.

13.2 Significance of Energy

Since its earliest beginnings, mankind has concentrated significant effort on mastering the raw materials, energy and resources found in its environment. Nomadic tribes provided for their sustenance through hunting and gathering. Invention and innovation, fueled by curiosity, led to the development of technologies which improved the yield from such efforts and reduced the time required to obtain the energy needed for sustenance. This beneficial cycle provided increasing amounts of energy to be expended and more time in which to expend it, leading to increased cultural sophistication, invention and innovation.

The development of both physical and social technologies has progressed in a similar positive cycle throughout most of human history. Quantum leaps forward in physical technologies mark this development, including the discovery of fire, agriculture, irrigation, metallurgy, water power, the steam engine, and the internal combustion engine to name a few. Quantum leaps in social technologies include spoken language, codes of laws, the written word, court systems & legal structures, and corporations. Each of these advancements significantly improved the ability of human beings to work together, extract resources and energy from the environment, and manipulate those resources to provide goods and services satisfying human needs and desires. Throughout human history two factors have remained constant: 1) as civilizations progress, the amount of energy consumed per capita continuously increases; 2) each increase in energy production appears to correspond with an increase in standard of living.

Modern society consumes significantly more energy per capita than preceding societies. This fact is readily apparent when one looks at the nearly limitless selection of products and services available to modern *Homo sapiens*, and considers the energy consumed in producing and distributing this bounty. But has this actually improved our standard of living? While this question could be considered somewhat open to interpretation, one cannot deny that today an individual dedicates

proportionally less of their time and effort to securing basic sustenance (food & shelter). A significant portion of their time is concerned not with survival, but rather with satisfying immediate desires including art and entertainment. This trend is illustrated by Fig. 13.1 and Table 13.1 (adapted for use here from (Cook 1971)), which depicts average energy expended per person for different activities and timeframes in human history.

Table 13.1 Daily Consumption of Energy Per Capita (x 1,000 kcal)

	Early Martian Man	Technological Man	Industrial Man	Advanced Agricultural Man	Primitive Agricultural Man	Hunting Man	Primitive Man
Transportation	111	63	14	1			
Industry and Agriculture	1574	91	24	7	4		
Home and Commerce	136	66	32	12	4	2	
Food	780	10	7	6	4	3	2

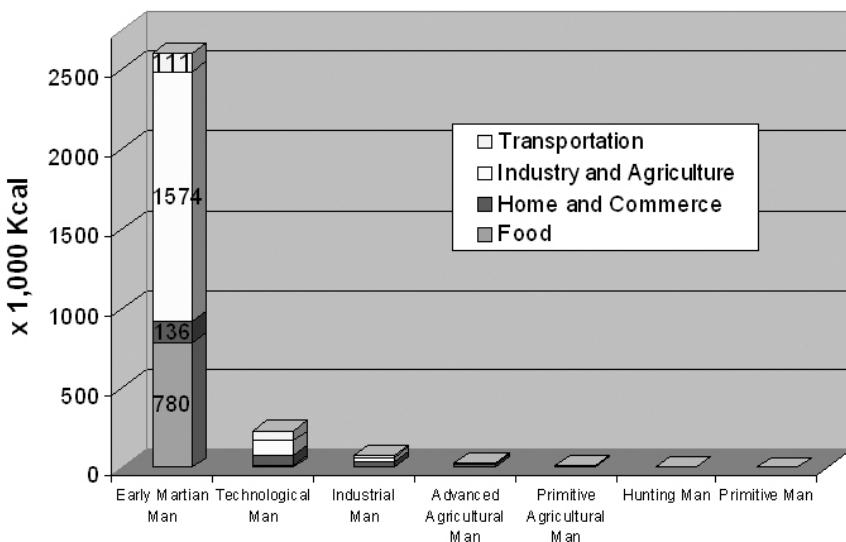


Fig. 13.1 Daily Consumption of Energy Per Capita

We have added the left most column (Early Martian Man), as an extrapolation of energy consumption and breakdown of activities for early Martian settlers. Early Martian Man will consume approximately 11 times the energy of modern

technological man. A significant portion of this energy is consumed by high-intensity agriculture (hydroponics, aquaponics) and by early industrial processes, where raw resources are extracted from the Martian environment and crafted into building materials and consumables. Note that energy for artificial lighting, required to achieve desired Martian crop yields, accounts for some of the increase in predicted energy consumption for food.

The assumptions have been made that the settlement is largely self sufficient, that trade has not yet developed between Earth and Mars, and that plentiful raw resources exist within close proximity of the settlement location (thus minimizing transportation requirements). It was also assumed that a relatively small number of people live on the planet (hundreds of people to a few thousand). Furthermore, the authors of this chapter anticipate a relatively high initial growth rate, with settlement size and population doubling approximately every four years. The population may grow through immigration and a high internal birthrate (10-20% yearly population growth) comparable to some developing countries on Earth. Population expansion will lead to economies of scale and other efficiencies. Growth rate fluctuation is also expected.

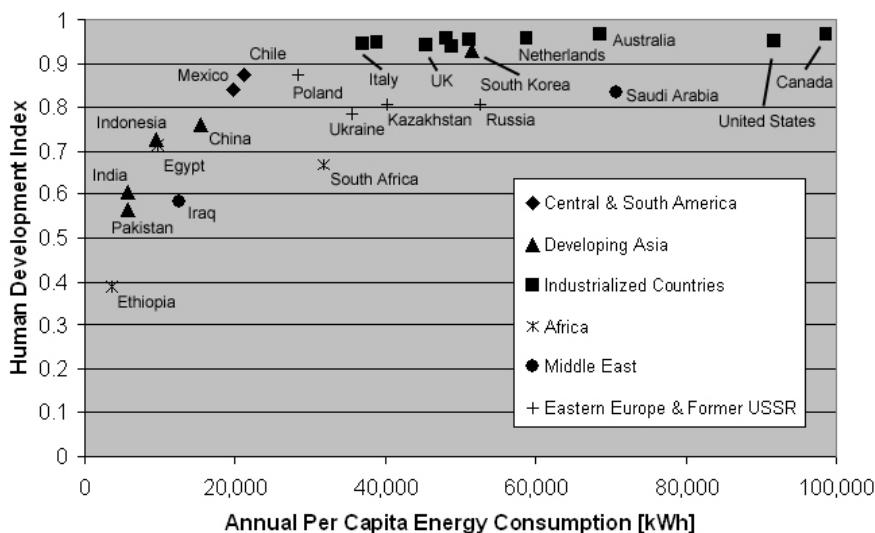


Fig. 13.2 Human Development Index vs. Per Capita Energy Consumption

Returning to our discussion of the relationship between standard of living and energy consumption, we find evidence of the correlation of these two factors by observing countries here on Earth. The United Nations currently assesses member countries through the Human Development Index (HDI), “a comparative measure of life expectancy, literacy, education and standards of living” (Wikipedia 2009b). Figure 13.2 is a plot of these index measurements (2005 data) versus annual per capita energy consumption for each country (2006 data) (IEA 2007). Industrialized countries such as the United States, Canada, UK, and South Korea have the

highest HDI values, while these are also the countries which consume the highest levels of energy per capita. Developing countries in Central America, South America and Asia have somewhat lower HDI values and correspondingly lower per capita energy consumption, while some locations such as Ethiopia have both low HDI and low energy consumption.

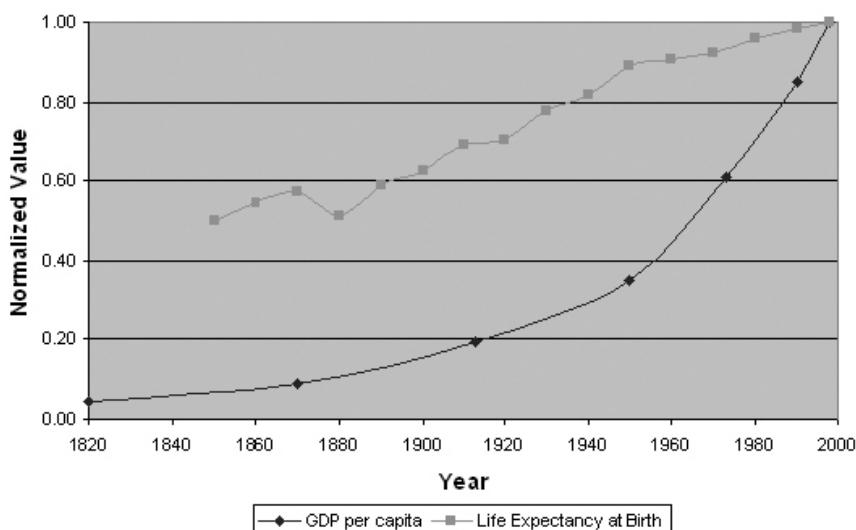


Fig. 13.3 Measures of Standard of Living in the U.S. vs. Time

Finally, we observe that over the past half-century, the standard of living in the United States has nearly tripled. We attribute the increase largely to higher per capita energy consumption. Figure 13.3 plots two measures of standard of living over time, specifically Gross Domestic Product (GDP) per capita and life expectancy at birth for native born males (Steckel 2005). These data have been normalized by the authors to the highest value recorded.

The data presented here, while certainly not comprehensive, supports our observation that standard of living is linked to per capita consumption of energy. It is possible that an investigation of other indicators would further substantiate this observation.

We believe that, as a species, the improvement of the standard of living for ourselves and for our fellow man is a desirable objective. We will therefore plan to establish a standard of living on Mars for the early settlers which is at least comparable to that found on Earth in western society. This objective will affect our requirements for energy consumption on the Martian surface. Due to environmental conditions which are inhospitable to life in the open, on Mars our reliance on mechanical devices will be more profound than at any previous point in human history. The availability of energy will literally mean the difference between life and death.

13.3 Demand as a Driver

It goes almost without saying that the scale of implemented energy systems on Mars will depend upon the scale of the operations and number of human inhabitants established there. Early human operations will likely focus on fulfilling science & exploration objectives, with a small number of explorers requiring modest amounts of power to support these activities. Round-trip human mission plans, including the NASA Design Reference Mission and the Mars Direct plan, call for 45 kWe (Drake 1998) and 100 kWe (Zubrin 1996) nuclear power systems, respectively, to produce methane for supporting surface operations and fueling the return craft.

Following early exploration, it is envisioned that permanent settlers will immigrate to Mars with the intention of building a new life for themselves. The 4Frontiers Corporation has completed two design studies of the first proposed settlement and has developed an understanding of its infrastructure requirements (Petrov et al. 2005). We will leverage that understanding throughout this chapter.

13.3.1 Early, Power-Rich Environments

By necessity, the first permanent Mars settlement will be a power-rich environment. Materials processing, life support systems, and food production will drive a large proportion of the energy requirements. Beyond these operational requirements, however, lies an even more important rationale for a power-rich environment. As stated above, unlike on the surface of the Earth, on Mars people will be dependent upon their mechanical devices, and therefore the energy powering those devices, for their very existence. As such, the only environment providing an acceptable level of safety for human inhabitants is one having plentiful available power and a rich diversity of power sources.

Due to the significant expense of importing materials from Earth, many of the materials required for the construction of infrastructure must be derived from the local Mars environment. A key advantage of locating a permanent settlement on Mars is that the Martian environment is a rich source for a multitude of the needed resources. Mars possesses the full range of chemicals and minerals found on the Earth, and the actions of water, wind, and volcanism have concentrated these resources in discrete locations. Much to the benefit of future Mars settlers, surface deposits of these resources are plentiful, unlike on the Earth where these surface deposits have been depleted by our industrial society. By transporting to Mars the key tools and equipment required to access these resources, one can establish the first significant industrial capability beyond the Earth's surface and leverage this investment by more than an order of magnitude in the mass of materials and products extracted from the Mars environment.

As we know from our terrestrial experience, the location, extraction, beneficiation, relocation and processing of resources in meaningful quantities, is an energy intensive process. Extracting a range of minerals from concentrated deposits, many of which will be located away from the main settlement, necessitates their transportation over significant distances. Water, perhaps the most critical resource,

is likely to be found at distances beneath the surface and will require energy intensive drilling and pumping operations to recover. Alternatively, water may be obtained through an even more energy intensive process, by baking it out of surface deposits of hydrated minerals. Carbon dioxide found in the Martian atmosphere is another key resource which will be compressed, stored, and broken into its constituent elements by more processes which will require energy. Finally, the wide range of mechanical and chemical processing required to refine all of these materials and manufacture final products from them will require yet more energy.

The 4Frontiers Corporation envisioned the first Mars settlement sustaining an initial group of 24 permanent settlers. Table 13.2 estimates the energy requirements for supporting the continuing needs of these inhabitants. More significantly, Table 13.3 estimates the amount of energy required to obtain and process raw resources and produce the building materials, goods, and products for constructing the settlement.

Table 13.2 Mars Settlement Power Loads (24 Settlers)

Area Type	Total Equiv. Cont. Load [kW]
Artificially Lit Greenhouse	439.9
Nuclear Plant BOP (Includes HVAC Loads)	69.6
Manufacturing (Includes Gas Plant Loads)	54.6
Misc Distributed Loads	27.3
Waste Processing	18.0
Air Handling	15.6
Inedible Plant Growing Areas	14.6
Lab Space	13.7
Water Plant (Domestic & Industrial)	12.8
Construction Habs	11.3
IT & Telecommunications	10.4
Mechanical Shop, Rover Garage, Suit Room	9.6
Living Quarters	4.6
Greenhouse Support Areas	3.9
Fish Farming	2.4
Medical Examination Areas	1.7
Sun Lit Greenhouse	1.7
Brewing, Fermenting & Distilling Areas	1.3
Total	712.8

The estimated continuous power demand of the first Mars settlement totals ~700kW of electricity. The artificially lit greenhouse consumes the largest amount of energy, by far. While Mars does receive sufficient solar irradiation to grow plants in naturally lit greenhouses, the decision was made to not rely on them exclusively. The 4Frontiers approach balances the risk of a long-duration global dust storm or a transient high-intensity solar radiation burst against the possibility of nuclear power system failure. Other significant settlement loads include the Heating,

Ventilation & Air Conditioning (HVAC) system (providing a comfortable environment within habitable spaces), the gas production facility (extracting gases from the Mars environment to replace those lost through seals in the settlement), the waste processing facility (processing the liquid and solid wastes produced by the settlement) and the air handling units (moving air throughout the facility).

Table 13.3 Energy and Power Requirements for Settlement Construction Materials

Material	Mass needed [kg]	Total energy [kWhe]	Power for 720 days [kWe]	Power for 400 days [kWe]	Power for 100 days [kWe]
Steel	78,000	273,504	16	29	114
Copper	28,303	134,677	8	14	56
Aluminum	12,391	342,578	20	36	143
Alumina	5,600	36,297	2	4	15
Silica	1,000	38,000	2	4	16
Silicon	2,000	68,000	4	7	28
Ethylene	58,981	369,614	21	39	154
Benzene	26,569	89,653	5	9	37
Polyethylene	46,400	51,968	3	5	22
Polyester	27,840	74,982	4	8	31
Epoxy	5,000	17,600	1	2	7
Polycarbonate	33,092	116,484	7	12	49
Basalt Fiber	80,033	28,537	2	3	12
Cement	195,000	225,688	13	24	94
Industrial H ₂ O	203,160	1,554,172	90	162	648
Settlement H ₂ O	300,000	2,295,000	133	239	956
Totals	1,103,369	5,716,756	331	595	2,382

Table 13.4 Mars Settlement Power Sources

Power Source	Max Power Capacity [kWe]	Operating Runtime Parameter	Net Power Capacity (Night) [kWe]	Net Power Capacity (Day) [kWe]
Nuclear Unit 1	400	1	400	0
Nuclear Unit 2	400	1	400	330
Nuclear Unit 3	400	1	400	400
Solar Arrays	15	0.3	4.5	5
RTGs (~6 units)	1	1	1	1
Available Power:			1205.5	735.3

4Frontiers assumes that three nuclear reactor units will provide primary power for the settlement (see Table 13.4). These units will each be capable of providing 400 kW of electricity, or 2,000 kW of thermal energy (supplied to industrial processes as up to 1100°C CO₂ working fluid). A modest sized solar array system and

several Radio Thermal Electric Generators (RTGs) will provide emergency “keep alive” power, in the rare occurrence that all three nuclear reactors became inoperative. It was also assumed that the artificial greenhouse lights (the largest electrical load of 440 kW) will only operate at night. During the day, turning off the greenhouse lights reduces the net electrical power required by the settlement, freeing one or more of the nuclear reactors to produce thermal or electrical energy for industrial processes. The excess energy could be provided for industrial use or produce a stored energy reserve of liquid or gaseous fuels.

Sixteen key materials have been identified as both essential for construction of the settlement, and as able to be produced using raw resources found in the Martian environment. The masses required of each of these materials have been estimated, as well as the total energy necessary to produce that mass (see Table 13.3). Material production rates will depend on many factors, including the size of the equipment and of the group making up the operating crew. Therefore, a calculation has been performed to show the amount of power required to produce the requisite mass of each material in 720, 400, and 100 days. Obviously, producing materials faster requires more power.

Figure 13.4 displays the relative percentages by mass of the required materials. Water, both for domestic and industrial use predominates, accounting for 46% of the required mass. Cement, basalt fiber and steel account for another 32%.

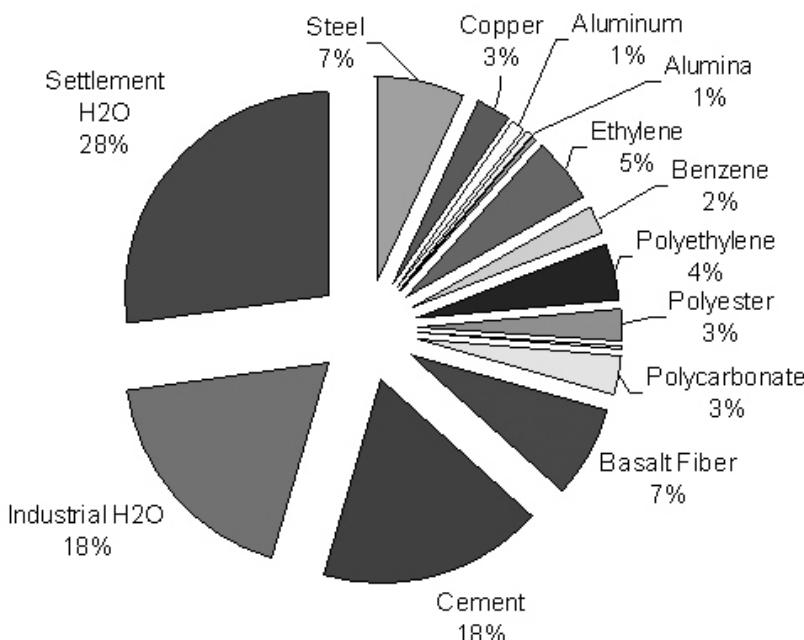


Fig. 13.4 Percent of Total Mass by Material

Figure 13.5 displays the relative percentages by energy required to produce the required materials. Again, water (as produced from hydrated minerals) predominates, as it requires 68% of the total energy allocated to producing materials for construction of the settlement. Aluminum, Ethylene, Steel and Cement account for another 21%.

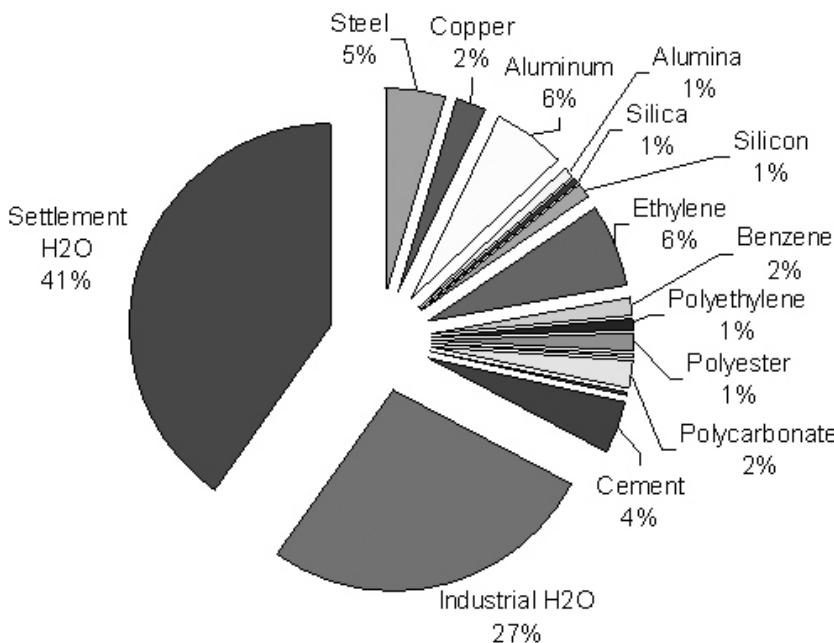


Fig. 13.5 Percent of Total Energy Needed by Each Material

The significant amount of energy (5.7 million kWh) required to support this modest level of industry and human habitation, underscores the importance of the subject matter of this work. A variety of power systems must be considered and evaluated to meet the early settlement power requirements.

As stated above, it was assumed that several small nuclear power plants will generate most of the power. These systems would be obtained from Earth and would require a minimum of setup and configuration after their arrival on Mars, thus providing the required power-rich environment with a minimum of effort.

As is the case with any power system, these reactors will have a finite operational lifetime, after which insufficient nuclear fuel will remain to continue energy production. Depending upon the reactor design, these reactors would need to either be replaced or refueled. Thus, if this approach is followed, Mars will become largely self sufficient from a mass perspective, while still depending on Earth for its source of energy (nuclear fuel). Given the aforementioned importance of energy, the means to achieve energy independence must be given serious consideration long prior to complete burn-up of the initially supplied nuclear fuel.

13.3.2 Growth and Expansion

The equipment and infrastructure established at the first permanent settlement will support the continued growth of the Martian frontier, providing building materials to enable expansion of the living quarters, food production, and other support infrastructure as immigration and local births increase the population. The first settlement will also provide supplies, equipment and materials to other groups who wish to establish their own outposts and independent settlements. As the Martian economy develops over time, so too will the energy requirements and the systems put in place to satisfy those requirements.

Throughout this discussion one fact has remained clear: plentiful energy can enable Martian growth and development. Conversely, we must recognize that a lack of energy will prove detrimental to the Martian economy and would limit the expansion rate in this frontier. The finite lifetime of nuclear energy supplies brought from Earth should drive investment in local sources of energy including solar, wind, geothermal, and perhaps even locally derived nuclear sources. Energy independence is important to the long-term viability of the Martian society, but in an environment such as Mars one must survive before one can thrive. Thus, mass independence should be addressed first, energy independence second, and perhaps, political independence third.

13.4 Energy as the Basic Metric

We have seen that energy will be an important and valuable commodity in any future Martian society. However energy will be just one commodity among many on Mars, each having differing “values” at the time, based in part upon relative supply and demand. How should one assign a value to a given material in a locale which is economically isolated or that only has the ability to receive materials but has little or no return trade?

First, we must recognize the difference between “value” and “cost”. Cost is directly traceable to tangible outlays required to produce the commodity. These outlays could include infrastructure, equipment, labor, consumables, transportation and depreciation. It is worthwhile to examine ways in which to determine cost, as it will provide some starting point when determining value. Unlike cost, “value” is less tangible, in that the value of a given material has a much less rigid connection with established cost. Value is determined by local supply and demand, which can at times be difficult to determine *a priori*.

We can estimate economic value in three ways. One possibility would be to assign the value of that same material on Earth. The authors of this chapter feel that this would be a poor choice. Even on Earth, materials that are difficult or impossible to transport have substantially different local values.

In the United States, natural gas is an example. Due to limitations on the importation of the gaseous product, combined with the relatively limited capacity of gas liquefaction plants overseas, limited transport capacity of liquefied gas and the lack of seaports capable of handling Liquefied Natural Gas (LNG), the price of natural gas in the United States is dependent on local and continental supply, not

worldwide capacity. As long as two locales remain isolated from each other, the value of a material in one locale has nothing to do with its value at the other. As less than 0.1% of US natural gas is imported as LNG, as of December 2005, natural gas pricing remains a local phenomenon (Lynch 2005).

A second concept of value may be the cost of material delivered at its final location. For example, we could define the value of steel produced on Mars as approximately the cost of steel delivered to Mars from Earth. This may seem a convenient method to assign value, as it is the closest approximation to true replacement cost. However, this approach fails to take into account the significantly lower cost of some locally derived materials. A key reason for establishing industrial infrastructure on Mars in the first place is because the energy cost of delivering and operating that infrastructure is significantly less than the energy costs of transporting the mass of materials that such infrastructure will produce over time.

The second concept of value is, however, a reasonable approximation for materials which are not derived from the Martian environment, due to lack of local infrastructure. Early industrial infrastructure will focus on producing those materials of which the highest mass is required for settlement construction and operation, as transport cost is directly proportional to mass. Highly processed and integrated goods, including semiconductors, sensors, and other low-mass and high-tech products will likely be imported from Earth for some time, while high-mass, low-tech materials including steel, aluminum, concrete, water, and a variety of polymers will be locally produced.

It is for these local materials that the third concept of value would be most appropriate. It involves approximating the cost of producing a given commodity (aluminum, ball bearings, wheat, etc.) on Mars. Determining the cost of the systems required to produce the given commodity, including infrastructure, consumables, labor, and depreciation, is challenging but not impossible. One could argue that such calculation will be only marginally more difficult than determining the cost of establishing and running similar operations here on Earth. For example, mining companies routinely consider logistical and economic factors of similar complexity when supporting operations in remote and hostile environments.

Once some sort of local active barter economy is created, one or more units of local value may act as currency, like wampum, tobacco or barrels of whiskey in colonial America, cigarettes in postwar Germany or units of proved land (land known suitable for agricultural production) in the novel "Farmer in the Sky" by Robert Heinlein (Heinlein 1950). The authors of this chapter argue here that the most significant unit of local value on Mars will be some fundamental unit amount of energy. Therefore, energy units should become the basis of Martian currency. We will call this unit a "QUID". The local barter economy will determine the value of an object (kg of steel, a methane tank, etc.), and the medium of exchange (QUID) is only a convenient local method of determining value.

By combining these concepts, one can determine the economic worth of a productive but isolated society. The cost of materials which must be imported from Earth will be determined by the established delivered cost. This cost can be defined by US dollars, or with no loss of specificity, by an equivalent energy cost (i.e. the energy expended to produce and transport the material). In contrast, the

cost of materials that are produced locally on Mars will be based upon their cost of local production. As we shall see, the most convenient method of measuring this cost is to equate all production factor costs into equivalent energy. The barter economy will relate the value of materials provided via these two approaches, and it will distinguish the value from individual costs.

Once bi-directional trade is active, evaluating economic worth becomes easier. If the currency used, for instance, is exchangeable in a free market on Mars, money will have an equivalent value on Mars and Earth. Usual classical models can than be utilized with values established in locations of creation, sale or use.

13.4.1 Equivalent Energy for Engineering Trade Analysis

As we have stated, the common unit of currency for defining value will be the Martian QUID. It makes sense for this currency to be based on an energy standard, considering that the costs of all materials, regardless of location or origin, can be defined in terms of energy. Energy is our common frame of reference.

This perspective provides us with a unique advantage when considering infrastructure planning, as will be illustrated by the example of the design and operation of a Martian greenhouse. The costs of operations, maintenance, and infrastructure establishment can be evaluated by applying equivalent energy analysis. We will see that this approach can be applied even to tradeoffs in engineering design.

13.4.2 Worked Example: Design of a Mars Greenhouse

The construction and operation of a Martian greenhouse illustrates the usefulness of energy as a basis for cost calculation on Mars. Greenhouse efficiency is difficult to compare on a direct level. For example, how does one quantify the tradeoff between a unit increase in light energy allowed to the plants (thus potentially producing more food) and an increase in structural steel mass to accommodate larger sheets of transparent polycarbonate? One way is to consider both quantities on an energy basis. Any sunlight provided to the plants can be compared as an energy gain from not having to supply that light artificially. This converted quantity can now be compared directly with the energy required for structural steel in order to better assess the advantages of a particular design.

In order to fully grasp the implications of implementing a particular greenhouse design on Mars, both initial energy costs and long term energy costs must be taken into consideration. A sample analysis of a greenhouse is detailed below:

Assumptions

- To simplify the problem, the “windows” are circular plates (disks) of polycarbonate, with an overlapping layer of steel around the outside for structural support (see Fig. 13.6). R_D is defined as the total disk radius, R_S as the visible radius (up to the support structure), and t as the thickness of the disk.
- The disks are to be laid out in a rectangular grid (see Fig. 13.6). While this is not the most efficient design, it simplifies the calculations, and could be adjusted later to any layout.

- This rectangular grid is assumed to be formed into a cylindrical shape, of large enough diameter that the seals can be assumed to have no angle between the structural steel and the disks
- The ends of the cylinder are ignored for the purposes of light, mass, and gas leaks
- The structural support material used is steel
- Energy to produce steel is 13.1 MJ/kg, energy to produce polycarbonate is 19.8 MJ/kg
- Total growth area is 1784.1m² (Owen-Going 2007)
- The artificial lights are Light Emitting Diodes (LED) Red Green Blue (RGB) arrays

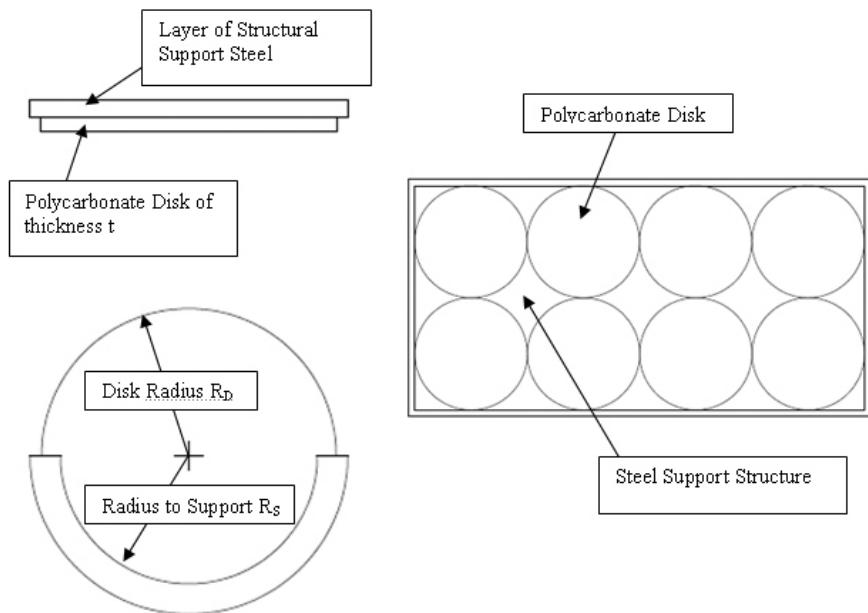


Fig. 13.6 Example of disk geometry with overlapping steel (left), and example of grid layout (right). The overall size of the disks is based on the disk radius R_D , while the effective useful sunlight area is based on the radius to support R_S .

Phase 1: Identifying the Players

The first step in the energy analysis is to identify all elements that provide or cost energy. One of the most significant costs is the fabrication of the structure itself. Masses of refined polycarbonate and steel cost significant energy to produce. Another cost, replacing gasses lost through leaks in the greenhouse seals, depends upon the number/layout of disks as well as the sealing technique. Another energy cost is the amount of artificial lighting needed to supplement the natural sunlight in order to maximize food production. The primary energy gain from the system is the energy of light let through the structure, which is taken up by the plants.

Note that some of these quantities are continuous (such as artificial lighting), and some of them are one-time costs (such as fabrication). One-time costs must be amortized over the lifetime of the structure. For example, if it is assumed that the entire structure needs to be replaced (from abrasion, etc.) every year, one can compare the total energy quantities over this period. Now the equation we desire to minimize can be constructed, as follows:

Energy Value = Energy required to fabricate structural steel (kWhe) + Energy required to fabricate structural plastic (kWhe) + Energy required to replace gas lost through leaks over a one year period (kWhe) + (Energy required to power LEDs to supply complete light to plants over a one year period (kWhe) – Energy input from sunlight over a one year period (kWhe)).

The smaller this energy value, the more efficient the energy use of the structure (corresponding to the lower total energy cost of the system). Now we must find these energy values.

Phase 2: Energy required to fabricate steel, E_s

To find the energy needed to fabricate the structural steel for the greenhouse, one can use the following equation:

$$E_s = k_s M \rho_p \pi R_D^2 t \quad (13.1)$$

where k_s is a conversion factor for the creation of steel ($= 3.5065 \text{ kWhe/kg}$), M - mass ratio, ρ_p = density of polycarbonate (kg/m^3), R_D - window disk radius (m), t - Window thickness (m).

Solving for the Mass Ratio

Mass Ratio is the ratio of mass of steel to mass of polycarbonate. This was determined based on a lunar structural analysis (Ruess et al. 2006), and adapted using material densities and yield strengths. The material used for both wall and framework in the lunar analysis was aluminum alloy (yield strength 500MPa). The yield strength of structural steel and polycarbonate are 250MPa (Wikipedia 2009c) and 65MPa (Boedeker Plastics 2009) for polycarbonate, about 1/2 and 1/6 the value of aluminum respectively. The density of aluminum alloy is 2.7 (Wikipedia 2009c), while that of steel and polycarbonate are 7.8 (Wikipedia 2009c) and 1.2 (Boedeker Plastics 2009), respectively (about 3 times and 0.5 times the value for aluminum). Therefore, it is assumed that relative mass in steel of about 6 times that of aluminum is needed. It is also assumed that a relative mass in polycarbonate of about 3 times that of aluminum (by the formula below) is needed.

Relative mass = (ratio of material density to that of aluminum)/(ratio of yield strength to that of aluminum)

Using this, we can apply a correction factor to the 24:7 mass ratio between wall and framework given in the lunar study. Since twice the mass of steel compared

to polycarbonate is needed (by looking at the ratio of the relative masses), one must double the mass of the framework. This yields a new mass ratio of 24:14 or about 1.714 kg polycarbonate/kg steel.

Phase 3: Energy required for fabrication of plastic, E_p

To find the energy for fabricating the greenhouse polycarbonate, one can use the following equation:

$$E_p = k_p \rho_p \pi R_D^2 t \quad (13.2)$$

where k_p is a conversion factor for the creation of polycarbonate ($= 5.50 \text{ kWhe/kg}$)

Phase 4: Energy required to replace the greenhouse atmosphere, E_g

To find the energy for replacing greenhouse gasses (lost by leaks in the seals), one can use the following equation:

$$E_g = k_g L_s G \quad (13.3)$$

where k_g - conversion factor in kWhe/kg for replenishing air in the greenhouse, L_s = seal length (m) and G = leak rate (kg/m) (see below).

Solving for k_g

Oxygen is produced at about 0.1714 kWhe/kg , and nitrogen is produced at 0.5313 kWhe/kg . Weighting this at approximately 0.785 for nitrogen and 0.215 for oxygen (and averaging), one gets an average value of about 0.454 kWhe/kg gas.

Solving For G :

Leak rate (1 atm) = $4.5 \text{ mL}/(\text{earth day})/0.414(\text{m seal length})$, as determined in a bell jar experiment (Rodriguez and Markovitch 1992). While the pressure differential of 1atm to vacuum is greater than our effective differential of 0.6 atm to 0.007 atm, we will use this value as a worst-case scenario to ensure a factor of safety.

Therefore, leak rate is given as $1.115 \cdot 10^{-4} \text{ m}^3/(\text{Mars day})/(\text{m seal length})$, where the Mars day is 24.62 earth hours. The density of air at 300K is 1.161 kg/m^3 . So, the mass of air lost is $1.2946 \cdot 10^{-4} \text{ kg}/(\text{Mars day})/(\text{m seal length})$. Over a 1 year period, $G = 0.0866 \text{ kg}/(\text{m seal length})$, where the Mars year is 668.6 Mars days.

Phase 5: Energy required for powering LEDs, E_L

We can estimate the energy for LEDs by calculating a weighted average of the light needs of the crops. After finding the weighted average in $\mu\text{mol}/\text{m}^2\text{s}$, a conversion factor (Folta et al. 2005) is used to compute the energy in Joules over a one year period necessary to meet this light requirement:

$$E_L = \frac{\sum_{i=1}^n (a_i y_i)}{n} Ct \quad (13.4)$$

where n is the number of crop types, y_i is the light requirement for each respective crop type, a_i is the weighting term defined as area of that crop type necessary to feed 24 people indefinitely (based on nutrition data) divided by total crop area (based on nutrition values), C is the conversion factor, and t is the amount of time. Note that this is assuming that distance is held constant between the plant and the LEDs at 5-7 cm.

Given 3 fixtures, at 36 LEDs per fixture, and 4W per LED, we can find that there is 0.432 kW/m² of needed light. After multiplying by the plant growth area, and accounting for the assumed one year period, this yields the energy required to provide LED lighting for the plants in kWe.

Phase 6: Energy input from sunlight, E_{SOL}

The energy input to the greenhouse is fairly straightforward to find, given the previous disk definitions. One can use the following equation:

$$E_{SOL} = I\pi R_s^2 t \quad (13.5)$$

where I is the irradiance (in kW/m²), which on average (for Mars) is found to be about 0.3 (Hublitz 2006), R_s is the radius of support of the windows and t is the amount of time. After again accounting for the one year period for the purposes of this analysis, we arrive at the total solar energy provided to the system in kWe.

Phase 7: Combining the Pieces

Using the equations defined above, one can find the Energy Value for a Martian greenhouse under any particular set of assumptions. Since energy has been used as a basis for cost, we may minimize cost directly for any system, by minimizing the following equation:

$$\text{Energy Value} = E_S + E_P + E_G + (E_L - E_{SOL}) \quad (13.6)$$

The energy value is graphed below (see Fig. 13.7) using MATLAB to find the optimal set of parameters for minimizing the energy required. Note, however, that using different materials or a different structural geometry may significantly alter the optimal values. The many assumptions that have gone into this analysis are reflected in the results.

It can be seen that the energy cost goes down dramatically with smaller disk radius and disk thickness. Note that both of these quantities are related, based on biaxial stress requirements. In order to support the pressure difference, a larger radius disk must have greater thickness for the same relative deflection. The recommended solution based on the model as described is therefore to have a fine mesh of disks held together with just enough steel to hold them. As the disks approach near zero radius and thickness, this could almost be thought of as a near-transparent fabric, or “bubble” type of structure. This result may not be particularly surprising, as there have been arguments on the efficiency of inflatable structures on the surface of Mars. This type of design minimizes the amount of polycarbonate and steel necessary for the structure, indicating that these are likely the primary driving factors for the energy value. On the surface of Mars, one might

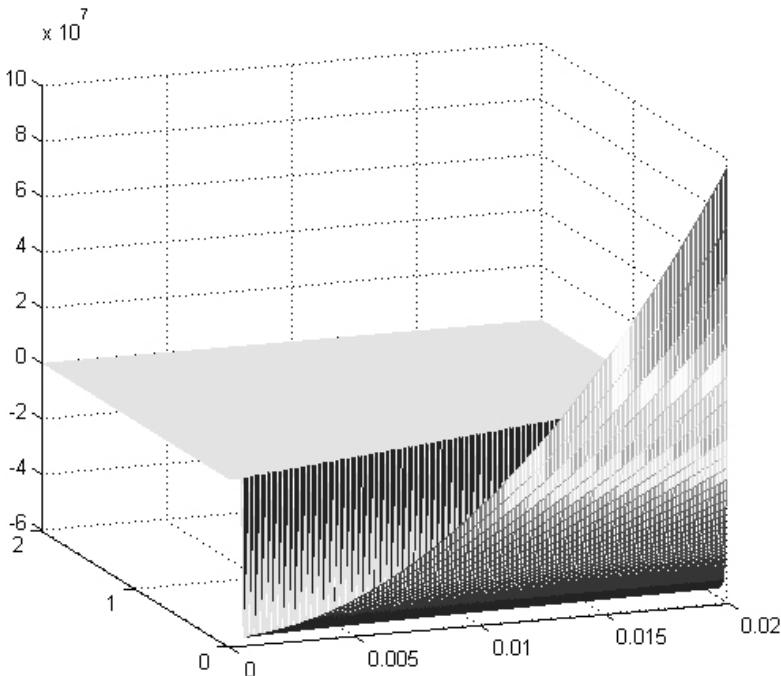


Fig. 13.7 Worked Example result showing the total energy cost of the system in kWe (vertical axis), dependent on disk radius in meters (lower left axis) and disk thickness in meters (lower right axis).

envision that the colonists could construct machines that would continually produce these tiny polycarbonate “chips” on a grand scale for the production of the greenhouses.

The result seen is a direct consequence of the assumptions used to drive the model described in this section. As part of the model, the mass of steel required for the structure is directly proportional to the mass of polycarbonate. In actuality, there is likely a diminishing return on the reduction of the mass of steel, and that it is probably more realistic to have an adaptive function for the mass of steel as the mass of polycarbonate changes. It is also important to note that the calculations for the rate of gas leakage out of the structure (proportional to the seal length) were based upon leak rates from bell jar experiments and International Space Station airlocks. These values are appropriate only if that level of excellent sealing is maintained on all the polycarbonate-steel interfaces. As the disks become extremely small, it is likely that this same quality of sealing will become extraordinarily difficult to maintain across all the interfaces. Defining a value for gas leakage that changes to account for this difficulty might be incorporated in future versions of the model.

It is interesting to note that some energy values (lowest on the vertical axis) are actually negative. This is possible due to the equation used for total energy cost.

We have subtracted the energy input due to sunlight, so if this energy value exceeds all of the energy costs of the material fabrication and greenhouse maintenance, the total cost will be negative. While this may seem unusual at first, it speaks to one of the primary reasons for the use of a greenhouse on long-term space missions. Sunlight utilization by the plants in the greenhouse turns into food production, waste removal, and atmospheric revitalization for the settlement as the plants grow. This type of energy graph illustrates that a properly designed greenhouse containing crops able to produce in the ambient Mars sunlight could more than pay for its own energy cost.

The model and results presented here do not represent a final product, but a starting place for the development for models defining projects on Mars in terms of their energy cost. Further iterations of the model would increase its reliability and potentially produce additional interesting results. The use of energy as a currency provides us with a way of quantifying the advantages and disadvantages of dissimilar concepts (i.e. how much polycarbonate is a 1kJ/day increase in sunlight worth?). This concept has been applied to a simple calculation involving a greenhouse structure, but could be employed to any number of projects on the Martian surface.

13.5 Energy of Human Endeavor – Impact on Design

To provide a complete accounting of energy costs on Mars, we must assign a cost to human effort. On Earth, when an employer assigns a cost for labor on a project, the required salary depends upon rates for comparable talent in the local market. Salary calculations include the cost of living in the area from which the labor is being drawn, competing talent available, and a host of other factors. On Mars, at least early on, the cost of labor will more intimately link to the cost of living. It is assumed that settlers will be living initially in a “company town”, where the employing company owns the support infrastructure. Thus, the cost of the infrastructure and consumables for supporting each human inhabitant will define the cost of human labor.

A difficulty is presented in the fact that there will be significant common infrastructure provided for the benefit of the entire settlement’s population. The cost of this infrastructure must be divided among the settlers in residence. Thus, the cost of labor on Mars will depend largely upon the scale, the complexity, and ultimately the cost of the settlement where the labor is located. Since the cost of infrastructure can be translated into an equivalent energy cost in the manner previously stated, one can arrive at a cost of human labor defined as an amount of energy per day. This cost can then be directly input into larger calculations of the cost of Martian operations.

A settler also incurs significant transportation costs in traveling to Mars in the first place. One may decide that this initial cost (which we assume was paid for by the company) should be amortized over some portion of the inhabitant’s lifetime. If this is the case, then one should add the amortized portion of this cost to the above mentioned support infrastructure and consumables cost when computing the cost of labor.

13.6 Mars Surface Trade

To provide economically for basic shelter, food, protection, and social interaction, early Martian human settlements will be concentrated in a small number of primary locations. Other locations will also be developed, remote from the main settlement(s), where a particular resource (e.g. water ice, hematite, aluminum rich clays, silica, etc.) is found in relative abundance. Small numbers of settlers, robots, and automated equipment at these remote sites will extract a particular resource, concentrate it, package it for shipping, and transport it to the primary settlements where power from the major energy generators will transform the resource into end product(s).

The allocation of people, robots, and extraction equipment to these remote sites will be done using an economic analysis that accounts for the costs of all equipment at the site, its maintenance, eventual retirement or replacement using some depreciation schedule, and other costs. Planners will measure these costs in a currency based on the value of energy. The Martian QUID will not be based on an amount of precious metal or on bank promissory notes, but rather on a quantity of energy (i.e. Joule or Watt-hr = 3600 Joule). For many years, energy will be the dearest resource on The Red Planet, perhaps until distributed renewable energy resources such as solar and wind power provide a major portion of Martian settlement needs.

The number of people at these remote sites should increase with time as occupants of the main settlements discover opportunities for providing services to the remote miners and processors. Services could include local food and energy production, equipment maintenance and repair, a general store or distribution center (think FedEx), an inn/bar/brothel, etc. As in the settlement of the American West, individuals or families will choose to move to the remote locations if the benefits of the move are economically quantifiable within the Martian currency system.

A major cost of operating these remote sites is the cost of transporting goods from the remote site to the settlement, including the cost of periodic resupply with people and equipment. The largest currently known sites of highly concentrated clay, silica, iron rich hematite, etc. on Mars are thousands of km apart. Even if the main settlements are built adjacent to a concentrated source of one resource, each settlement will likely be hundreds, if not thousands, of km from others. The use of a methane-oxygen internal combustion engine has been suggested as an effective means of land transportation for wheeled or rail vehicles. However, we must analyze the energy cost of the transportation infrastructure, including that of supplying caches of fuel-oxidizer along the route and of establishing an effective road or rail to each site. Furthermore, we must consider the cost to extract a particular resource from a location that is perhaps less rich (less concentrated) in the resource but is closer to the settlement. Where the mass of the transported resource is small, alternative transportation via atmospheric hoppers or buoyant airships may provide a more cost-effective transportation method.

In all these considerations, after establishing and documenting the energy cost of each activity, standard economic analysis practices will be used to allow the tradeoff among several alternatives. As stated earlier, individuals may make their

decisions based on the microeconomics of their particular interests, experiences and assessment of future opportunities with little or no formal economic analysis, but the settlement as a whole will require more formal analyses and attention paid to gathering the data that allows the energy cost to be quantified. Also, social and political issues will sometimes trump the economic cost considerations.

13.7 Trade between Mars and Other Solar System Domains

From the considerations above, and some extrapolation from our own history on Earth during the industrial age, we can project how a distributed set of human settlements on Mars will trade with each other using a common currency based on energy or energy credits. But now one must ask whether the trade between Mars and other Solar System domains (Earth, asteroids, Luna, Earth orbit hotels, etc.) is subject to the same sorts of assumptions that are carried with us in this modern world.

Goods and services that Mars might trade with these other domains should require roughly the same amount of energy to produce if they are extracted from locations that are at least as rich as those available on Earth. Chemistry and the necessary chemical engineering infrastructure dominates the energy content of mined, refined, or synthesized goods, independent of where the goods are produced. With the appropriate manufacturing infrastructure in place, an ingot of iron produced on Earth or Mars or the Moon has about the same energy cost to produce.

Yet Mars needs an external source of trade revenue to become an independent place of settlement and not just a well-subsidized colony of Earth. It has been suggested that a trade network among these Solar System domains could be established (Sylvan et al. 2009; Zubrin 1996). In the paragraphs that follow it will be shown that the cost of delivered trade goods in the solar system will be dominated by the cost of transportation, especially when the product is produced on a planet or large moon like Luna where the delivery entails escape from a deep gravity well. Deep gravity wells around planets require high impulse rocket engines firing for short times to generate a change in velocity sufficient to orbit or escape the planet's gravity. The cost of transportation via rocket propulsion is greater from starting locations with a deep gravity well. Based on projections for the reusable Falcon 9 Heavy launch vehicle (still in design/test), the near-future cost of delivering a payload to Low Earth Orbit will be approximately \$3,000 per kg (by shuttle it is now \$15,000/kg). The cost of building materials (e.g. cement, steel beams, aluminum framing) ranges from \$0.10 to \$10 per kg., so *the space transportation cost to Low Earth Orbit (LEO) is 300-30,000 times the material/product cost on the planet of its manufacture – Earth in this case.* It seems logical to conclude that, if material resources are found on a planet or large moon, it will be cheaper by far to extract them and use them locally because the cost of transporting them to orbit or another planet or moon is huge compared to the cost of manufacture.

However, Mars enjoys a transportation cost advantage for delivering certain goods to LEO and Luna compared to the cost of delivering the goods from Earth. LEO contains no local resources except solar power. On Luna, oxygen and structural materials such as steel and titanium can be processed from the regolith, but

sufficient quantities of water, hydrogen, nitrogen, carbon, or organic chemicals and polymers made from C,H,O,N are unconfirmed. Significant opportunities will exist someday for delivery of supplies and materials from Mars to LEO and to Luna, but to understand why, let's review a little bit of orbital mechanics.

Anyone who grew up in the Space Age is familiar with the idea that given a sufficient positive change in tangential velocity, a rocket can launch from Earth and enter into a stable orbit around the planet. Even when the rocket stops firing and accelerates toward the Earth under the influence of gravitational attraction, it is traveling fast enough tangentially to remain in circular or elliptical orbit around the Earth instead. This same rocket could be fired again to gain a second increment in velocity, perhaps reaching a total velocity relative to Earth called ΔV_{c3} – the hyperbolic velocity, at which the rocket is moving fast enough to escape the gravity of Earth all together. At this point the rocket is in orbit around the Sun at about the same radius as Earth. To send the rocket to Mars, the least energetic approach uses the Hohmann transfer orbit – the lowest energy elliptical path from one circular orbit to another via two high thrust firings – one at earth orbit and the second when approaching Mars orbit. Hohmann transfer orbits require less total energy compared to other maneuvers with low thrust rockets such as ion engines, so the authors of this chapter will continue to focus on use of high impulse – high thrust rockets in this discussion. It turns out that the Hohmann transfer orbit maneuver allows us to hand off the rocket from one two-body gravity problem to another without having to deal with multibody gravity mechanics. To summarize with high impulse-high thrust rockets and Hohmann orbit transfer, one can describe the energy required to reach each destination as simply a scalar change in magnitude of velocity.

This allows the construction (Francis 2008a, 2008b; Wikipedia 2009a) of a Δv (km/s) diagram (see Figure 13.8) for the Sun, Earth, Moon, Mars, and Ceres – the largest dwarf planet of the Solar System (formerly known as an asteroid). The diagram uses circles to illustrate Low Orbits around Earth, Mars, and Ceres. It also shows some Δv segments between Earth and Mars that are direction dependent. In going from Earth to Mars, a spacecraft can aerobrake in the atmosphere of Mars to shed most of its velocity prior to landing; minimal retro-rocket mass or fuel is required. The figure represents the maximum Δv possible by aerobraking at Mars as those segments directed towards Mars along the path from Earth. Similarly one can reduce the total Δv needed to go from Mars to Earth without expending much rocket fuel to change the velocity along the segments directed towards Earth along the path from Mars.

To determine the relative costs of these various journeys, we will use the rocket equation developed by Tsiolkovsky (Wikipedia 2009d). Assuming that multistage rocket design can be applied on rockets from Mars as well as those from Earth, we will ignore the details of multistage rocketry as a means to maximize payload for a rocket engine with fixed thrust. Instead, we focus on the mass of propellant needed to change the velocity of a rocket using a rocket engine with an exhaust velocity of v_{ex} .

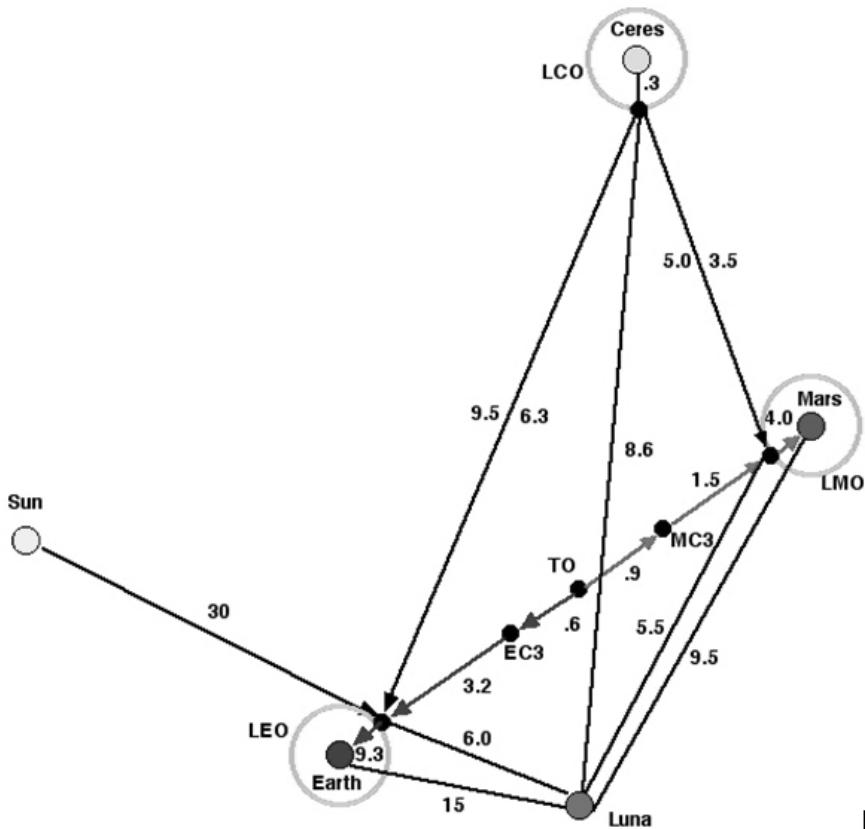


Fig. 13.8 ΔV Diagram for the Sun, Earth, Moon, Mars, and Ceres

The Tsiolkovsky chemical rocket equation establishes that the ratio of an initially fueled rocket of mass m_0 to the final dry rocket mass m_f (which includes the payload mass m_{pl} as well as the structural mass m_s) is an exponential function of $\Delta v/v_{ex}$ in the form:

$$\frac{m_0}{m_f} = e^{\left(\frac{\Delta v}{v_{ex}}\right)} \quad (13.7)$$

It is assumed that the exhaust velocity will be 3.7 km/s corresponding to a L_H-L_{OX} fueled chemical rocket. A rocket of initial mass m_0 carries fuel of mass m_p that typically accounts for ~90% of the total mass of the rocket. The remaining mass is the payload mass and the mass of the rocket structure. It is desirable to convert the rocket equation into a form in which one can compare costs instead of mass.

First, one should note that, for reusable rockets amortized over many space flights, the effective cost of the rocket's journey is dominated by the cost of the

fuel-oxidizer propellant used. Therefore, the effective cost is proportional to the propellant mass m_p . If the unit cost of the propellant is C then:

$$\text{Total cost } \sim Cm_p$$

The mass of the payload m_{pl} will be taken to be a fixed ratio of the structural mass $m_{pl}=m_f/R$. The cost per unit mass of payload is then

$$\begin{aligned} Cm_p / m_f / R &= RC(m_p / m_f) = RC(m_o - m_f)m_f \\ &= RC \left[\left(m_o / m_f \right) - 1 \right] \end{aligned} \quad (13.8)$$

Since the cost of producing fuel from water by chemical electrolysis is the same wherever you do the electrolysis and since R is a constant of the rocket design

$$\text{Cost per unit payload} = RC \left[\left(m_o / m_f \right) - 1 \right] = RC \left[e^{(\Delta v / v_{ex})} - 1 \right] \quad (13.9)$$

With the Δv s of Fig. 13.8 and the derived payload cost equation above, we obtain Table 13.5 which compares the costs of transportation from Earth and from Mars to several other destinations in the Solar System. This comparison is expressed as a ratio of costs incurred from Mars vs. Earth (so that the constant RC drops out in the comparison).

Let's focus on the bolded items in Table 13.5. For those that are more comfortable using dollars rather than Martian energy credits in QUID, you may calibrate these costs by assuming a \$3000/kg cost of payload from Earth's surface to LEO as projected for the Falcon 9 Heavy. The equivalent cost from Mars surface to LEO is 41% of Earth cost or \$1230/kg of payload. The advantage of carrying payload to Geosynchronous Earth orbit (GEO) from Mars' surface is even greater. The cost ratio is .13 times the cost of carrying the payload from Earth's surface. The transportation cost of supplying a Moon base from Mars is 0.21 that of Earth. Finally the cost of setting up a mining operation on Ceres from Mars is only .07 the cost of doing it from Earth, even if eventually the transfer of final goods from Ceres to a place like Luna is 0.84 the cost of delivering the same goods from Mars.

To summarize, space trade is all about the space transportation cost. Because Earth has a deeper gravity well, conducting trade from a Mars base provides an amazing cost advantage over conducting the same trade from Earth. One should be realistic regarding other aspects of commerce and supplier decisions that may favor a more costly but timely supplier. A good example of this is the decision we all make regarding the timely delivery of a product (air vs land mail, FedEx Overnight, etc). Delivery of goods from Earth to Luna may cost four times as much as from Mars, but delivery will take a week instead of 6 months to 2 years. Similarly, delivery of goods from Earth to LEO may cost twice that of goods delivered from Mars, but the goods can arrive the same day rather than after 6 months to 2 years after placement of the order. Martian traders must take advantage of trading partners at LEO and the Moon willing to use long term delivery contracts that can save them thousands of dollars per kg.

Table 13.5 Comparison of Transport Cost Between Earth and Mars to Other Destinations

Path	Earth		Mars		Earth/Mars cost ratio	Comment
	Δv	Mass ratio	Δv	Mass ratio		
Surface to LO of planet	9.30	12.35	4.00	2.95	0.17	
Surface to C3 of planet	12.50	29.32	5.50	4.42	0.12	
Surface to LEO	9.30	12.35	6.40	5.64	0.41	with aerobrake
Surface to GEO	13.40	37.40	6.40	5.64	0.13	with aerobrake
Surface to Luna surface	15.00	57.63	9.50	13.03	0.21	
LO to Luna surface	6.00	5.06	5.50	4.42	0.84	
Surface to Mars surface	16.50	86.44				no aerobrake
Surface to Mars surface	13.10	34.49				with aerobrake
Surface to Ceres	18.60	152.48	8.90	11.08	0.07	
Ceres LO to LO	9.48	12.95	4.96	3.82	0.24	no aerobrake
Ceres to LO	6.60	5.95	3.80	2.79	0.36	with aerobrake
Ceres surface to Luna surface	8.90	11.08			0.84	*Mars vs Ceres cost ratio

13.8 Economic Models to Aid Mars Infrastructure Planning

In itself, the development of models for the space economy is a sign of rapid change in the uses of space and of the expanding potential of space to impact human activity.

By examining historical trends of world gross domestic product and extrapolating through the initial decades of space frontier development, one can create a set of economic models of the process. A base model can use both historical analogies and a variety of assumptions concerning social and scientific progress, increase of population, etc. A series of derived models can then be developed to evaluate positive effects such as improved productivity, as well as growth of economic activity in the inner solar system. This section intends to define and consider examples of both these base and derived models.

In general, economic growth beyond Earth orbit will reflect several factors, including changes in space transportation speed and expense, new technology, energy conversion efficiency, and the growing scarcity of certain metals on Earth. This economic growth is predicted to take place on four current and emerging frontiers: Earth orbit; the Moon; the Mars system; and the Asteroids. Mining asteroids and creating a permanent Mars settlement are considered particularly significant events (Sylvan et al. 2009).

13.8.1 Measures of Economic Activity

Economic activity may be measured in several ways:

- **Classic:** Gross / Net Domestic Product (GDP/NDP) where World GDP/NDP is the global economic activity, either gross or net, of an economic entity or area.
- **Total Economic Activity:** Counts some activity in more than one location but may be a better measure of actual activity for a given location within this environment.

To evaluate the value of exploration, one must assign an economic value to data. The best available models are programs where NASA outsources a request for data. NASA effectively creates a value for the particular data requested. One can extend this concept to create a minimum measure for the value of data and how to attribute it.

For sake of this model, economic value of data will be defined as mission cost. A successful mission therefore has at least twice the value of a failed mission. This includes both value of creating the mission, as well as the value of the generated data, plus the value of any extensions of the mission and the value of the data created in those extensions.

Having given a value to data, over and above mission cost, one must then locate where to attribute its value. In classic economics, value of data either does not exist or lies in economic rents of other sources, is attributed to the location where the activity is utilized or to the mission's objective(s). In a model of total economic activity, value may be attributed to both origin and objective. For example, a Mars scientific mission is attributed to both Earth (where the data is organized and utilized) and to Mars (where the mission is sent and the scientific information is created). The mission is part of the Mars economy, while the data are part of the overall Earth intellectual database.

13.8.2 Economic Zones of Interest in Initial Models

1. **Earth:** All activity originating from the center of the Earth to that point in Earth's atmosphere where orbit is possible. In its current nascent form, we consider suborbital flight an Earth activity. If suborbital activity leads to orbital activity, (which may occur soon) the location of such activity within the model may need reevaluation. **Earth Zone Value = EV**
2. **Earth orbit:** All activity relating to objects in Earth orbit. This includes economic activity resulting from the process of placing the object in orbit, or by processes using the orbiting object, or by processes involved in the servicing or visiting of the object. **Earth Orbit Zone Value = EOV**
3. **Moon:** All activity relating to objects passing, orbiting, or on the surface of the Moon. The economic activity attributed to the Moon includes launches going to the Moon with free return, those that go into orbit around the moon and those that land on the moon. **Moon Zone Value = MOV**
4. **Asteroid:** All activity relating to asteroids, comets or other objects not considered to be planetary bodies. The asteroid-based economy is defined as

economic activity arising from asteroid development and operations. This includes economic activity resulting from the process of placing an object in orbit around or in contact with an asteroid, processes involved in visiting an asteroid, processes involving in obtaining and/or using materials extracted from an asteroid and by the servicing of equipment placed on or near an asteroid. For the sake of this model, Pluto is a planet. At this stage of model development, activity on the moon of a non-Terran planetary body is attributed to that planet. **Asteroid Zone Value = ASV**

5. **Mars:** All activity originating from center of Mars to the most distant point where stable Mars orbit is possible. The economic activity attributed to Mars includes launches going to Mars with free return, those that go into orbit around Mars and those that land on Mars. This includes economic activity resulting from the process of placing an object in orbit around or on the surface of Mars, processes involving in obtaining and/or using materials extracted from Mars and by the servicing of equipment placed on or near Mars. At this level of model formation, any activity relating to probes in orbit or on Phobos and Deimos are considered Mars activity. **Mars Zone Value = MAV**
6. **Other:** Activity at Lagrange Points and activity in other areas are clustered together. Activities that are directed to one of the Lagrange points or take place at one of the Lagrange points will be considered Lagrange point economic activity. If a particular region or body shows substantial activity, it may need a separate line within the model. **Other Zone Value = OTV**
7. SEV or the total space economy value = EOY + MOV + ASV + MAV + OTV

13.8.3 Space Economy Model – General Form

The general form of the space economy model is described by:

$$\begin{aligned} SEV(v_1, \dots, v_n, t) = & EOY(v_1, \dots, v_n, t) + MOV(v_1, \dots, v_n, t) \\ & + MAV(v_1, \dots, v_n, t) + OTV(v_1, \dots, v_n, t) \end{aligned} \quad (13.10)$$

Here, v_1 to v_n are the variables being evaluated as inputs in any particular model. Variables may be direct and locally calculable such as local methane production, or of indirect effect such as reduction in the cost or time of transport. “ t ” is the variable time.

13.8.4 Inner Solar System Trade

Several previous publications (Zubrin 1995, 1996) address early inner solar system trade. They generally predict active trade outside of Earth orbit in addition to the transfer of highly processed goods and settlers from Earth in return for raw materials and lightly processed goods from the space frontier. One possible relationship is included in Figure 13.9.

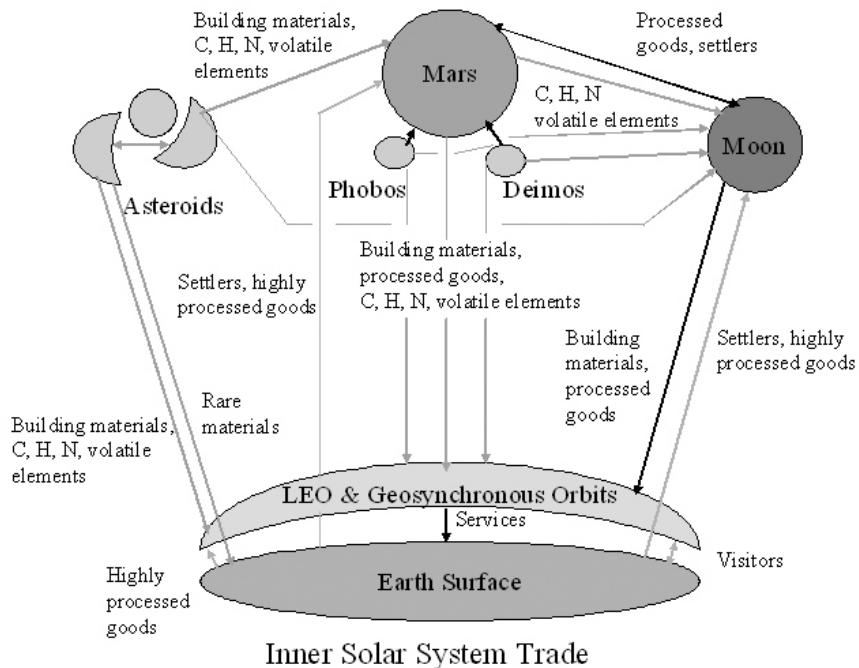


Fig. 13.9 Inner Solar System Economic Domains & Interactions (Sylvan et al 2009)

13.8.5 Modeling Results

To provide a sense of the possible use of a model of the space economy, the following section presents several examples of modeling results including an initial model with subsequent modifications.

MODEL 1

Assumptions. The model begins with the year 2005 and is followed through 2050 at initially five then ten year intervals. In 2005, world GDP is \$38 trillion, the space economy is worth \$180 billion. GDP grows at 3.3% per year and the space economy grows at twice that rate (as predicted by Sylvan et al. (Sylvan et al. 2009)). Dedicated moon activity starts to be meaningful in 2010, Mars activity in 2020, and asteroidal activity in 2025.

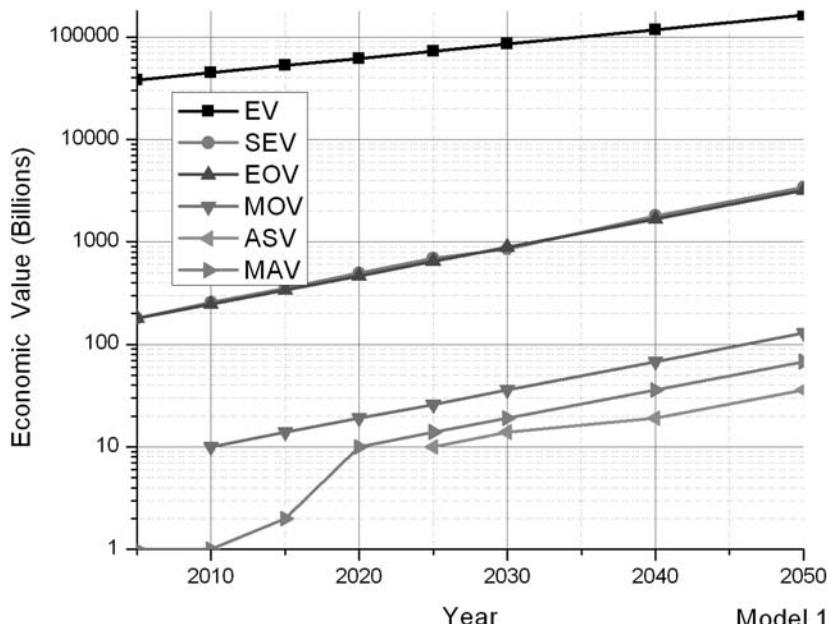
The base equation is:

$$\text{Econ. Comp. (year } n) = \text{Econ. Comp. (year } b) * (1 + kr)^{(n-b)} \quad (13.11)$$

where b is the base year, n is the year being evaluated, k is the growth rate of the world economy and r is the ratio of the space economy growth rate to that of the world economy.

Table 13.6 Model 1 Economic Projections (Values in Billions)

	2005	2010	2015	2020	2025	2030	2040	2050
EV	38,000	44,688	52,553	61,802	72,680	85,471	117,917	162,770
SEV	180	257	355	496	693	854	1799	3409
EOV	179	246	339	467	643	885	1676	3176
MOV	inc. in EOV	10	14	19	26	36	68	129
ASV					10	14	19	36
MAV	1	1	2	10	14	19	36	68
OTV	included in EOV							

**Fig. 13.10** Model 1 Economic Projections (Sylvan et al. 2009)

In this model, the base year is 2005, the assumed growth rate k is 3.3% per year and the ratio r is 2. As this is a simple exponential equation, it is easily graphed in logarithmic format. Simplified, the model therefore is:

$$\text{Econ. Comp. (year } n) = \text{Econ. Comp. (year } b) * (1.066)^{(n-b)} \quad (13.12)$$

If for example, the base year is 2005, the year of interest is 2015 and the base space economy value is 180 billion dollars, the equation becomes:

$$\text{Space economy (2015)} = \$180 \text{ billion} * (1.066)^{(2015-2005)} \quad (13.13)$$

$$= \$180 \text{ billion} * (1.066)^{(10)}$$

Model 1 gives us a snapshot of how the space economy may develop under a particular set of assumptions and initial conditions. At the commencement of the time period in question, the space economy was about 0.5% of the overall level of the earth economy. Even using the extremely conservative estimates of this particular model, the space economy, already a significant force in world economics has grown to over 2% of the global economy. This would place the space economy just outside of the top ten nations in GDP / Net National Product (NNP). In addition, while the dominant portion of the space economy remains the near earth segment, the Moon and Mars segments combine to surpass the size of the current space economy by mid-century.

MODEL 1 ALTERNATE

Assumptions. Identical to Model 1, Model 4 increases the rate of Martian growth by 2% per year after 2030 due to the creation of a permanent settlement that acts as a center of commerce for raw material and industrial production.

Table 13.7 Model 1 Alternate Economic Projections

	2005	2010	2015	2020	2025	2030	2040	2050
MAV (Model 1)	1.0	1.4	1.9	10.0	13.8	18.9	35.9	67.8
MAV (Model 4)	1.0	1.4	1.9	10.0	13.8	18.9	43.6	98.2

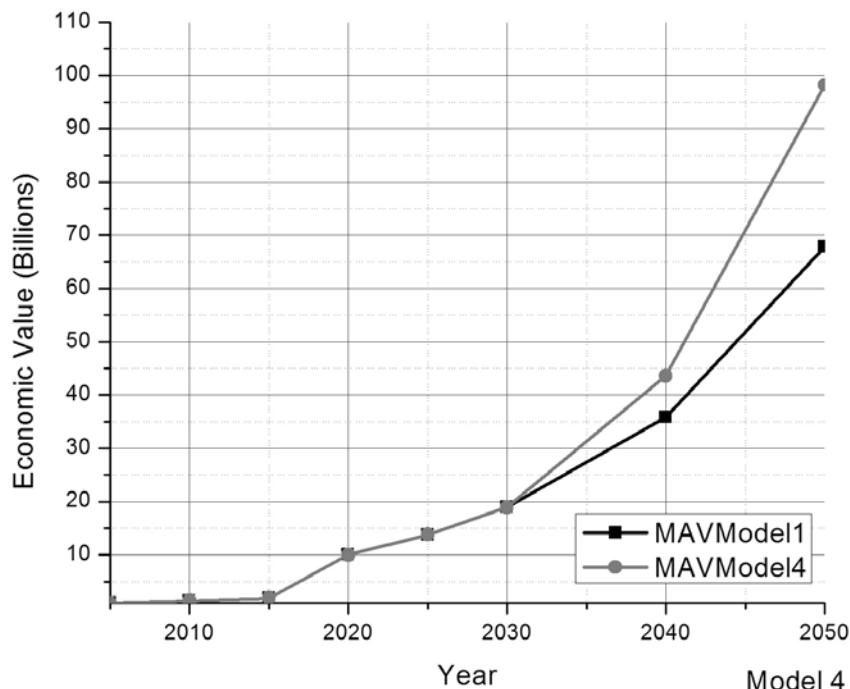


Fig. 13.11 Model 1 Alternate Economic Projections (Sylvan et al. 2009)

Note that a small change in growth rate, brought about by a timely investment (in this case on Mars), creates a substantial alteration in the size of the Martian economy within a mere two decades.

These preliminary examples show just some of the types of questions that may be addressed through the development of space economy models. Technology in general and the space economy in particular have significant potential to act as driving engines for future economic growth on Mars, and indeed wherever human beings have the desire to go and the determination to settle (Collins 2002; USDOC 2001; Zubrin 1995, 1996).

13.9 Conclusion

Our ever increasing ability to extract and make use of energy has enabled us to gain greater mastery of the world around us. In a similar manner, it shall empower us to extend human presence and action to the new world which awaits us. But with the technologies which enable our control comes increased dependence, both on the technologies themselves and on the energy required for their operation. On Mars, this will be the case to a greater extent than ever before in human history.

Rich sources of energy will enable early human inhabitants of The Red Planet to provide for their essential sustenance. Further, it will power the fires of industry which will process the raw resources of the planet into the materials needed to transform a barren world into a new land of promise and opportunity. As population increases, the demands for energy will change and grow, and local sources of energy will be exploited. First material independence and later energy independence will provide the Martian society will the essential strength required for a truly self sufficient branch of human civilization.

Future human settlement and operations on the planet Mars must be planned and executed by those with a thorough understanding of the economics of energy, as such understanding will provide critical insight. Energy is the common frame of reference into which the costs of all materials and operations can be translated, regardless of the planet of origin of those elements. As such, and considering its intrinsic value to Martian society, energy provides an ideal basis for the Martian unit of currency. It will enable Martian planners to make intelligent decisions about the design and development of infrastructure, of the details of operations, and of the dynamics of trade both on the Martian surface and with other domains within the inner solar system.

References

- Boedeker Plastics, Polycarbonate Specifications (2009),
http://www.boedeker.com/polyc_p.htm
- Collins, P.: Meeting the Needs of the New Millennium: Passenger Space Travel and World Economic Growth. *Space Policy* 18(3), 183–197 (2002)
- Cook, E.: The Flow of Energy in an Industrial Society. *Scientific American* 224(3), 135–143 (1971)

- Drake, B.G.: Reference Mission Version 3.0. Addendum to the Human Exploration of Mars: The Reference Mission of the NASA Mars Exploration Study Team. NASA JSC Exploration Office Report. NASA/SP-6107-ADD (1998), http://ston.jsc.nasa.gov/collections/TRS/_techrep/SP-6107-ADD.pdf
- Folta, K.M., Koss, L., McMorrow, R., Kim, H., Kenitz, J.D., Wheeler, R.M., Sager, J.C.: Design and Fabrication of Adjustable Red-Green-Blue LED Light Arrays for Plant Research. *BMC Plant. Biol.*, 5–17 (2005)
- Francis, E.M., Delta, V.: Nomogram v01. Atomic Rockets (2008a),
<http://www.projectrho.com/rocket/dvNomogram01.pdf>
- Francis, E.M.: What's the Mission? Atomic Rockets (2008b),
<http://www.projectrho.com/rocket/rocket3b.html>
- Heinlein, R.: Farmer in the Sky. Charles Scribner's Sons, New York (1950)
- Hublitz, I.: Heat and Mass Transfer of a Low Pressure Mars Greenhouse: Simulation and Experimental Analysis. Ph.D. Dissertation, University of Florida, Gainesville (2006)
- IEA (International Energy Agency) Statistics Division, Energy Balances of OECD Countries (2008 edn.) and Energy Balances of Non-OECD Countries (2007 edn.). IEA, Paris (2007), <http://data.iea.org/ieastore/default.asp>
- Lynch, D.L.: Natural Gas Treads a Global Path. *USA Today*, Money (2005), http://www.usatoday.com/money/industries/energy/2005-12-19-lng-usat_x.htm
- Petrov, G., Homnick, M., Mackenzie, B., Palaia, J.: A Permanent Settlement on Mars: The Architecture of the Mars Homestead Project. In: 35th Int. Conf. on Environ Systems (ICES), Rome, Italy, paper 05ICES-466 (July 2005)
- Owen-Going, T.N.: The Agricultural Concept: A Nutritional Review and Greenhouse Organization Design Outlook. 4Frontiers Corp. internal report, April 21 (2007)
- Rodriguez, P.I., Markovitch, R.: Space Station Freedom Seal Leakage Rate Analysis and Testing Summary: Air Leaks in Ambient Versus Vacuum Exit Conditions. NASA TM-103604 (1992)
- Ruess, F., Schaenzlin, J., Benaroya, H.: Structural Design of a Lunar Habitat. *J. Aerospace Eng.* 19(3), 133–202 (2006)
- Steckel, R.H.: A History of the Standard of Living in the United States. EH.Net Encyclopedia (July 2005),
<http://eh.net/encyclopedia/?article=steckel.standard.living.us>
- Sylvan, R., Komerath, N., Woellert, K., Homnick, M., Palaia, J.: The Emerging Inner Solar System Economy. *World Future Review*, World Future Society, March 2009 Issue (2009)
- US Department of Commerce. Trends in Space Commerce. Office of Space Commercialization report prepared by Futron Corporation (2001)
- Wikipedia (2009a), http://en.wikipedia.org/wiki/Delta-v_budget
- Wikipedia, List of countries by Human Development Index (2009b), http://en.wikipedia.org/wiki/List_of_countries_by_Human_Development_Index#Complete_list_of_countries
- Wikipedia, Tensile Strength (2009c),
http://en.wikipedia.org/wiki/Tensile_strength
- Wikipedia, Tsiolkovsky rocket equation (2009d),
http://en.wikipedia.org/wiki/Tsiolkovsky_rocket_equation
- Zubrin, R.: The Economic Viability of Mars Colonization. *J. British Interplanetary Soc.* 48(10), 407–414 (1995)
- Zubrin, R.: The Case for Mars. Simon & Shuster, New York (1996)

Chapter 14

Tumbleweed: A New Paradigm for Surveying the Surface of Mars for In-situ Resources

Kimberly R. Kuhlman¹, Alberto Behar², Jack Jones², Penelope Boston³, Jeffrey Antol⁴, Gregory Hajos⁴, Warren Kelliher⁴, Max Coleman², Ronald Crawford⁵, Lynn Rothschild⁶, Martin Buehler⁷, Greg Bearman⁸, and Daniel W. Wilson²

¹ Planetary Science Institute, Tucson

² Jet Propulsion Laboratory, Pasadena

³ Earth & Environmental Science Dept., New Mexico Tech, Socorro, NM

⁴ NASA Langley Research Center, Hampton, VA

⁵ University of Idaho, Moscow, ID

⁶ NASA-Ames Research Center, Moffett Field, CA

⁷ Decagon Devices, Pullman, WA

⁸ Snapshot Spectra, Pasadena, CA, USA

14.1 Introduction

Mars missions to date have interrogated the planet at very large scales using orbital platforms or at very small scales intensively studying relatively small patches of terrain. In order to facilitate discovery and eventual utilization of Martian resources for future missions, a strategy that will bridge these scales and allow assessment of large areas of Mars in pursuit of a resource base will be essential. Long-range surveys of in-situ resources on the surface of Mars could be readily accomplished with a fleet of Tumbleweeds - vehicles capable of using the readily available Martian wind to traverse the surface of Mars with minimal power, while optimizing their capabilities to perform a variety of measurements over relatively large swaths of terrain. These low-cost vehicles fill the niche between orbital reconnaissance and landed rovers, which are capable of much more localized study. Fleets of Tumbleweed vehicles could be used to conduct long-range, randomized surveys with simple, low-cost instrumentation functionally equivalent to conventional coordinate grid sampling. Gradients of many potential volatile resources (e.g. H₂O, CH₄, etc.) will also tend to follow wind-borne trajectories thus making the mobility mode of the vehicles well matched to the possible target resources. These vehicles can be suitably instrumented for surface and near-surface interrogation and released to roam for the duration of a season or longer, possibly on the residual ice cap or anywhere orbital surveillance indicates that usable resources may exist. Specific instrument selections can service the exact exploration goals of particular survey missions. Many of the desired instruments for resource discovery are currently under development for in-situ applications, but have not yet been miniaturized to the point where they can be integrated into Tumbleweeds. It

is anticipated that within a few years, instruments such as gas chromatograph mass spectrometers (GC-MS) and ground-penetrating radar (GPR) will be deployable on Tumbleweed vehicles. The wind-driven strategy conforms to potential natural gradients of moisture and potentially relevant resource gases that also respond to wind vectors. This approach is also useful for characterizing other resources and performing a variety of basic science missions. Inflatable and deployable structure Tumbleweeds are wind-propelled long-range vehicles based on well-developed and field tested technology (Antol et al., 2005; Behar et al., 2004; Carsey et al., 2004; Jones and Yavrouian, 1997; Wilson et al., 2008). Different Tumbleweed configurations can provide the capability to operate in varying terrains and accommodate a wide range of instrument packages making them suitable for autonomous surveys for in-situ natural resources. Tumbleweeds are lightweight and relatively inexpensive, making them very attractive for multiple deployments or piggybacking on larger missions.

14.2 History and Development of Tumbleweed Vehicles

Tumbleweeds are large, lightweight, spherically shaped wind-propelled vehicles that can enable exploration of vast areas of Mars. A variety of vehicles referred to as Tumbleweeds and inspired by the Russian thistle (*Salsola tragus*) have been investigated by numerous groups of investigators. Jacques Blamont of NASA's Jet Propulsion Laboratory (JPL) and the University of Paris originally conceived the first known Mars wind-blown ball in 1977, shortly after the Mars Viking Landers discovered that Mars has a thin CO₂ atmosphere with relatively strong winds (Blamont, 1977). Blamont's "Mars Balls" were conceived as relatively large, 3-to 10-meter diameter inflatable balls that could carry payloads, of 20-30 kg for distances of at least 100 km (Janes, 1989). These proposed balls could be powered either by the wind or powered and steered by an inner drive mechanism.



Fig. 14.1 Original 3-wheeled inflatable rover shown with inventor, Jack Jones, NASA Jet Propulsion Laboratory (Jones et al., 1999).

14.2.1 Inflatable Tumbleweeds

In 2000, Jack Jones of NASA JPL was testing a three-wheeled inflatable rover (Fig. 14.1) in a windy sand dune area in California's Mojave Desert when one of the wheels broke off and took off over the sand dunes, while Jones' crew chased the ball with a dune buggy (Jones, 2001). The renegade 1.5 m diameter ball was able to climb steep slopes, over large boulders, and through the jagged brush without hesitation. This seemingly unlucky incident produced the inspiration for the current Tumbleweed vehicle (Behar et al., 2004). JPL then went on to measure performance of a 1.5 m sphere in the Mojave Desert (Jones, 2001), which was confirmed by theoretical analyses performed by the University of Southern California (Wang et al., 2002). The inflatable Tumbleweed has since successfully been tested in Greenland in 2003 and in Antarctica in 2004 (Fig. 14.2). The latest version of the rover was deployed in Greenland in May 2004, where it autonomously traveled more than 200 km across an ice sheet during a 4-day period. Communicating via the Iridium satellite network, the vehicle successfully and reliably relayed live GPS, temperature, and pressure data to a ground station at JPL.

Modeling and testing have shown that an inflatable 6 meter diameter Tumbleweed is capable of climbing 25° hills, traveling over 1 meter diameter boulders, and ranging over a thousand kilometers of terrain (Wang et al., 2002). Tumbleweeds



Fig. 14.2 NASA JPL Tumbleweed test deployment in Antarctica in 2004. Image courtesy Alberto Behar, NASA JPL.

have a potential payload capability of about 10 kg and could potentially generate 10-20 W of power by means of using an internal kinetic energy production device (Jones, 2009). Stopping for measurements can be accomplished using partial deflation or other braking mechanisms (Fig. 14.3). Carnegie Mellon University (CMU) has also conducted empirical testing of the JPL inflatable concepts (Apostolopoulos et al., 2003). The primary purpose of CMU's test was to

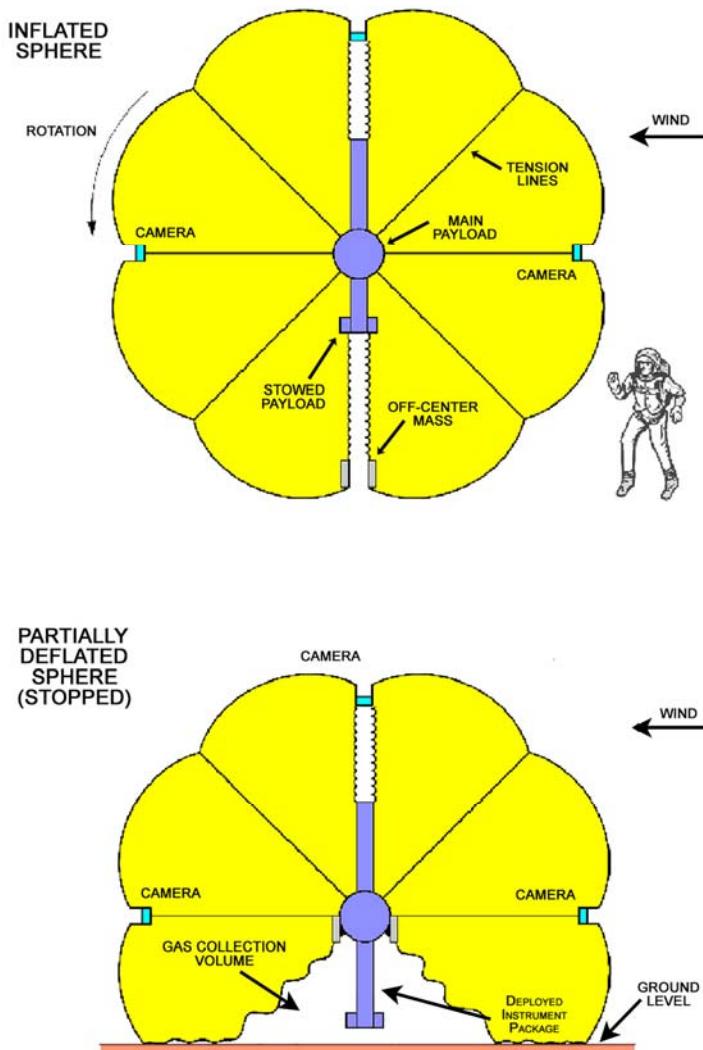


Fig. 14.3 An inflatable Tumbleweed can be stopped by partially deflating the ball and pulling on one of the central payload tension cords to create the “Turtle mode.” This mode of stopping also creates a collection chamber for gas measurements (Jones, 2001).

characterize the rolling resistance, drive torque, drive power and tire wear of a single inflatable sphere, for use on the three-wheeled rover, which utilized the inflatable balls for wheels. A testbed apparatus was developed for these tests that allowed variation of tire design, wheel speed/acceleration, tire pressure, soil/obstacle properties and traverse length.

Inflatable Tumbleweeds have been tested by NASA JPL in Greenland and Antarctica carrying a complete central payload consisting of batteries, inflation/deflation pumps, communications, and a winch (similar to Fig. 14.4). The winch can be used to pull on one or more of the central payload tension lines while the ball deflates. This will create a “turtle” shape and allow the ball to stop, forming a volume underneath that can be used to collect gases emanating from the soil that might indicate hidden subsurface resources (Fig. 14.3). This chamber is very conducive to the collection of gases because its volume to basal area ratio can be very small, providing more rapid feedback to the concentration gradient driving molecular diffusion across the surface (Livingston and Hutchinson, 1995).

14.2.1.1 Example of Tumbleweed Capability: Terrestrial Test Payload

The current terrestrial Tumbleweed payload is suspended at the center of the vehicle (Fig. 14.4). The payload currently consists of a motherboard, a liquid crystal display (LCD), a 900 MHz serial transmitter, an Iridium modem with integrated global positioning system (GPS) receiver, an omni-directional Iridium antenna, an active GPS antenna, a lithium battery pack, a pulse modulated voltage regulator board, a Darlington transistor board, and an air pump. A composite flange with pliable rubber gaskets attaches the air intake to the nylon bag. Similar flanges and gaskets will be implemented on the skin to house the moisture-sensing units.

14.2.1.1.1 Ground Station

To facilitate the field-testing of Tumbleweed, as well as the long-range terrestrial deployments, a highly functional, robust, and user-friendly ground station software has been developed using LabVIEW (Virtual Instrument Engineering Workbench). The LabVIEW ground station software as written, allows for near real-time data processing and distribution via the Internet.

14.2.1.1.2 Electronics Package

The motherboard consists of numerous components that serve to control Tumbleweed, as well as take scientific data. Mounted on the board are two pressure transducers (one for ambient pressure and the other for monitoring the membrane’s internal pressure), a thermocouple (for recording ambient temperature), three 2-axis accelerometers (to determine the orientation of the Tumbleweed at the time of acquisition), and a real-time clock (for noting the time at which the readings were made).

The heart of the rover’s electronics package is the Basic Stamp microcontroller, which is also mounted on the motherboard. The microcontroller takes temperature, pressure, accelerometer, battery level, time, and position data once every second. Every fifteen minutes, the microcontroller attempts to make a call with the Iridium modem, which includes an integrated GPS receiver. If a connection is

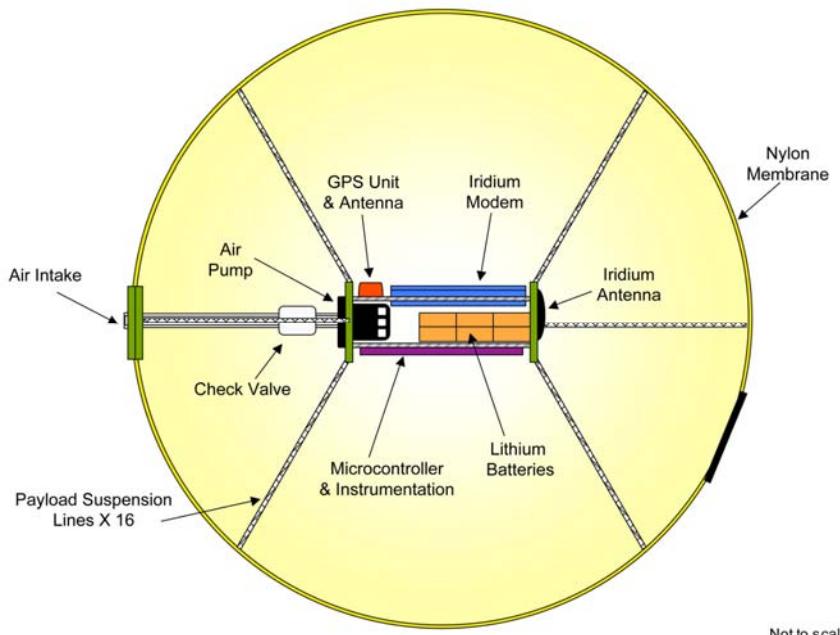


Fig. 14.4 A system diagram showing the inflatable Tumbleweed internal components. Many of these components are part of Tumbleweeds that have been well tested in extreme terrestrial environments by NASA JPL.

made, the Basic Stamp begins sending the stored strings of data. If the connection is unsuccessful, the Stamp will try once more. If the second attempt is unsuccessful, the Stamp will continue taking new data and wait fifteen minutes before making the next attempt.

The power system of the Tumbleweed is designed to satisfy the electrical requirements of six different subsystems currently used for housekeeping and general measurements during polar testing. To handle the various power sources, a pulse modulated voltage regulator is used to knock any voltage from 6 to 16 V down to 5 V.

14.2.1.1.3 Batteries

The factors affecting battery choice are numerous and varied. The weight and number of batteries are given special consideration in order to maintain a balanced load in Tumbleweed such that it rolls along a preferred axis. The maximum weight and number are also affected by the limitation placed on the position of the batteries by the requirement that the modem antenna must remain centrally located to ensure good signal quality. The modem itself experiences current spikes over 4 amps, a level that most batteries are not capable of providing.

Lithium batteries are the clear choice when it comes to power density, weight, and operating temperature, however there are not many lithium batteries available able to

handle the current spikes induced by the modem. Although the lithium batteries are the same physical size as Ni-MH, they operate at more than twice the voltage.

14.2.1.1.4 Pumps

In the long-range deployment scenarios planned, Tumbleweed may experience a significant change in altitude. If measures are not taken to counteract the loss of internal pressure occurring from altitude change, not to mention small leaks, the vehicle could deflate to the point that it is no longer possible to roll, well before reaching its final destination. To alleviate this problem, a small pump is included in the central payload to actively inflate Tumbleweed during the course of its journey.

A second pump/valve system has been designed to actively inflate and deflate the Tumbleweed membrane at a relatively high rate as a means of controlling its speed, as well as providing a mechanism for stopping the vehicle. This system is controlled via the onboard Stamp microcontroller and allows the user to command Tumbleweed (via the ground station software) to travel at a certain speed. The microcontroller then compares this desired speed with the actual speed determined by the onboard navigation unit. If the actual speed is not within range of the desired speed, Tumbleweed automatically adjusts its level of inflation (i.e. -- inflate to speed up, deflate to slow down).

14.2.2 Deployable Structure Tumbleweeds

Tumbleweeds utilizing lightweight deployable structures to harness the wind for mobility have been developed by NASA Langley Research Center (LaRC). LaRC engineers were inspired by the Mars Pathfinder airbag landing system, which traveled a significant distance across the surface of Mars (much farther than the wheeled Sojourner rover ultimately would travel) before coming to a rest and deflating. Various methods were considered for maintaining the rolling motion, with the Martian wind appearing to be the most promising (Antol et al., 2003). Leveraging LaRC's expertise in lightweight structures, several notional concepts of Tumbleweeds (Figs. 14.5 and 14.6) were defined with the goal of providing vehicles with superior aerodynamic properties for capturing the wind (Antol, 2005).

The "Box Kite" concept uses fabric sails, similar to a kite, but with the sails attached to spring hoops to provide increased rolling capability. The "Dandelion" concept was biomimetically inspired with the objective of creating a branch structure similar to that of a Tumbleweed plant. However, the configuration evolved into a symmetric array of legs extending from a spherical core and having pads at the ends to prevent sinking into soft surfaces, thus resembling a dandelion more than a Tumbleweed. A variation of the Dandelion that more closely resembles the Tumbleweed plant is the "Eggbeater Dandelion," which replaces the legs with multiple curved struts resembling eggbeaters or whisks. The "Tumble-cup" consists of open-ended cones around a spherical core to maximize aerodynamic surface area while reducing rolling resistance. The open configurations of the deployable structure Tumbleweed concepts have the additional advantage of allowing unobstructed access to the environment for scientific instrumentation.

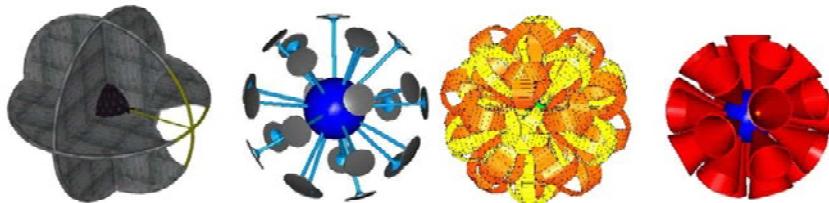


Fig. 14.5 NASA LaRC Tumbleweed Concepts (left to right): Box kite, Dandelion, Egg-beater, and Tumble-cup. Image courtesy NASA/AMA Inc.

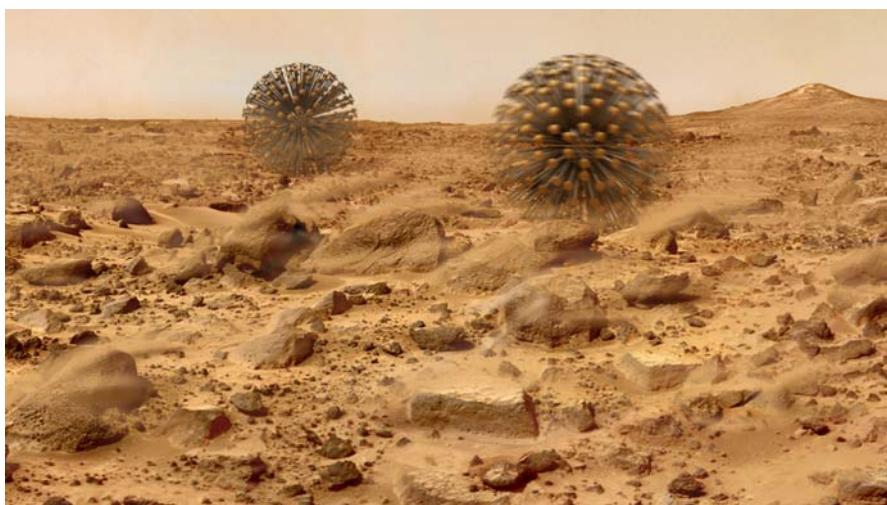


Fig. 14.6 Artist depiction of NASA Langley Research Center Dandelion Tumbleweeds on Mars. Image courtesy NASA/AMA Inc.

Preliminary analysis and wind tunnel testing of the LaRC notional concepts (see Sect. 14.2.2.1) determined that the deployable structure Tumbleweeds require a diameter of approximately 4-6 meters (m) with a mass of no more than 20 kg, including subsystems and instruments, to capture the thin Martian atmosphere and achieve mobility (Antol et al., 2005). The Box Kite Tumbleweed has emerged as the most promising configuration because of its drag properties, packaging efficiency, and open architecture that maximizes access to the environment for sensors and instrumentation.

The first scale prototype of a deployable structure Tumbleweed, based on the Box Kite concept, was developed by students from the North Carolina State University (NCSU) Department of Mechanical and Aerospace Engineering in cooperation with LaRC (Fig. 14.7). Known as the Tumbleweed Earth Demonstrator (TED), the TED includes a central instrument core with temperature/pressure sensors, accelerometers, a Global Positioning System (GPS) package, and a data

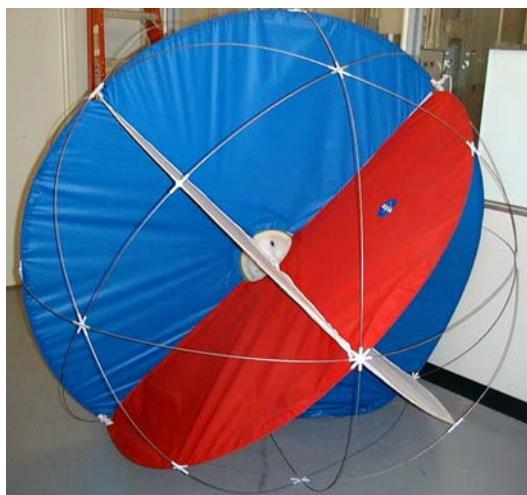
Fig. 14.7 Tumbleweed Earth Demonstrator designed and fabricated by North Carolina State University. Image courtesy NCSU.



acquisition/communication system for collecting data during Earth-based field tests of the concept (Hanrahan et al., 2003). NCSU has also developed a second and third generation TED, employing them in field tests on a variety of terrain (Wilson et al., 2008).

Based on preliminary analysis results and leveraging knowledge gained from development of the NCSU TED, NASA LaRC constructed a prototype box kite (Fig. 14.8) for demonstrating mission concepts and science measurement techniques (Antol et al., 2006a). As with the NCSU TED, the LaRC Box Kite prototype has a central core with sensors and a data acquisition/communication system. The primary

Fig. 14.8 NASA LaRC Box Kite Tumbleweed Prototype.



difference between the NCSU TED and the LaRC prototype is the use of titanium rods in the hoop structures for reduced weight and durability.

The Texas Tech University (TTU) Department of Mechanical Engineering is developing extremely lightweight, miniaturized, deployable structure Tumbleweed concepts, which would permit large numbers to be carried as secondary payloads on Mars missions. The TTU Tumbleweeds, with the configuration of a sector-removed sphere (Fig. 14.9), would be approximately 12 inches in diameter and have a mass less than a kilogram. Constructed of lightweight materials, the TTU Tumbleweed could be covered with solar cells and have electronics and sensors embedded in the structure (Rose et al., 2006).

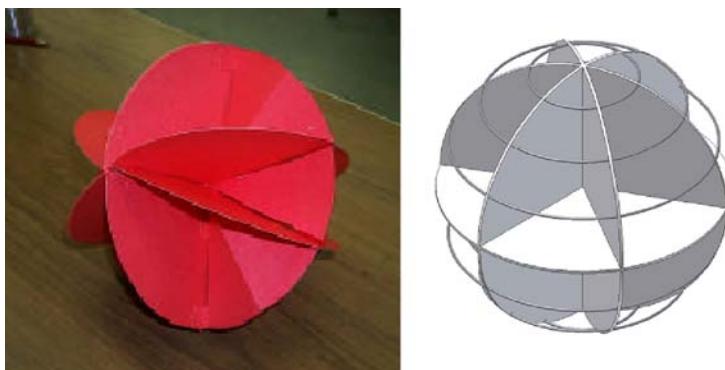


Fig. 14.9 TTU Tumbleweed concept. Image courtesy TTU.

14.2.2.1 Test and Analysis

NASA LaRC, NCSU, TTU, and the Biorobotics Laboratory at Case Western Reserve University (CWRU) have conducted extensive analysis and wind tunnel testing of Tumbleweed concepts. Dynamic simulations have been developed to investigate the potential mobility of Tumbleweed concepts in a Martian environment. Wind tunnel testing was conducted to obtain drag characteristic data on the various Tumbleweed concepts for input to the dynamic simulations. Another objective of the aerodynamics research is to determine which configurations achieve the highest drag coefficient (C_d), using a simple, smooth sphere (~0.5) as the baseline.

LaRC wind tunnel testing was conducted in the Basic Aerodynamics Research Tunnel (BART). Multiple configurations of the Dandelion, Tumble-cup, and Egg-beater Dandelion were tested in a free stream flow at the expected Reynolds numbers for the martian surface ($Re = 50,000 - 125,000$) (Fig. 14.10).

Analysis of the test data showed the majority of Tumbleweed configurations exceeded the C_d of a sphere. The Box Kite displayed an angle of attack dependency, with the C_d varying from 0.8 up to 1.2 depending on the model orientation. The other concepts, which are symmetric in shape, were consistent across all angles of attack. The C_d for the Tumble-cup and Dandelion models ranged from

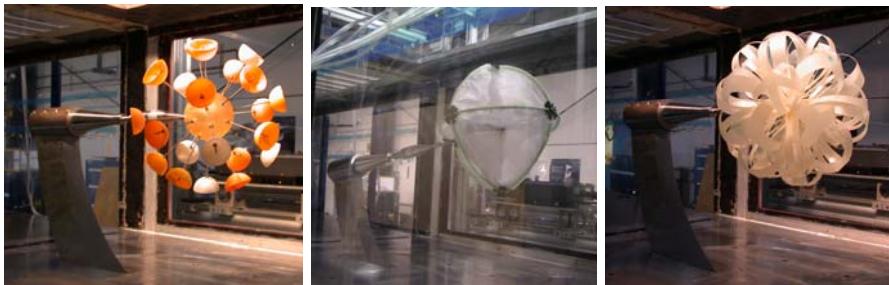


Fig. 14.10 NASA LaRC Dandelion, Box Kite, and Eggbeater Danelion in the Basic Aerodynamics Research Tunnel.

0.8 to 1.0, varying with the number of legs/cups employed, while the C_d for the Eggbeater Dandelion models ranged from 0.6 to 0.85. Several alternative configurations of the Dandelion, equipped with cupped feet, performed no better than a sphere (Antol et al., 2006b).

Testing was also conducted in the TTU Atmospheric Boundary Layer (ABL) Wind Tunnel to study the surface boundary layer effects on the Tumbleweed drag coefficients (Fig. 14.11). Several models, including the NASA LaRC Dandelion and Tumble-cup, were fully submerged in a simulated Martian atmospheric boundary layer based on a theoretical NASA model of a Mars surface wind boundary layer. A comparison of the results with those from testing conducted in the free stream reveals a uniform decrease in the drag coefficients between 6 and 16% in the atmospheric boundary layer testing versus the free stream (Rose et al., 2006).

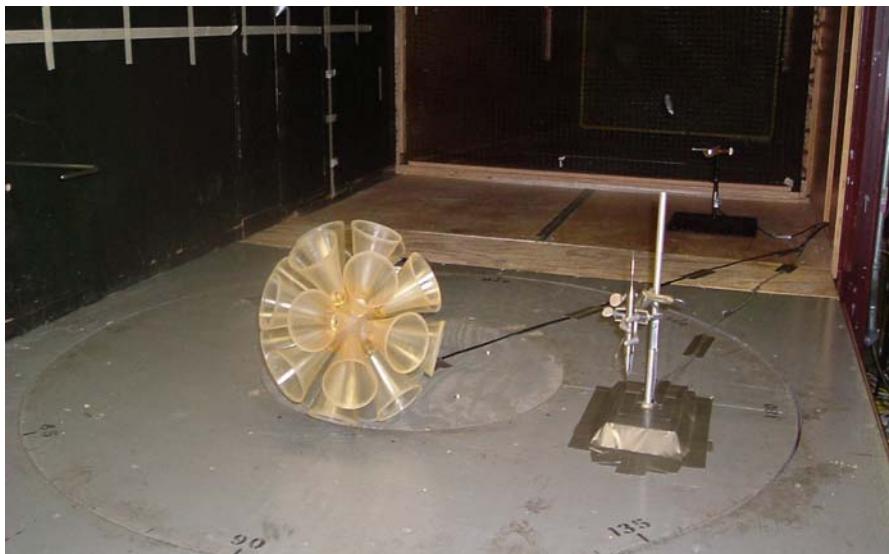


Fig. 14.11 NASA LaRC Tumble-cup model in Texas Tech University Atmospheric Boundary Layer Tunnel. Image courtesy TTU.

14.3 Tumbleweed Deployment

Launch and deployment of Tumbleweed vehicles promises to be simple and relatively low-cost. A small number of Tumbleweeds could piggyback along with other missions, or fleets of Tumbleweeds could be launched together and released at one or multiple locations depending upon the coverage desired. Mobility of a wind driven system is inversely proportional to mass and directly proportional to atmospheric density and drag coefficients. To increase the system mobility, mass should be minimized. Increasing the size of a Tumbleweed will increase the driving force from the wind, but increases the structural mass, which will limit the additional mobility. Communication systems tend to need high levels of electric power. Power systems tend to be heavy to begin with and increase in mass with increased capacity. The preferred approach is to reduce the size and mass of system components such as the structure, power and communication systems.

In order to quantify the potential mobility of deployed Tumbleweed vehicles, NASA LaRC has examined the motion of various Tumbleweed designs over various terrains (sand dunes, rocky desert, glacial ice, etc) to determine the optimal size for instrument suites and anticipated topography (Antol et al., 2005). The *Mars Tumbleweed Monte Carlo Simulator* is software that models the dynamics of a vehicle interacting with the environment of Mars to provide a moderate fidelity, end-to-end mission simulation (Flick and Toniolo, 2005). The software allows users to select value ranges for a series of parameters (e.g., mass properties, wind model parameters, terrain properties, etc.), which define bounds on the simulation environment. Random values are then assigned to each parameter during the simulation for each Monte Carlo iteration based on these user-defined value ranges. The purpose is to analyze closely related scenarios in order to identify common behaviors or trends. Figure 14.12 shows the combined result of 2000 Monte Carlo simulations of dispersing Tumbleweeds 120 seconds after release (assumptions: 5m radius, rigid-body Tumbleweed, constant wind velocity of 7 m/s, 1-2m rocks spaced approximately 10m apart in a randomized distribution based on the Viking 1 landing site data). Approximately 890 Tumbleweeds stuck between two rocks in the first 100m (major groupings of 380 at 0-20m, 280 stuck at 80-100m). The Tumbleweeds would remain stuck until the wind shifts direction or the wind speed increases until the resultant force overcomes the resistance to motion and the Tumbleweed rolls over the obstacles. Mars Pathfinder data show a diurnal wind sweep of at least 180 degrees, so a stuck Tumbleweed can become free within one Sol. Mars Pathfinder data also indicates that wind speed cycles throughout a Sol with an average of approximately 7 m/s and peaks between 10 m/s and 15 m/s (seasonally dependent). The software can be enhanced to use the most recent Mars environment models, including, but not be limited to the Mars terrain (Mars Orbiter Laser Altimeter (MOLA) data), rock distribution (Golombek and Rapp, 1997), wind direction/magnitude, boundary layer effects (ARC Global Circulation Model (GCM)) (Haberle et al., 2003), and MSFC Mars Global Reference Atmospheric Model 2001 (Justus et al., 2006).

LaRC also developed a 3 dimensional (3-D) Matlab/Simulink simulation using a lumped mass dynamics model to study the rolling characteristics of the deployable

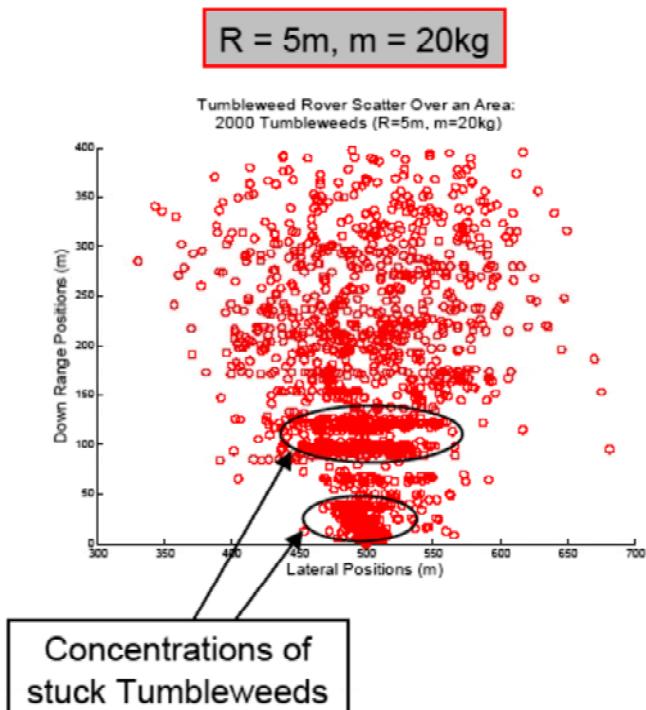


Fig. 14.12 Monte Carlo simulation of Tumbleweed trajectories run for a total of 2000 rovers, with diameter 5 m and mass of 20 kg mass.

structure Tumbleweed concepts. To refine the lumped mass dynamics model approach of the 3-D Simulation and to validate modeling assumptions, a series of empirical tests were conducted at LaRC using photogrammetry techniques. Three Tumbleweed concepts (a simple sphere, a Box Kite, and an Eggbeater Dandelion) were modeled and rolled down a 24-foot plywood ramp and the rolling/bouncing movement recorded using a 6-camera photogrammetry system. Cameras were located to capture 6 degree-of-freedom (DOF) dynamics at start of motion, middle of the ramp, and at the end of the ramp. Three-dimensional motion data was extracted from the test data and used to refine the assumptions and validate the 3-D model.

A stochastic simulation of a group of Tumbleweeds interacting with obstacles was developed by students of the CWRU Biorobotics lab. Mars Orbiter Laser Altimeter (MOLA) mission data was used to create the general topography of the martian surface. Valleys and other geological obstacles were simulated using a Gaussian randomization of the entire surface with a standard deviation of twenty-five degrees. The simulation also employed a group behavior algorithm with Tumbleweed vehicles having limited capability for stopping/starting and steering. A mission scenario to explore Martian gullies, based upon the Dao Vallis topographical

data, was analyzed to assess Tumbleweed mobility and group behavior capabilities. An animation of the mission scenario also was created (Fig. 14.13) (Hoeg et al. 2006).

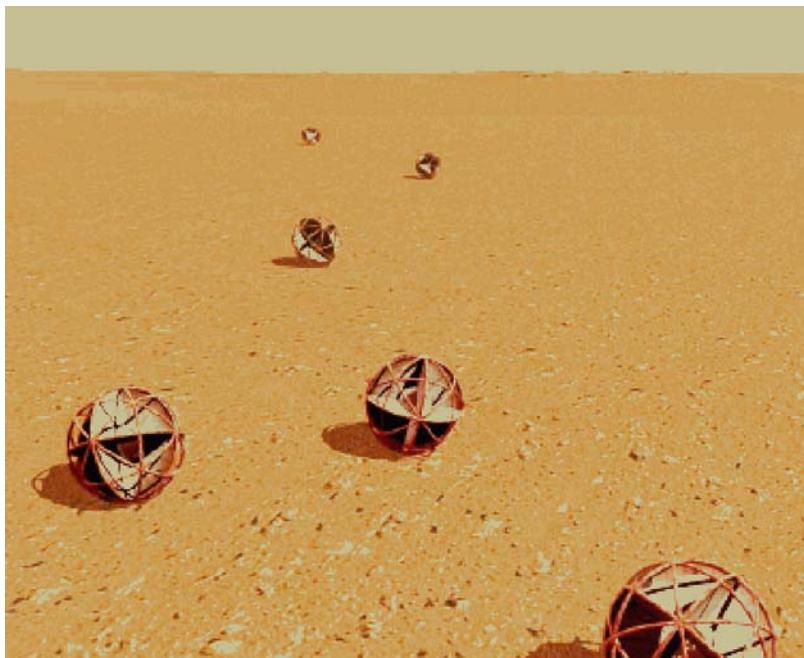


Fig. 14.13 Artist depiction of a group of Box Kite Tumbleweeds crossing a Martian plateau. Image courtesy CWRU.

A preliminary assessment of navigation system options by NASA LaRC identified several possibilities for accurately determining the location of a Tumbleweed rover on Mars, including: Doppler radiometric techniques using existing Mars orbiting assets, odometry using Inertial Measurement Units (IMU), optical terrain mapping, and optical celestial navigation (Antol, 2005). Dead-reckoning using an IMU or optical position updates from terrain mapping could be used for instantaneous navigation, combined with radiometric techniques for regular position updates. The recommended configuration is a Tumbleweed equipped with a Proximity-1 radio with Doppler capability, a roll sensor, Micro-Electro-Mechanical Systems (MEMS) accelerometers, a MEMS IMU (if the technology matures), and 6 MEMS Sun sensors. It is important to note that this recommendation imposes no new Mars infrastructure; however, technology advancements may be needed in the above components to meet Tumbleweed mass and power requirements.

14.4 Target *In situ* Resources

Since the potential resources of Mars have been discussed in detail in other chapters in this book, we will briefly touch upon the resources that can potentially be

surveyed using a fleet of Tumbleweed vehicles. It should also be noted that Tumbleweeds and their highly configurable payloads could perform a variety of basic science missions. The most important resource for human habitation is arguably water. A well-defined supply of water is deemed a necessity for a base that is expected to flourish and grow (Taylor, 2001). Tumbleweeds can survey for water in two ways: surface mounted sensors and ground penetrating radar (GPR). Both will be discussed in the following section on instrumentation. GPR in particular will be capable of characterizing the location and abundance of near-surface water, critical for locating a future Martian base (Jakosky and Zent, 1993). In addition to ice and underground aquifers, water may be present in the pore spaces of clays and other hydrated minerals, making these materials potential important water resources on Mars (Baker et al., 1993). Tumbleweeds outfitted as proposed below would be capable of identifying areas of hydrothermal activity and thus deposits of hydrated minerals, such as those found in Nili Fossae by Mars Express (Bibring et al., 2006; Poulet et al., 2005) and Mars Reconnaissance Orbiter (MRO) (Mustard et al., 2008).

Hydrogen and oxygen (both likely derived from water) are key components in the synthesis of fuels and propellants (Stoker et al., 1993), Tumbleweeds would also be useful components in the search for quantities of water that will be needed to fuel a base on Mars and provide transportation off the surface. Tumbleweeds equipped with gas sensors could also be used to locate sources of methane plumes observed on Mars in recent years (Krasnopolsky et al., 2004; Mumma et al., 2009). These sources could be important for production of fuels and propellants.

Another of the future resources that human colonists will need to have on Mars is soil for agricultural activities (McKay et al., 1993; Taylor, 2001). Ever since Viking, there have been indications that the regolith contains montmorillonite (a swelling clay) and other components that may make it difficult to use for agriculture (Bibring et al., 2006; Poulet et al., 2005). Information about pH (seeking the neutral soil), clay types, heavy metal contents, etc. would be target information that would pertain to this goal. Both a multispectral imager and X-ray fluorescence spectrometer would allow Tumbleweeds to survey these properties of the martian regolith (Marshall et al., 1997). These instruments would also be critical to locating sources of nitrates and phosphates, which will be critical to the agricultural enhancement of regolith local to a martian base (Stoker et al., 1993; Taylor, 2001). Tumbleweeds would also be well suited to survey for sulfur compounds, which would be important for a variety of industrial processes that one would assume to be important to a base on Mars.

Conversely, recognition of perchlorate (and possibly chlorate) as discovered recently by the Phoenix mission (Hecht et al., 2008) would be important for two reasons. Natural perchlorate on Earth is thought to be formed by photolytic processes in the atmosphere (Erickson, 1981). Similar processes produce nitrate. Thus, perchlorate could be a pathfinder for much larger nitrate deposits. Both of these materials are very soluble salts and can accumulate on the Earth's surface in very arid regions, like the Atacama Desert, where the average perchlorate forms less than 0.05% of nitrate deposits (Erickson, 1983). While nitrate is a valuable fertilizer, perchlorate and, to a much greater extent, chlorate, are herbicides. Concentrations

of perchlorate of up to 0.8% have been reported in nitrate fertilizer from Chile (Bohlke et al., 1997). Perchlorate is a powerful oxidant and could be a useful energy source in the same way as it is used in solid fuel rockets.

Tumbleweeds outfitted with similar instruments will also be useful for locating resources of building materials. Aggregates, calcium-sulfates and carbonates are all important materials for building infrastructure (Schmitt, 2004; Taylor, 2001). Calcium-rich sulfates, most likely gypsum, have been observed in several locations on Mars by OMEGA/Mars Express (Gendrin et al., 2005; Langevin et al., 2005). Carbonates have recently been observed at the Phoenix lander site (Boynton et al., 2009; Kounaves et al., 2009; Sutter et al., 2009). Even after a base is established, Tumbleweeds may be used to survey for long-term resources such as ore-bodies, sources of metals, organic compounds and extensive clay deposits. Tumbleweeds will be particularly useful for long-range surveys for resources that require intensive, global geological exploration suggested by Taylor, 2001: sedimentary deposits, hydrothermal deposits, and differentiated igneous provinces. Taylor (2001) suggests that these global searches for resources be started early in the process in order to attract capital for martian investment. Fleets of low-cost Tumbleweeds could play an integral role in such reconnaissance.

14.5 Instrumentation for ISRU Surveys

A suite of instrumentation can be envisioned for a fleet of Tumbleweeds for deployment on Mars for *in situ* resource surveys. An example instrument suite could include surface mounted soil moisture sensors (SMSMS), ground penetrating radar to characterize subsurface layering, aquifers and voids, sensors for a variety of useful gases, a miniature X-ray fluorescence spectrometer for elemental analysis of martian regolith and a multispectral imaging system for characterizing grain size and shape distributions as well as surface mineral composition. Other suites of instruments could be envisioned based upon the requirements of specific survey scenarios.

14.5.1 Surface Mounted Soil Moisture Sensors (SMSMS)

During the wind driven phase, surface mounted soil moisture sensors (SMSMS) embedded in the skin of Tumbleweed will measure the soil volumetric water content (VWC) to a depth of 5 cm. Layers of water ice may well exist below 5 cm as buried snow, ice lenses, or ice wedges (Feldman et al., 2008). The SMSMS will be used in combination with real-time ground penetrating radar (GPR) as suggested in the following section. The measurement time for SMSMS is of the order of 10 ms so the results can be used for exception-based monitoring. That is, as Tumbleweed moves along, the VWC will be monitored so that when measurements exceed a threshold value, the motion of Tumbleweed can be stopped. The Tumbleweed will then transform to operate in the Turtle Mode as discussed earlier in the chapter and shown in Fig. 14.3. In this mode, more detailed measurements of the regolith can be acquired using Tumbleweed's suite of analytical instruments.

The SMSMS probes are implemented as nubs on the surface of Tumbleweed between the treads. The probes' electric field will penetrate the soil to a depth of

about 5 cm so the instrument will sample the water content of this region. Thus, the sensor design will allow Tumbleweed to measure the soil moisture at a depth well below the surface crust. SMSMS is capable of measuring dielectric permittivity that varies from 1 for air, 3 to 6 for dry soils, 30 to 35 for water saturated soils, and 78 for water. The measurement time is approximately 10ms. Its mass is less than 100 g, and it requires less than 100 mW of power. The instrument has two rugged nubs that are separated by approximately 3 cm. With this spacing, SMSMS is capable of sampling the water content up to a depth of 5 cm (Fig. 14.14).

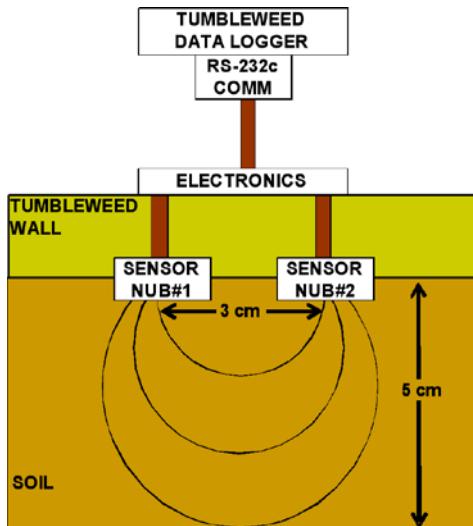


Fig. 14.14 SMSMS embedded in the wall of Tumbleweed showing the nubs separated by 3-cm and the current stream lines penetrating to a depth of 5-cm into the soil.

Operationally, SMSMS can also serve as a tachometer indicating the speed of Tumbleweed. As Tumbleweed rotates, SMSMS will be in air (dielectric permittivity of 1) most of the time and briefly it will be next to the ground (minimum permittivity of 3). Knowing the diameter of Tumbleweed and the time between the ground pulses allows a calculation of Tumbleweed's velocity. Unlike traditional land and Martian rover systems, Tumbleweed allows intelligent sampling through the integration of the SMSMS into its surface. The SMSMS monitors moisture content from 0-5 cm in depth in real time while Tumbleweed is traveling. In turn, anomalous moisture content triggers sampling events. This type of sampling regime, termed "exception-based" sampling, uses real-time information to determine sampling locations rather than site selection based on prior data. The SMSMS system allows Tumbleweed to implement the NASA paradigm, "Follow the Water" in a manner that is *immediately responsive* to its changing surroundings as it travels. We believe this to be a very efficient way of finding and analyzing areas that have the potential for *in situ* resources.

The circuitry for the soil moisture sensors is based on two existing soil moisture sensors; they are the Thermal and Electrical Conductivity Probe (TECP) and ECH₂O Probe. The TECP was built by Decagon Devices, Inc. for the NASA Phoenix Scout Lander, which landed on Mars May 25, 2008 (Cobos et al., 2004). The TECP instrument was designed to measure the dielectric permittivity of the martian regolith, which was assumed to be between 1 and 20; the instrument can resolve the dielectric permittivity to within 0.02. Detecting a dielectric permittivity greater than the anticipated permittivity of 3 to 6 for the regolith is highly suggestive of the detection of liquid water.

The ECH₂O probe, also manufactured by Decagon Devices, Inc measures the dielectric permittivity of the media surrounding two electrodes (Kizito et al., 2008). The soil water content is calculated from the dielectric permittivity measurement of the soil-moisture mixture. The dielectric permittivity of dry soil is between 3 and 4 and the dielectric permittivity of water is 80; so the presence of moisture in the soil is apparent as an increase in the dielectric permittivity. The measurements are comparatively fast, requiring approximately 10 ms for data acquisition.

14.5.2 Ground Penetrating Radar

Ground penetrating radar (GPR) is a non-invasive technique for imaging shallow and midrange subsurface materials and features (Conyers and Goodman, 1997; Doolittle, 1982; Duke, 1990; Huffman III, 1992). Pulses of ultra high frequency microwaves (300 to 1,000 MHz) are transmitted to the ground through a transducer to ~10 m (high resolution). Deep GPR uses low frequency antennae (25 to 200 MHz) with poorer resolution. Different materials reflect differently depending upon differing electrical conductivities and dielectric constants (Kolecki and Landis, 1996; Olhoeft, 1979; Wright et al., 1996). Returning waves (received by the same antenna) are compared with two-way travel time, amplified, and plotted. Portable units have been tested with data collection rates of several hundred scans per second with a time resolution of 5 picoseconds (Kathage et al., 2005), which means that GPR can also be used for real-time exception-based monitoring as discussed in the previous section. High conductivity materials (e.g. shales, clay, etc.) can be imaged only to about one meter, which is of limited utility. Depth of imaging ranges from less than one meter in clay soils (Nettwork, 2002) to greater than 5,400 meters in polar ice (Olhoeft, 2000). The depth of investigation that is possible increases with decreasing frequency but with decreasing resolution. Typical depths of useful investigation in clay-free sands saturated with fresh water or low conductivity materials (granites or dry particulates) are about 30 m.

The ability of GPR to reveal the fine structure of layers and small subsurface features has been documented in a wide variety of applications, including the study of centimeter-scale annual layers in a coastal barrier (Moore et al., 2004). Fine structure and layering in soils was an early target of GPR studies (Boll et al., 1996; Kung and Lu, 1993) including measuring the soil water content profile in sandy soil (Lambot et al., 2004). Periglacial terrains have been examined including spatial variability in glaciers (Palli et al., 2002), and in permafrost (Brandt et al., 2007). GPR has successfully resolved layers in sandstone (West and Truss, 2006) and it has been used increasingly in locating karst features like sinkholes

and caves (Ezersky et al., 2006). The methodologies of GPR continue to be further refined at a rapid pace including integration with other methods like Resistivity Image Profiling (RIP) (Yang et al., 2006) advanced algorithms for handling multiple modes of GPR data (Stanley et al., 2004), and physical laboratory testing of GPR capabilities (Capizzi and Cosentino, 2008).

Both 2D and 3D imaging of up to hundreds of m²/day is done with small portable units. Typically, electrical properties are much more important than magnetic properties in controlling propagation speed and amplitudes. Rock or soil density, chemistry, state (liquid/gas/solid), distribution (pore space connectivity) and H₂O content all contribute. Electrical properties of planetary surfaces (Kolecki and Landis, 1999; Olhoeft, 1991) and lunar regolith H₂O (Olhoeft, 1976) are of great interest. GPR use for Mars has been suggested (Olhoeft, 1998). Permafrost (Olhoeft, 1975, 1977), sand dune structures (Schaber et al., 1986; Schenk et al., 1993), organics (Olhoeft, 1986), and oxide minerals critical for human inhabitation (Lindsley, 1991) are all distinctive. Small GPR units (e.g. GeoModel, Inc., GPRS, Inc., GRORADAR™, Enviroprobe Service, Inc., NGPRS Inc., Geophysical Survey Systems, Inc.) are able to test capability of detecting near subsurface structure, composition and inhomogeneities. While present technology is not mountable on prototype Tumbleweeds, future miniaturization may make this a reasonable choice for a Mars-bound Tumbleweed to survey for *in situ* resources (Grant et al., 2003). The smallest currently commercially available system is the RAMAC X3M by Mala Geoscience, which is used for mapping lake sediments, manmade objects, detection of subsurface cavities and for mapping other small-scale features in detail. At a mass of 5 to 7 kg, this unit is approaching the size that could be mounted on a Tumbleweed.

14.5.3 Gas Sensors

Numerous gases will be useful *in situ* resources on Mars as discussed in Chaps. 15, 19 and 20. The handheld VRAE gas surveyor from RAE Systems could be used as the core of an instrument to measure resource relevant gases; oxygen, ammonia and carbon dioxide. A variety of sensor types are used for the various different gases; catalytic sensors for so-called combustible gases (primarily CH₄), thermal conductivity sensors for percentage volume of combustible gas, and electrochemical sensors for oxygen and toxic gases (including H₂S, SO₂, NO, N₂O, NO₂, Cl, HCN, NH₃, and PH₃). The unit is small enough to fit within the central tube of the inflatable Tumbleweed with a flexible probe that will fit alongside an imaging spectrometer. The VRAE weighs slightly more than 0.5 kg and is easily integrated into the data collection system using a USB connection. The measurements can be made while the Tumbleweed is stopped in Turtle Mode and the winches inside have created a gas collection chamber from the skin of the Tumbleweed. The measurements will be made as a function of time in an attempt to determine the concentration of gases in the soil as a function of depth (Livingston and Hutchinson, 1995).

The detection of water vapor in the soil and atmosphere immediately above the soil is desired to extend the range at which ambient water can be measured using Tumbleweed. Methane and hydrogen peroxide are other important resource gases

that could be surveyed by Tumbleweeds. The JPL Tunable Laser Spectrometer (TLS) is a component of the Sample Analysis at Mars (SAM) analytical suite scheduled to fly on the Mars Science Laboratory (MSL) mission in 2011 (Tarsitano and Webster, 2007). The TLS uses a multipass Herriot cell and four laser sources -- near-infrared, interband cascade (IC) and quantum cascade (QC) – to measure methane, water, hydrogen peroxide, nitrous oxide and carbon dioxide. The TLS is extremely sensitive. For example, it is capable of measuring methane abundances to 0.01 ppbv (Tarsitano and Webster, 2007).

Other miniature tunable diode lasers (TDL) are being developed commercially. For instance, Physical Sciences Incorporated fabricates a GasScan™ Miniature Diode Laser-based Ambient Gas Sensor based upon near-IR tunable diode laser absorption spectroscopy (Physical Sciences Incorporation, 2009). This detector can be configured to detect a variety of gases including water and methane with sensitivities of 1.0 ppm.

14.5.4 Miniature X-ray Fluorescence Spectrometer

Tumbleweeds equipped with miniature X-ray fluorescence spectrometers (XRF) will be capable of performing elemental analyses of martian regolith while the Tumbleweed is in the stopped mode. NASA LaRC has been involved in building planetary flight X-ray Fluorescence Spectrometers (XRFS) since its participation in the development and delivery of flight XRFS units for the Mars Viking Project. The XRFS could also be replaced by an XRF coupled with a miniaturized X-ray diffraction (XRD) detector to provide complete mineralogical identification. Examples of potential instruments include the Mineral Identification and Composition Analyzer (MICA) (Martin et al., 2008) and the Chemistry and Mineralogy instrument (CheMin), which is part of the Mars Science Lander (MSL) payload (Blake et al., 2007). An XRFS unit could be mounted inside the central tube of a Tumbleweed such as the inflatable version shown in Fig. 14.3 and could be coupled with a drilling unit. When the Tumbleweed is deflated into a stopped mode, the drill and XRFS unit could be deployed.

Most recently, an XRFS was miniaturized to be able to be inserted into a 27.1mm diameter drilled hole to conduct elemental analysis of Martian regolith strata layers (Elam et al., 2008). The borehole XRFS provides superior performance to the X-ray elemental analyzers on the Mars rovers Spirit and Opportunity in that analysis time has been reduced to 1000 seconds with the borehole instrument versus the 20 hours required by the rover instruments. The analysis time can be further reduced to 100 seconds, making it an ideal instrument for a Tumbleweed science complement in Turtle Mode. Analytical performance of the borehole instrument was determined by checking the lower limits of detection for several elements over a wide range of the periodic table. The performance of the borehole XRFS was tested in a simulated Mars atmosphere using several terrestrial soil Standard Reference Materials (SRM) (Fig. 14.15). Minimum detection limits were below 10 ppm for most elements. No filters or other optics were used in the incident beam, and the detector has an internal collimator to restrict the

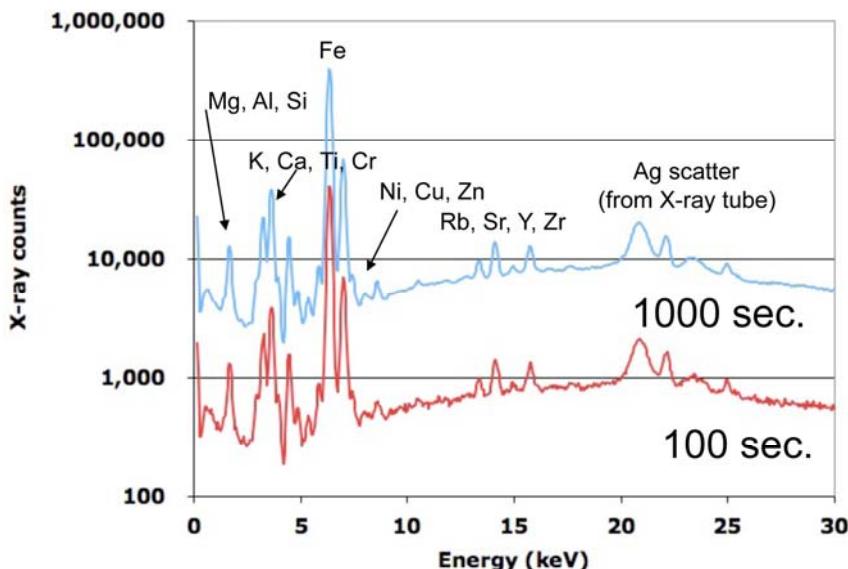


Fig. 14.15 Spectra of SRM 2709 tested using the borehole XRFS in a simulated martian atmosphere. Note that 100 sec spectra provides as much basic info as the 1000 sec does.

beam to the center of the diode. Data collection time was 1000 sec for the upper spectrum and 100 sec for the lower spectrum in Fig. 14.15. Note that the majority of the information is still available even with the 100 sec data collection time. This short data collection will greatly facilitate its use as an instrument on a Tumbleweed vehicle.

14.5.5 Multispectral Imaging

Media for agriculture, aggregates, and structural materials are also subjects for long-range surveys of Mars in preparation for the location of a human outpost. A computed tomography imaging spectrometer (CTIS) designed by NASA JPL will be capable of enabling snapshot spectral imaging by capturing spatial and spectral information in a single frame (Bearman et al., 2007). There are no moving parts or narrow-band filters, and nearly all collected light is passed to the detector at all times. A CTIS captures a scene's spatial and spectral information by imaging the scene through a two-dimensional grating disperser as in Fig. 14.16. This produces multiple, spectrally dispersed images of the scene that are recorded by a focal plane array (FPA) detector. From the captured intensity pattern, CT algorithms can be used to reconstruct the scene into a cube of spatial (x and y) and spectral (wavelength) information.

Operation of the CTIS is illustrated in Fig. 14.16. With this technique, diffractive optics disperses the spectral and spatial information of each pixel onto an imaging

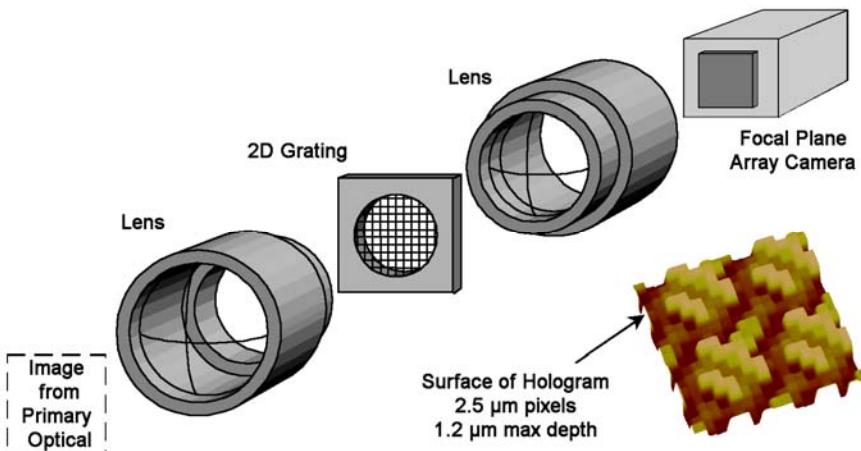


Fig. 14.16 Operation of the computed tomography imaging spectrometer (CTIS).

sensor; an image cube in wavelength space is reconstructed from a *single* image. Note that the CTIS uses just two lenses and a focal plane detector.

The mathematics of the reconstruction is the same as tomographic imaging. It is important to note that each image is not simply composed of single wavelengths; spatial and spectral information from each object pixel is multiplexed over the entire array – Figure 14.17 shows how the spectrum of a single pixel is distributed by the diffractive disperser. A single image contains all the information required to reconstruct the spectral image cube. A CTIS can operate over a large wavelength range, easily from 400 to 800 nm, and with the proper detector can operate in the IR or UV. The data from a single image can be reconstructed in a variety of ways to adjust image size and wavelength bands. It is also possible to reconstruct the image over a narrow spectral range and place the sampled wavelengths at desired locations.

A CTIS device is ideal for a Tumbleweed platform. The snapshot nature of the device means that object or platform motion is not an issue and that one can obtain spectral scenes from a randomly moving and rotating platform. Camera sensors can trigger the CTIS when appropriately pointed, either down at the contact point or sideways at terrain (not up at sky). For example, a CTIS is being used now for retinal spectral imaging, acquiring a complete image cube in the eye in 3 ms (Johnson et al., 2007). The CTIS easily handles eye saccades and other patient motions. Since the CTIS only requirement is the delivery of an image to an aperture plane that defines the viewed image, the CTIS could even be fiber fed so it does not have to be located on the periphery of the Tumbleweed. Location away from the edge would reduce shock and vibration requirements on the instrument.

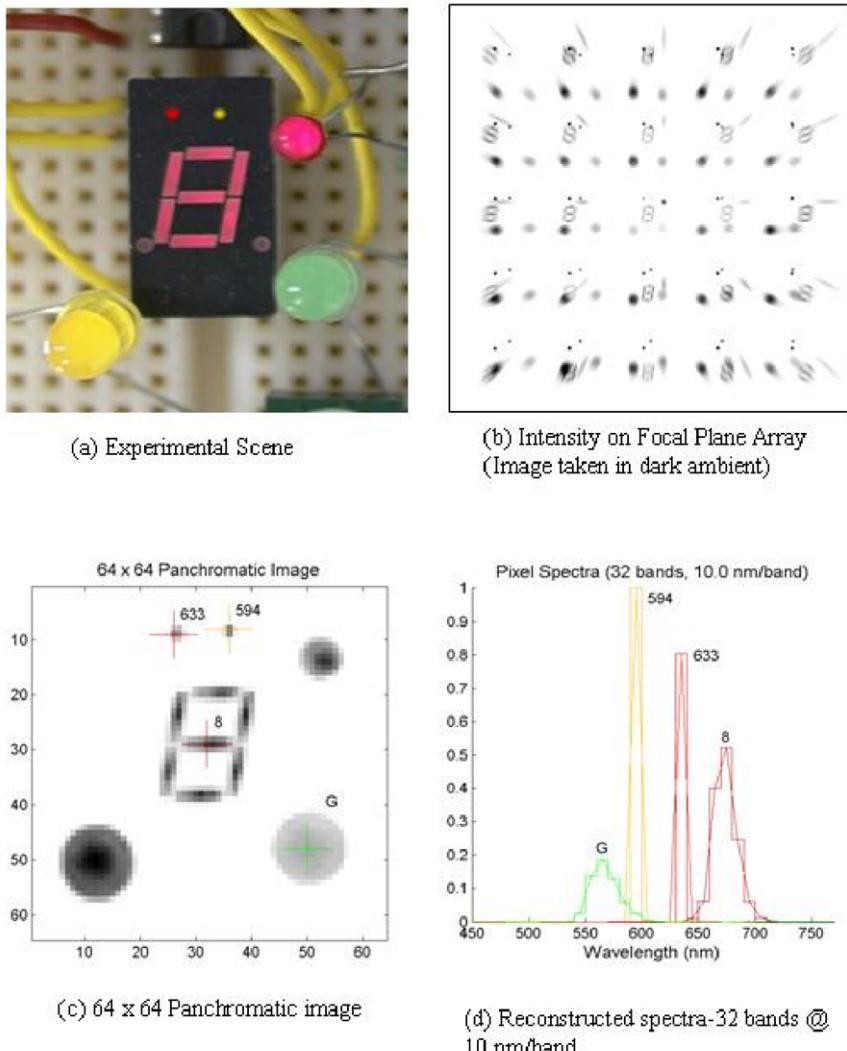


Fig. 14.17 Computed Tomographic Imaging Spectrometer (CTIS). The observed scene is in 2a, consisting of 3-color LEDs, a HeNe laser spot and an 8-segment indicator. 2b shows the resulting image from the monochrome focal plane. There is a zero order image, 2c, which can be used for focusing as well as providing an initial starting point for the reconstruction of the image cube. Recovered spectra are shown in 2d.

14.6 Conclusions

A variety of wind-driven Tumbleweed concepts have been proposed, studied and tested in extreme terrestrial environments as potential scouts for various mission

scenarios on Mars. Multiple Tumbleweed rovers could be outfitted to autonomously survey large areas of Mars for *in situ* resources at relatively low cost. Similar missions could also be configured for performing basic science. Communications and navigation would be dependent upon the assets in orbit about Mars at the time of deployment.

Tumbleweed rovers could also be networked together to provide additional communication and navigational support. Since the Tumbleweed rovers are significantly lower mass and compactable than traditional wheeled robotic rovers, many more of the Tumbleweeds can be deployed on the Mars surface during a single mission. A group of Tumbleweed rovers could survey a particular region of Mars with each Tumbleweed having a unique sensor or long-range communications capabilities. When something interesting is detected by a particular Tumbleweed, it would communicate its findings to the others, activating a swarm intelligence-based algorithm that would direct the others to proceed to the same general area and conduct additional sensing with their unique instruments. Such networking and swarming behavior is currently being studied for robotic systems by several groups (Bae et al., 2005a, b; Baxter et al., 2006; Clark et al., 2003; Hashimoto et al., 2008). An added benefit of a swarm of multiple Tumbleweeds is that a stuck rover would be to act as a fixed facility to gather temporal data while other Tumbleweeds proceed.

Low-mass, highly mobile autonomous vehicles capable of making survey measurements will fill the current void between orbital reconnaissance and landed rovers with limited range. Thus, Tumbleweeds are an attractive option for performing surveys of potential *in situ* resources available on Mars.

References

- Antol, J.: A New Vehicle for Planetary Surface Exploration: The Mars Tumbleweed. In: 1st Space Exploration Conference: Continuing the Voyage of Discovery, Orlando, Florida, AIAA-2005-2520 (2005)
- Antol, J., Calhoun, P., Flick, J., Hajos, G.A., Kolacinski, R., Minton, D., Owens, R., Parker, J.: Low Cost Mars Surface Exploration: The Mars Tumbleweed, August 2003, NASA/TM-2003-212411 (2003)
- Antol, J., Calhoun, P.C., Flick, J.J., Hajos, G.A., Keyes, J.P., Stillwagen, F.H., Krizan, S.A., Strickland, C.V., Owens, R., Wisniewski, M.: Mars Tumblweed: FY 2003 Conceptual Design Assessment. NASA Langley Research Center (2005)
- Antol, J., Chattin, R.L., Copeland, B.M., Krizan, S.A.: The NASA Langley Mars Tumbleweed Rover Prototype. In: 44th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, January 9-12, AIAA-2006-0064 (2006a)
- Antol, J., Harris, S.B., Hajos, G.A., Strickland, C.V.: Wind Tunnel Tests of Evolved Mars Tumbleweed Concepts. In: 44th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, January 9-12, AIAA-2006-0069 (2006b)
- Apostolopoulos, D., Wagner, M.D., Heys, S., Teza, J.: Results of the Inflatable Robotic Rover Testbed. Carnegie Mellon University (2003)
- Bae, J., Lee, S.: IEEE, Active sensing based mobile robot exploration. Monterey, CA, Jul 24-28, pp. 934–939 (2005a)

- Bae, J., Lee, S.: IEEE, Exploration algorithm for multiple robots. Seattle, WA, Jul 17-20, pp. 895–900 (2005b)
- Baker, V.R., Gulick, V.C., Kargel, J.S.: Water Resources and Hydrogeology of Mars. In: Lewis, J.S., Matthews, M.S., Guerrieri, M.L. (eds.) *Resources of Near-Earth Space*, pp. 765–797. The University of Arizona Press, Tucson (1993)
- Baxter, J.L., Burke, E.K., Garibaldi, J.M., Norman, M.: Multi-robot search and rescue: A potential field based approach, Dec. 11–14, 2006, pp. 9–16. Palmerston, New Zealand (2006)
- Bearman, G.H., Johnson, W.R., Fink, W., Wilson, D.W.: An Intelligently Reconfigurable Snapshot Imaging Spectrometer for Planetary Exploration. In: 38th Lunar and Planetary Science Conference, Houston, Texas, Abstract #1103 (2007)
- Behar, A., Carsey, F., Matthews, J., Jones, J.: NASA/JPL Tumbleweed Polar Rover. In: IEEE Aerospace Conference, Big Sky, Montana, p. 1003 (2004)
- Bibring, J.P., Langevin, Y., Mustard, J.F., Poulet, F., Arvidson, R., Gendrin, A., Gondet, B., Mangold, N., Pinet, P., Forget, F., team O: Global mineralogical and aqueous mars history derived from OMEGA/Mars express data. *Science* 312, 400–404 (2006)
- Blake, D.F., Sarrazin, P., Bish, D.L., Chipera, S.J., Vaniman, D.T., Ming, D., Morris, D., Yen, A.: CHEMIN: A Definitive Mieralogy Instrument on the Mars Science Laboratory (MSL 2009) Rover. In: 7th International Conference on Mars, Pasadena, California, July 9–13 (2007) Abstract #3220
- Blamont, J.: Personal Communication, Pasadena (1977)
- Bohlke, J.K., Erickson, G.E., Revesz, K.: Stable isotope evidence for an atmospheric origin of desert nitrate deposits in northern Chile and southern California, USA. *Chem. Geol.* 136, 135–152 (1997)
- Boll, J., van Rijn, R.P.G., Weiler, K.W., Ewen, J.A., Dalipathy, J., Herbert, S.J., Steenhuis, T.S.: Using ground-penetrating radar to detect layers in a sandy field soil. *Geoderma* 70, 117–132 (1996)
- Boynton, W.V., Ming, D.W., Sutter, B., Arvidson, R.E., Hoffman, J., Niles, P.B., Smith, P.: Evidence for Calcium Carbonate at the Phoenix Landing Site. In: 40th Lunar and Planetary Science Conference, Houston, TX (2009) Abstract #2434
- Brandt, O., Langley, K., Kohler, J., Hararan, S.E.: Detection of buried ice and sediment layers in permafrost using multi-frequency Ground Penetrating Radar: A case examination on Svalbard. *Remote Sensing of Environment* 111, 212–227 (2007)
- Capizzi, P., Cosentino, P.L.: GPR multi-component data analysis. *Near Surface Geophysics* 6, 87–95 (2008)
- Carsey, F., Boston, P.J., Rothschild, L.J., Coleman, M., Jones, J.A., Behar, A.E., Antol, J., Hajos, G.A., Rudisill, M., Parker, J.R., Kelliher, W.C., Carlberg, I.A.: Tumbleweed: Wind Driven Sampling on the Surface of Mars. *International Journal of Astrobiology*, 85–86 (2004)
- Clark, C.M., Rock, S.M., Latombe, J.C.: Motion planning for multiple mobile robots using dynamic networks. In: Proceedings of 2003 IEEE International Conference on Robotics and Automation, vol. 1–3, pp. 4222–4227 (2003)
- Cobos, D., Campbell, G., Campbell, C.: Thermal and Electrical Conductivity Probe (TECP) for Phoenix 2007 Scout. *Geophysical Research Abstracts* 6 (2004)
- Conyers, L.B., Goodman, D.: *Ground Penetrating Radar: A Primer for the Archaeologist*. AltaMira Press (1997)
- Doolittle, J.A.: Characterizing soil map units with the ground-penetrating radar. *Soil Survey Horizons* 23, 3–10 (1982)

- Duke, S.K.: Calibration of ground penetrating radar and calculation of attenuation and dielectric permittivity versus depth. Dept of Geophysics. Colorado School of Mines, Golden, Colorado, 236 (1990)
- Elam, W.T., Kelliher, W.C., Carlberg, I.A.: A New X-ray Fluorescence Spectrometer for Planetary Exploration. In: 39th Lunar and Planetary Science Conference, Houston, TX, Abstract #1824 (2008)
- Erickson, G.E.: Geology and origin of the Chilean nitrate deposits. U.S. Geological Survey Professional Paper P 1188 (1981)
- Erickson, G.E.: The Chilean Nitrate Deposits. *Am Scientist* 71, 366–374 (1983)
- Ezersky, M., Bruner, I., Keydar, S., Trachtman, P., Rybakov, M.: Integrated study of the sinkhole development site on the Western shores of the Dead Sea using geophysical methods. *Near Surface Geophysics* 4, 335–343 (2006)
- Feldman, W.C., Mellon, M.T., Gasnault, O., Maurice, S.,Prettyman, T.H.: Volatiles on Mars: Scientific Results from the Mars Odyssey Neutron Spectrometer. In: Bell, J.F. (ed.) *The Martian Surface*, vol. 9, p. 636. Cambridge University Press, Cambridge (2008)
- Flick, J.J., Toniolo, M.D.: Preliminary Dynamic Feasibility and Analysis of a Spherical, Wind-Driven (Tumbleweed). In: *Martian Rover 43rd AIAA Aerospace Sciences Meeting and Exhibit*, Reno, Nevada, AIAA-2005-0250 (2005)
- Gendrin, A., Mangold, N., Bibring, J.P., Langevin, Y., Gondet, B., Poulet, F., Bonello, G., Quantin, C., Mustard, J., Arvidson, R., LeMouelic, S.: Suffates in martian layered terrains: the OMEGA/Mars Express view. *Science* 307, 1587–1591 (2005)
- Golombek, M., Rapp, D.: Size-frequency distributions of rocks on Mars and Earth analog sites: Implications for future landed missions. *J. Geophys. Res-Planets* 102, 4117–4129 (1997)
- Grant, J.A., Schutz, A.E., Campbell, B.A.: Ground-penetrating radar as a tool for probing the shallow subsurface of Mars. *J. Geophys Res-Planets* 108 (2003); art. no.-8024
- Haberle, R.M., Hollingsworth, J.L., Colaprete, A., Bridger, A.F.C., McKay, C.P., Murphy, J.R., Schaeffer, J., Freedman, R.: The NASA /Ames Mars General Circulation Model: Model Improvements and Comparison with Observations. In: *Mars Atmosphere Modeling and Observations*, Granada, Spain, January 13-15 (2003)
- Hanrahan, H., Minton, D., DeJarnette, F., Camelier, M.: Mars Tumbleweed: A New Way to Explore Mars. In: *Planetary Probe Atmospheric Entry and Descent Trajectory Analysis and Science Workshop*, Lisbon, Portugal, October 6-9 (2003)
- Hashimoto, H., Aso, S., Yokota, S., Sasaki, A., Ohyama, Y., Kobayashi, H.: Stability of Swarm Robot Based on Local Forces of Local Swarms, Chofu, JAPAN, Aug 20-22, pp. 1201–1204. IEEE, Los Alamitos (2008)
- Hecht, M.H., Kounaves, S.P., Quinn, R.C., West, S.J., Young, S.M., Clark, B.C., DeFlores, L.P., Kapit, J.A., Gospodinova, K., Smith, P.H.: Discovery of Perchlorate at the Phoenix Landing Site. *Eos Transactions AGU* 89 (Fall Meeting Suppl.) (2008) Abstract U14A-04
- Hoeg, T., Southard, L., Boxerbaum, A., Reis, L., Antol, J., Heldmann, J., Quinn, R.: Tumbleweed Rover Science Mission to Dao Vallis. In: *44th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, Nevada, January 9-12 (2006) AIAA-2006-0070
- Huffman III, A.C.: Characterization of three-dimensional geological heterogeneities using ground penetrating radar. Department of Geophysics. Colorado School of Mines, Golden, Colorado, 189 (1992)
- Jakosky, B.M., Zent, A.P.: Water on Mars: Its History and Availability as a Resource. In: Lewis, J.S., Matthews, M.S., Guerrieri, M.L. (eds.) *Resources of Near-Earth Space*, pp. 737–763. The University of Arizona Press, Tucson (1993)

- Janes, D.M.: The Mars Ball: A Prototype Martian Rover. *The Case for Mars III, Part II* 75, 569–574 (1989)
- Johnson, W.R., Wilson, D.W., Fink, W., Humayun, M., Bearman, G.: Snapshot hyperspectral imaging in ophthalmology. *Journal of Biomedical Optics* 12 (2007)
- Jones, J., Yavrouian, A.: *JPL New Technology Report* (1997)
- Jones, J.A.: Inflatable Robotics for Planetary Applications. In: *6th International Symposium on Artificial Intelligence, Robotics and Automation in Space, I-SAIRAS*, Montreal, Canada (2001)
- Jones, J.A.: Personal Communication, JPL, Pasadena, CA (2009)
- Jones, J.A., Saunders, S., Blamont, J., Yavrouian, A.: Balloons for controlled roving/landing on Mars. *Acta Astronaut* 45, 293–300 (1999)
- Justus, C.G., Duvall, A., Keller, V.W.: Validation of Mars Global Reference Atmospheric Model (Mars-GRAM 2001) and planned new features. *Middle and Upper Atmospheres, Active Experiments, and Dusty Plasmas* 38, 2633–2638 (2006)
- Kathage, A., Niessen, J., White, G., Bell, N.: Fast Inspection of Railway Ballast By Means of Impulse GPR Equipped with Horn Antennas. In: *Railway Engineering 2005*, Edinburgh, Scotland, June 29–30, 2005, vol. 10(9) (2005),
<http://www.ndt.net/article/v10n09/kathage/kathage.htm>
- Kizito, F., Campbell, C.S., Campbell, G.S., Cobos, D.R., Teare, B.L., Carter, B., Hopmans, J.W.: Frequency, electrical conductivity and temperature analysis of a low-cost capacitance soil moisture sensor. *Journal of Hydrology* 352, 367–378 (2008)
- Kolecki, J.C., Landis, G.: Electrostatic Charging of the Mars Pathfinder Rover and Electrical Charge Phenomena on Mars. In: *37th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, Nevada (1999) AIAA-99-0377
- Kolecki, J.C., Landis, G.P.: Electrical Discharge on The Martian Surface. NASA Lewis Research Center, Cleveland, Ohio (1996)
- Kounaves, S.P., Catling, D., Clark, B.C., DeFlores, L.P., Gospodinova, K., Hecht, M.H., Kapit, J.A., Ming, D.W., Quinn, R.C.: Aqueous Carbonate Chemistry of the Martian Soil at the Phoenix Landing Site. In: *40th Lunar and Planetary Science Conference*, Houston, TX, Abstract #2489 (2009)
- Krasnopolsky, V.A., Maillard, J.P., Owen, T.C.: Detection of methane in the martian atmosphere: evidence for life? *Icarus* 172, 537–547 (2004)
- Kung, K.J.S., Lu, Z.B.: Using Ground-Penetrating Radar To Detect Layers Of Discontinuous Dielectric-Constant. *Soil Science Society of America Journal* 57, 335–340 (1993)
- Lambot, S., Rhebergen, J., van den Bosch, I., Slob, E.C., Vanclooster, M.: Measuring the soil water content profile of a sandy soil with an off-ground monostatic ground penetrating radar. *Vadose Zone Journal* 3, 1063–1071 (2004)
- Langevin, Y., Poulet, F., Bibring, J.P., Gondet, B.: Sulfates in the north polar region of Mars detected by OMEGA/Mars express. *Science* 307, 1584–1586 (2005)
- Lindsley, D.H. (ed.): Oxide minerals: petrologic and magnetic significance. *Minerological Society of America* (1991)
- Livingston, G.P., Hutchinson, G.L.: Enclosure-based measurement of trace gas exchange: applications and sources of error. In: *Biogenic Trace Gases: Measuring Emissions from Soil and Water*, Oxford, Great Britain, pp. 14–51 (1995)
- Marshall, J., Koppel, L., Bratton, C., Metzger, E., Hecht, M.: In: *Situ Identification of Mineral Resources with an X-ray Optical “Hand-Lens” Instrument*. *ISRU II Technical Interchange Meeting*, Houston, TX, November 18–19, Abstract Number 9028 (1997)
<http://www.lpi.usra.edu/meetings/ISRU-II-97/ISRU-II-97.3rd.html>

- Martin, J.P., Marshall, J.R., Mason, L.W., Scheld, D.: In-Situ Geologic Analyzer for Lunar and Martian Surfaces. In: Earth & Space 2008: Engineering, Science, Construction, and Operations in Challenging Environments, Long Beach, California, March 3-5 (2008) doi:10.1061/40988(323)100
- McKay, C.P., Meyer, T.R., Boston, P.J., Nelson, M., MacCallum, T., Gwynne, O.: Water Resources and Hydrogeology of Mars. In: Lewis, J.S., Matthews, M.S., Guerrieri, M.L. (eds.) Resources of Near-Earth Space, pp. 819–843. The University of Arizona Press, Tucson (1993)
- Moore, L.J., Jol, H.M., Kruse, S., Vanderburgh, S., Kaminsky, G.M.: Annual layers revealed by GPR in the subsurface of a prograding coastal barrier, southwest Washington, USA. *Journal of Sedimentary Research* 74, 690–696 (2004)
- Mumma, M.J., Villanueva, G.L., Novak, R.E., Hewagama, T., Bonev, B.P., DiSanti, M.A., Mandell, A.M., Smith, M.D.: Strong Release of Methane on Mars in Northern Summer 2003. *Science* 323, 1041–1045 (2009)
- Mustard, J.F., Murchie, S.L., Pelkey, S.M., Ehlmann, B.L., Milliken, R.E., Grant, J.A., Bring, J.P., Poulet, F., Bishop, J., Dobrea, E.N., Roach, L., Seelos, F., Arvidson, R.E., Wiseman, S., Green, R., Hash, C., Humm, D., Malaret, E., McGovern, J.A., Seelos, K., Clancy, T., Clark, R., Des Marais, D., Izenberg, N., Knudson, A., Langevin, Y., Martin, T., McGuire, P., Morris, R., Robinson, M., Roush, T., Smith, M., Swayze, G., Taylor, H., Titus, T., Wolff, M.: Hydrated silicate minerals on mars observed by the Mars reconnaissance orbiter CRISM instrument. *Nature* 454, 305–309 (2008)
- Nettwerk, Underground mapping, pipeline location technology, and condition assessment. Univ. Birmingham Infrastructure Engineering & Management Research Centre, Civil Engineering, School of Engineering, Edgbaston, Birmingham, UK (2002)
- Olhoeft, G.R.: The electrical properties of permafrost. Department of Physics. University of Toronto, Toronto, Canada. 172 (1975)
- Olhoeft, G.R.: Effects of water on the electrical properties of planetary regoliths. In: Colloquium on Water in Planetary Regoliths, Hanover, NH, Dartmouth College, pp. 139–142 (1976)
- Olhoeft, G.R.: Electrical properties of natural clay permafrost. *Canadian Journal of Earth Sciences = Journal Canadien des Sciences de la Terre* 14, 16–24 (1977)
- Olhoeft, G.R.: Impulse radar studies of near surface geological structure. In: Lunar and Planetary Science Conference X, Houston, Texas, pp. 943–945 (1979)
- Olhoeft, G.R.: Direct detection of hydrocarbon and organic chemicals with ground penetrating radar and complex resistivity. In: NWWA/API Conference on Petroleum Hydrocarbons and organic chemicals in ground water – prevention, detection and restoration, Dublin, OH, November 12–14, 1986, pp. 284–305 (1986)
- Olhoeft, G.R.: Magnetic and electrical properties of martian particles. In: Sand and Dust on Mars, NASA Conf. Publ. 10074, Greenbelt, Maryland, pp. 44–46 (1991)
- Olhoeft, G.R.: Ground penetrating radar on Mars. In: Proc. of GPR 1998, 7th Int'l. Conf. On Ground Penetrating Radar, May 27–30, 1998, pp. 387–392. The Univ. of Kansas, Lawrence, KS, USA (1998)
- Olhoeft, G.R.: Maximizing the information return from ground penetrating radar. *Journal Of Applied Geophysics* 43, 175–187 (2000)
- Palli, A., Kohler, J.C., Isaksson, E., Moore, J.C., Pinglot, J.F., Pohjola, V.A., Samuelsson, H.: Spatial and temporal variability of snow accumulation using ground-penetrating radar and ice cores on a Svalbard glacier. *J. Glaciol.* 48, 417–424 (2002)
- Physical Sciences Incorporation, PSI: Products, vol. 2009 (2009)

- Poulet, F., Bibring, J.P., Mustard, J.F., Gendrin, A., Mangold, N., Langevin, Y., Arvidson, R.E., Gondet, B., Gomez, C., Omega, T.: Phyllosilicates on Mars and implications for early martian climate. *Nature* 438, 623–627 (2005)
- Rose, S.E., Moody, C.B., James, D.L., Barhorst, A.A.: Drag measurement and dynamic simulation of Martian wind-driven sensor platform concepts. *Journal of Fluids and Structures* 22, 21–43 (2006)
- Schaber, G.G., McCauley, J.F., Breed, C.S., Olhoeft, G.R.: Shuttle imaging radar: physical controls on signal penetration and subsurface scattering in the eastern Sahara. *Institute of Electrical and Electronic Engineers Trans Geoscience and Remote Sensing GE-24*, 603–623 (1986)
- Schenk, C.J., Gautier, D.L., Olhoeft, G.R., Lucius, J.E.: Internal structure of an aeolian dune using ground penetrating radar. In: Lancaster, N., Rye, K. (eds.) *Aeolian Sediments: Ancient and Modern*, vol. 16, pp. 61–69. Blackwell Publishing, Malden (1993)
- Schmitt, H.H.: *Potential Resources of Mars* (2004)
- Stanley, R.J., Ho, D.K.C., Gader, P.D., Wilson, J.N., Devaney, J.B.: Advances in EMI and GPR algorithms in discrimination mode processing for handheld landmine detectors. In: *Proceedings of the SPIE*, vol. 5415, pp. 874–882 (2004)
- Stoker, C.R., Gooding, J.L., Roush, T., Banin, A., Burt, D., Clark, B.C., Flynn, G.J., Gwynne, O.: *The Physical and Chemical Properties and Resource Potential of Martian Surface Soils*. In: Lewis, J.S., Matthews, M.S., Guerrieri, M.L. (eds.) *Resources of Near-Earth Space*, pp. 659–707. The University of Arizona Press, Tucson (1993)
- Sutter, B., Lauer, H.V., Golden, D.C., Boynton, W.V., Morris, R.V., Niles, P.B., Ming, D.W.: Thermal and Evolved Gas Behavior of Calcite under Mars Phoenix TEGA Operating Conditions. In: *40th Lunar and Planetary Science Conference*, Houston, TX (2009) Abstract #1843
- Tarsitano, C.G., Webster, C.R.: Multilaser Herriott cell for planetary tunable laser spectrometers. *Applied Optics* 46, 6923–6935 (2007)
- Taylor, G.J.: Human Exploration for Resources on Mars. *Science and the Human Exploration of Mars*, Greenbelt, Maryland (2001),
<http://www.lpi.usra.edu/publications/reports/CB-1089/taylor.pdf>
- Wang, H., Yang, B., Jones, J.: Mobility Analysis of an Inflated Tumbleweed Ball under Wind Loads. In: *3rd AIAA Gossamer Spacecraft Forum*, Denver, Colorado, April 22–25, AIAA-2002-1556 (2002)
- West, L.J., Truss, S.W.: Borehole time domain reflectometry in layered sandstone: Impact of measurement technique on vadose zone process identification. *Journal of Hydrology* 319, 143–162 (2006)
- Wilson, J.L., Mazzoleni, A.P., DeJarnette, F.R., Antol, J., Hajos, G.A., Strickland, C.V.: Design, analysis, and testing of Mars Tumbleweed rover concepts. *Journal of Spacecraft and Rockets* 45, 370–382 (2008)
- Wright, D.L., Olhoeft, G.R., Grover, T.P.: Velocity, attenuation and dispersion electromagnetic tomography in fractured rock. In: *U S Geological Survey Toxic Substances Hydrology Program Review*, Colorado Springs, CO, September 20-23, 1993, pp. 111–118 (1996) USGS WRI Report 94-4015
- Yang, C.H., Tong, L.T., Yu, C.Y.: Integrating GPR and RIP methods for water surface detection of geological structures. *Terrestrial Atmospheric and Oceanic Sciences* 17, 391–404 (2006)

Chapter 15

Drilling and Excavation for Construction and In-Situ Resource Utilization

Kris Zacny¹ and Yoseph Bar-Cohen²

¹ Honeybee Robotics Spacecraft Mechanisms Corporation, New York

² Jet Propulsion Laboratory/Caltech, Pasadena, USA

15.1 Introduction

As the exploration of the solar system by the international space agencies (NASA, ESA, JAXA and others) continues to expand there are increasing considerations to include humans as part of the future missions. These missions may consist of short term site visits as was done during the landing on the Moon or for extended stay as part of permanent human habitation. At the present time, the planetary body that is considered for a permanent habitation is the Earth's Moon. However, human presence on the Moon is also sought as part of a broader plan where it would serve as a stepping stone to colonizing other bodies in our solar system, the most intriguing one being Mars. Such a presence is quite challenging and will require the establishment of the critical infrastructure to support the essential capabilities that are often taken for granted on Earth. The tasks enabling human settlement will be quite complex due to the harsh conditions on Mars such as low temperature, atmospheric pressure, and the radiation environment. In addition, long communication delays will significantly impact the ability for any system to be teleoperated from Earth. The cost and associated challenges will require cooperation of the international space agencies and this may follow some of the approaches and procedures that were used in the development and operation of the International Space Station (ISS).

To help make planetary habitation more practical it would also be important to insure commercial benefits from the related missions to provide an impetus to a permanent presence on extraterrestrial bodies. The use of robots helps minimize risk to humans by performing tasks that are too risky, demanding or difficult for humans. New robotic capabilities that may use telepresence will be essential, and these may include scenarios where robots will act as human surrogates and perform tasks in extreme environmental conditions (Bar-Cohen and Hanson 2009). Performing some tasks on Mars will also be more difficult than on Earth because of the fact that the gravity is three times lower. This means that to exert the same excavation or drilling force a system has to be three times as massive as on Earth. The development of the required tools and capabilities will be an evolutionary process where challenges will be addressed and solutions will be implemented as the confidence in the in-situ resource utilization (ISRU) grows and the benefits

expand. The topic of in-situ resource utilization is quite broad and covers all the aspects of assuring sustainable human life including mining and processing of local resources. The focus of this chapter is on mining and excavation technologies that will enable human exploration and long term presence on Mars.

15.2 How Mars Is Different to Earth

Human presence on Mars, especially over extended time periods, necessitates a thorough understanding of the in-situ environment. Besides protecting the astronauts and making sure they are able to survive in the extreme conditions it is also necessary to assure the operability of the required equipment and tools. This is, in a way, analogous to operation in extreme environments on Earth. Trucks driving in the northern parts of Alaska are modified to withstand the harsh cold environment where the temperature reaches -50°C and lower. In the same way, trucks traversing the hot outback of Australia are modified to deal with extreme heat reaching temperatures of 60°C and more. Equipment that will be used on the surface of Mars will also have to be modified or designed having Mars in mind. The issues that will have to be taken into account include low temperatures, large thermal fluctuations of up to 100°C in just a few hours, large vertical thermal gradients (the temperature on the ground is much higher than the temperature one meter above the ground), low gravity, low atmospheric pressure and a high radiation environment, which can damage electronic systems and living cells (humans). The following subsections will address each of the above issues in more detail.

15.2.1 The Effect of Temperature

Mars is a very cold planet, with surface temperatures reaching as low as -133°C at the poles during the winter (Clancy et al. 2000). This necessitates keeping many of the critical electronic subsystems warm. The electronic systems that are currently developed for use on Mars are designed to operate above -55°C and to survive exposure to temperatures above -65°C. If temperatures below these values are encountered, the critical subsystems have to be kept warm by placing them inside a warm box. The heat is normally generated by radioisotope heater units (RHU). An RHU is a small device that is about 3.2 cm long by 2.6 cm in diameter and weighs about 40 grams. It provides approximately one watt of thermal power (heat), derived from the decay of a few grams of a radioactive element such as plutonium 238 (for details see Chap. 1).

Not all systems can be placed inside a warm box or kept warm via other means. For example, keeping warm an excavator bucket that is exposed to the outside cold environment would require too much heat. For this reason, the related structural components have to be designed to operate efficiently even at temperature of -100 °C or lower. The most critical issues that need to be addressed when designing for these extreme temperatures are lubrication, thermal expansion, and material embrittlement.

Viscosity (resistance to flow) of lubricants increases as temperatures is reduced (i.e. lubricants become highly viscous at low temperatures). For that reason, the

use of mining equipment in the Arctic is suspended once the temperature drops below -50°C or so to prevent engines and other system components with moving parts from being damaged. Possible solutions to the lubrication problem can be the use of a dry lubricant or low friction material such as PTFE, or anti-galling materials such as beryllium copper. In many aspects, dry lubricant is inferior to wet lubricant but for certain applications it may be sufficiently effective.

The issue related to thermal expansion comes into play when two materials with different coefficients of thermal expansion are in contact with each other. As the temperature drops the two materials will shrink but at different rates. A material with a large coefficient will shrink more than a material with a lower coefficient of thermal expansion. Thus, at the interface between them large lateral stresses will be developed. These stresses, if sufficiently high, will cause material failure such as fracture or delamination. Thus, when a system has tight clearances, materials with similar coefficients of thermal expansion should be used. This is especially important on Mars, where thermal fluctuations of over 100 °C can occur in a period of six hours (Martin et al. 2003).

Some materials such as low carbon steels change their mechanical characteristics with temperature. At higher temperatures, they are ductile (can be bent without breaking) while at low temperature, below the Ductile to Brittle Transition Temperature (DBT), they become brittle. For example, the hull of the steamship Titanic was made of steel with the DBT of +32°C. Thus at the temperatures below +32°C the steel making up the hull of the Titanic was brittle. When an iceberg hit its hull and created a small crack, the crack propagated very easily along the rest of the hull. If the steel had been ductile, the crack would have been arrested (would not propagate) because the steel would simply have deformed plastically. Fortunately, most modern steels become brittle only at temperatures below -27°C (Bruzel 2000). For Mars applications, because of the extremely low temperatures, material selection becomes even more critical. Some materials, such as Aluminum or Titanium (and many stainless steels) do not show the ductile to brittle transition and are therefore better suited for lower temperatures than steels, even though they are not as strong as steel.

15.2.2 The Effect of Low Atmospheric Pressure

In addition to the issue of low temperatures on Mars there is also a low atmospheric pressure, ranging from 0.1 kPa to 1.5 kPa (Smith and Zuber, 1998) (for other details see Chap. 2). These particular pressure levels are very important because the triple point of water is located at 0.63 kPa and 0°C. The triple point of water refers to temperature and pressure where water can be simultaneously in all three forms: ice, liquid, and gas. When the pressure is below the triple point, if temperature of an ice sample reaches 0°C, the ice will sublime directly to vapor. However, if the atmospheric pressure is above the triple point, the same ice will melt first upon reaching 0°. Only further increase in temperature will cause it to vaporize. At the pressure of 1 atmosphere (i.e. atmospheric pressure on Earth) this temperature is 100°C. The particular pressure regime on Mars allows for some innovative means of water mining and extraction, the details of which will be described later in this chapter.

The low atmospheric pressure also has an adverse effect on heat dissipation. On Earth, heat is transferred or dissipated mainly via gas convection. On Mars, however, gas convection is very low and gas conduction (heat transfer due to transfer of gas molecules) and to a lesser extent radiation are the main heat transfer methods (Presley and Craddock, 2006). Heat dissipation via gas conduction on Mars is at least two orders of magnitude lower than heat dissipation via gas convection on Earth, and in turn cooling systems for electric motors or Radioisotope Thermoelectric Generators (RTG) will be required.

Lower atmospheric pressure can also lead to a reduction in the coefficient of sliding friction due to desorption of adsorbed gases, such as water vapor (Zacny and Cooper 2007). This is in fact beneficial because lower friction means less wear and less heat generation. However, when the low pressure is combined with high temperature (e.g. when materials rub against each other and become hot), the activation energy for the decomposition of surface oxides present on surfaces of rubbing/sliding materials will be reduced. This means that protective surface oxide layers will be very unstable and will inevitably be removed, exposing the material surface to the environment. The result is an increase in friction coefficient, and in turn resistance to sliding. In a practical sense, higher friction means that higher forces need to be applied to move materials past each other (as in pressure cylinders, where seals are sliding past the cylinder).

The thin Martian atmosphere also requires spacecraft landing approaches which are different from those used on the Moon (the Moon lacks an atmosphere) and which are similar to those used on Earth. The entry, descent and landing of the spacecraft on Mars (EDL) can use atmospheric air drag (by using parachutes or even the large surface area of the spacecraft during ballistic entry) to slow down before deploying other landing system such as airbags or retro rockets.

15.2.3 Large Distance from the Earth and Communication Delay

The distance from Earth to Mars varies considerably, from 57,936,384 to 402,336,000 kilometers. This has not only direct implications to the high cost of placing a payload on the surface of Mars, but also affects communications with the spacecraft on Mars. The communication can only occur at the speed of light, which means, that when Mars is furthest away, the transmission delay can be up to 20 minutes each way. The round-trip delay of 40 minutes (assuming that information is send directly to an instrument on the surface of Mars and not relayed via an orbiter around Mars) is prohibitively long for any form of teleoperation. Teleoperation is possible for closer bodies such as Moon. In fact, the Soviet rovers, Lunokhod 1 and 2 were remotely controlled on the Moon from Earth and successfully traversed tens of kilometers. This will not really be an option for Mars, although Mars Exploration Rovers (MER) were teleoperated to a certain extent. Once the command was issued given the destination, the rovers would use their on board software to decide which way to get there. Teleoperation could be feasible for general instructions such as these, but not in dynamic situations where rapid response is required – it's the need for rapid response that determines the usefulness of in situ autonomous capability.

The communication delays can be dealt with in one of three ways including a) direct human operation (astronaut physically sitting on an excavator or a rover); b) human teleoperation from the surface of Mars or from the orbit around Mars; and c) full autonomy (equipment makes the decision where and how to excavate and drive on the surface). Each of these alternatives of course has its advantages and disadvantages and probably the optimum method will depend on the level of autonomy that we, humans, would entrust to a machine and also on the cost of implementing each of these methods. Ultimately it is the required speed of response that influences the choice.

15.2.4 Geological Uncertainty

Mars in a way is very similar to Earth, in that it is quite difficult to predict what one can find when digging a meter or two in the ground. This unknown is called geological uncertainty. As opposed to Earth or Mars, the Moon is quite predictable. We know that the first few meters are covered by dry regolith: crushed basaltic rocks mixed with agglutinates (fused glass and rock fragments). Thus for the designers of excavators and drills, the Moon would be an easy matter. However, Mars is different. Mars has water, and clay minerals, which make the soil significantly different. In addition, having an atmospheric pressure close to the triple point complicates the matter even more. For example, a drill penetrating icy soil may temporarily melt the ice and once the drilling stops, the water will refreeze around the drill possibly trapping it inside the hole. Clay minerals and salts, such as perchlorate that was found by the Phoenix 2008, reduce the ice melting temperature. In particular, perchlorate solutions can be liquid below -75°C. This makes Mars much more interesting from the exploration stand point, but also more challenging in terms of excavation and mining. In addition when soil is fully saturated with water ice, its strength is as high as that of very strong sandstone (Mellor 1971, Zacny et al. 2007).

15.2.5 One Third the Gravity of Earth

Gravity on Mars is one third of that on Earth and it means that everything weighs three times less (one third of its earth weight). For example, a 60 kg person, weighing 600 N on Earth, would weigh only 200 N on Mars (for some consequences see Chap. 22). Alternatively, to apply the same vertical force, the mass of an object on Mars has to be three times more. This large reduction in gravitational acceleration has far reaching consequences on the choice of excavation methods.

On Earth, most excavators use their own weight to apply the down force necessary to penetrate the ground. On Mars, to generate the same digging forces, the excavator would have to be three times more massive. Bringing a large excavator onto Mars, not only will be very expensive, but may also not be feasible. For example, landing the 1000 kg Mars Science Laboratory rover created many new challenges to the EDL (Entry Descend and Landing) team. The NASA/JPL team determined that the only viable way to land such a large payload on the surface of Mars would be by using a skycrane. This is somewhat similar concept to a helicopter lowering a payload onto the ground using a winch and a set of lines, but

instead of helicopter rotor blades, the skycrane would use rocket thrusters for keeping itself hovering at the required height above the surface.

15.3 Space Heritage: Examples of Excavation Systems Deployed on Extraterrestrial Bodies

The space heritage of excavation systems that were used on extraterrestrial surfaces is quite limited. So far, there have been only a handful of scoops and drills (Bar-Cohen and Zacny, 2009) that were used on the Moon and Mars. Scoops are very useful for collecting loose soil but have an inherent drawback in that they cannot dig into hard rocks, and struggle in highly compacted soils. Drills, on the other hand, can penetrate even the hardest rocks known to humans, however, they may require additional sample collection features when acquiring samples of rocks, cores or soil. This section reviews the scoops and drills that were deployed on the surfaces of other extraterrestrial bodies.

15.3.1 Planetary Scoops

The first scoops that were ever deployed extraterrestrially were onboard the lunar reconnaissance landers, called Surveyors. Surveyors were launched to the Moon between 1966 and 1968 as precursors to the Apollo missions (Fig. 15.1). There were only three successful missions: Surveyor 1 on June 2, 1966, Surveyor 3 on

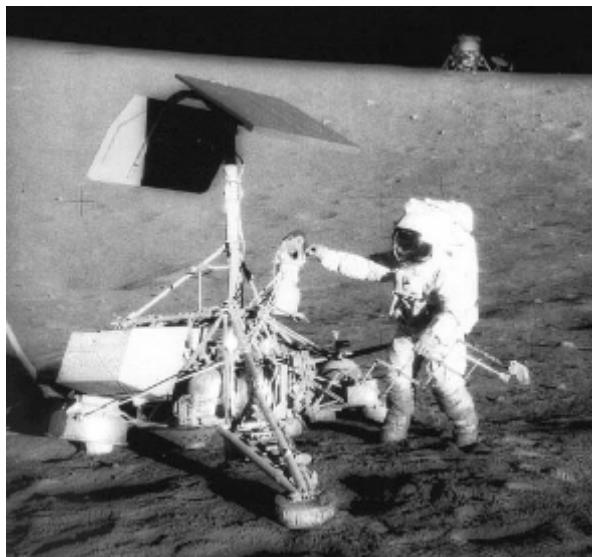


Fig. 15.1 Apollo 12 landed near the Surveyor 3 landing site. The scoop is in front of the astronaut, while the Apollo 12 Lunar Module is in the background. Credit: NASA, Kennedy Space Center.

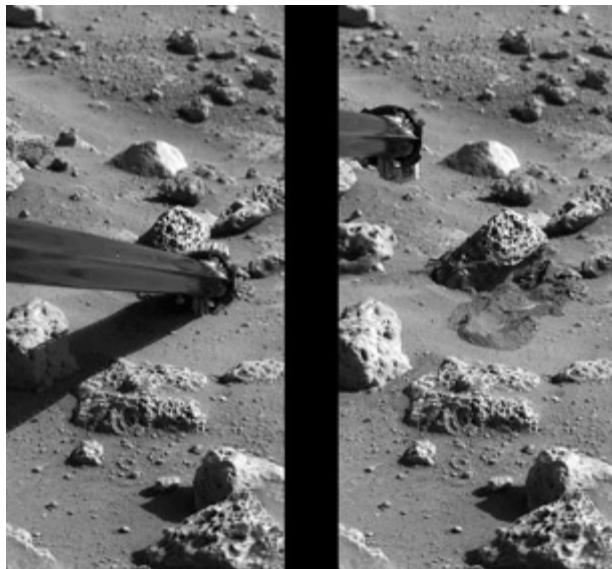


Fig. 15.2 Viking 2 landed at Utopia Planitia on September 3, 1976. Left: Viking 2 lander's arm scooping up soil; Right: the ram being retracted back with a soil sample (note the trench left behind on the surface). Courtesy NASA/JPL.

April 19, 1967, and Surveyor 7 on Jan 9, 1968. The purpose of having a scoop was to perform geotechnical experiments on the surface of the Moon with the purpose of measuring bearing strength of the top soil and also to determine the related excavation forces. Only Surveyor 3 and 7 directly measured actual digging and trenching forces. By pulling the scoop toward the spacecraft while pushing down to dig trenches, excavation forces could be measured.

In 1976, two landers used scoops for acquiring soil samples on Mars. The Viking program included two identical orbiters and two identical 600 kg landers, which landed on nearly opposite sides of Mars with a goal of acquiring soil samples to test for the presence of life indicators. Viking 1 landed in Chryse Planitia on 20 July 1976, while Viking 2 landed in Utopia Planitia on 3 September of 1976 (Fig. 15.2). Although the soil test results were ambiguous, the scoops worked flawlessly.

The last mission that included a scooping device was the Phoenix 2008 lander (Fig. 15.3), which touched down on Mars in May 2008. This mission's main goal was to search for environments suitable for microbial life on Mars, and to research the history of water. To support these two goals the lander was equipped with two main instruments: the Thermal and Evolved Gas Analyzer (TEGA) and The Microscopy, Electrochemistry, and Conductivity Analyzer (MECA). The Phoenix scoop at the end of the Lander's Robotic Arm was called the Icy Soil Acquisition Device (ISAD), because it was designed not only to scoop a loose overburden but also to sample a hard icy layer. For sampling the hard icy layer, the scoop was equipped

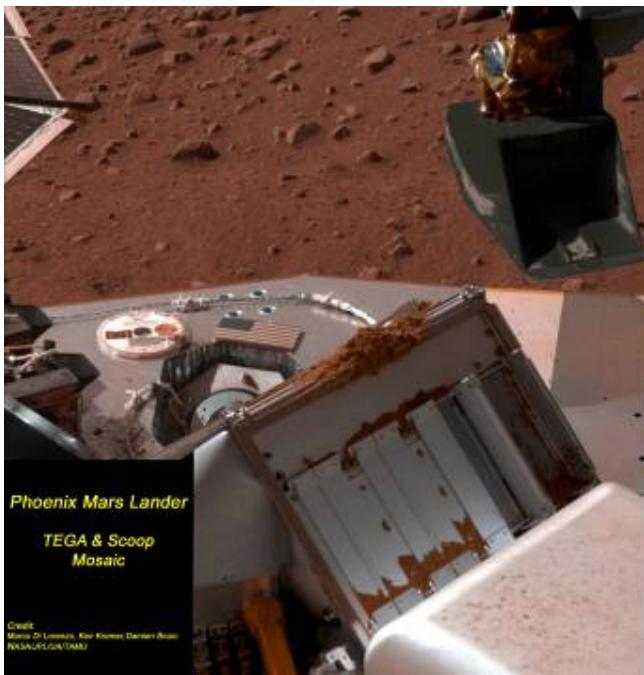


Fig. 15.3 NASA's Phoenix Mars Lander's solar panel and the Lander's Robotic Arm with a sample in the scoop. Image credit: Marco Di Lorenzo, Kenneth Kremer, Damien Bouic NASA/JPL/UA/TAMU/Aviation Week.

with a small tungsten carbide drill bit, called, the RASP (Fig. 15.4). The high speed drill bit when pushed into an icy surface threw the cuttings in all directions due to the centrifugal force. Some of these cuttings fell through a small slot and into the back of the scoop. The cuttings were then transferred to the front of the scoop (for easy delivery to scientific instruments) through an elaborate labyrinth of plates built in at the back end of the scoop. The scoop weighted less than 700 grams and had a capacity of 350 cc. The RASP required 15 Watts of power to acquire about 1cc of material in just 30 seconds. The RASP was also activated during the delivery of the sample to scientific instruments. Vibrations that were produced by the RASP helped with the flow of soil sample out of the scoop.

One of the more interesting findings during the Phoenix 2007 mission was the fact that soil tended to stick to the surface of the scoop. A thorough analysis revealed that the stickiness was caused by water-ice in the soil partially melting and refreezing onto the surface of the scoop. The solution to soil adhesion was solved by keeping the scoop in the shadow of the lander, i.e. keeping it away from the direct sunlight that was the source of heat.

Both the operation of the RASP and the scoop were very successful and both were instrumental in making a number of scientific discoveries.



Fig. 15.4 The Phoenix Icy Soil Acquisition Device with a small drill bit, called RASP, for acquiring of icy samples. Image courtesy Honeybee Robotics.

15.3.2 Planetary Drills

Only two countries deployed planetary drills on other bodies in the solar system; these are the United States and the Soviet Union. Both drills were deployed on the Moon in the 1970s.

The Apollo program used an astronaut operated rotary percussive drill, called the Apollo Lunar Surface Drill (ALSD), to acquire a continuous regolith core from up to 3 meters and also to deploy two heat flow probes to a depth of 2.4 meters (Fig. 15.5). These 400 Watt and greater power drills were some of the first battery operated drills made by Black and Decker. Only the last three missions, Apollo 15, 16 and 17, used the ALSD. The faulty auger designed for the Apollo 15 heat flow probe caused the drill to seize in place. The fault was attributed to a discontinuity in the auger flutes at the junctions between consecutive drill tubes that prevented chips from flowing further up the hole. This problem was solved in Apollo 16 and 17, and on both missions, astronauts were able to reach the target depth of 2.4 meters quite easily. In fact, penetrating the lunar regolith was easier than initially thought. It took a longer time to disconnect the drill head and to add additional drill string sections than to penetrate the regolith.

However, pulling the drill out with a continuous core was hard work. During the Apollo 15 mission, David Scott sprained his shoulder trying to pry the drill out. For this reason Apollo 16 and 17 had a jack for lifting the drill string out of the hole. Yet, even with this mechanical jack assistance, it was still hard work. The reason for these problems was the very high relative density of the lunar regolith. The soil was packed so much, that it could not be pushed into the sidewall but had to be moved out of the hole. In addition soil cohesion driven by the highly angular shape of the soil grains and also high relative density did not make the drilling any easier.



Fig. 15.5 Mock up of Apollo astronauts at NASA JSC visitor's center. The astronaut on the left holds the Apollo Lunar Surface Drill. The drill strings in front obscure part of the drill. Image courtesy: K. Zacny.

The Soviet shallow drills (tens of centimeters) were part of the Luna 16 and 20 missions while a 2 meter deep drill was part of the Luna 24 mission. All three missions were successful in acquiring and returning regolith samples back to Earth.

15.4 Operating Outposts on Mars and Other Bodies

Technologically, humans have progressed tremendously over the past few hundred years (Bar-Cohen and Zacny, 2009). Especially in the post World War II times, the progress driven by the Cold War and other conflicts led to the development of new technologies that eventually helped to make our lives relatively comfortable. The availability of life essentials has been enabled through the development of a well established and effective infrastructure. The infrastructure includes buildings, transportation, fresh water supply and sewage systems, shelters for protection against extreme weather conditions, communication systems such as the phone and internet, utilities such as gas and electricity as well as variety of foods that can be purchased in stores in our neighborhoods. All these elements and many more that are considered essentials of the infrastructure are not available on other planets and their availability will need to be established. Bringing everything from Earth will be quite expensive and may take a long time and therefore effective utilization of local resources will be highly beneficial. The Louis and Clark's "living off the land" approach will also be applicable to Mars. The initial development of capabilities allowing permanent human presence on the Moon may pave the road for establishing the foundations for exploration of other planets, such as Mars. The Earth's Moon may serve in some cases as the testbed for various technologies while reducing the associated development cost and risk to extended human operation directly to Mars. However, not all systems developed for the Moon

will be directly applicable to Mars. For example many In-Situ Resource Utilization technologies using lunar regolith will not work on Mars because of the different properties of the martian soil. The essential infrastructure components will need to be prioritized according to their criticality to the life of the astronauts.

Establishing outposts with livable conditions will require breathable air, comfortable temperature and air pressure, running water for drinking and hygiene, food processing capability, and other needs. These systems to a certain extent have been developed for and used in the International Space Station (ISS). However, the ISS is just 350 km up in the sky, and if something goes wrong, like for example if the water purification system fails completely, the astronauts can reach Earth in a matter of hour or less. On Mars, no such safety net will be available and thus any life support system has to be very robust. Some of the specific needs for future habitation of Mars may therefore include:

- Need for harvesting local resources for consumable supply (Oxygen, water, fuel including hydrogen and methane, food, etc.)
- Recycling of all the resources. This includes a centralized water and waste recycling system.
- Communication infrastructure.
- Power generation facilities that may use either local resources such as solar power or a more stable power supply from a nuclear reactor.
- Storage facilities for various resources (including liquids and gases).
- Landing sites with protective berms.
- Protection from radiation and extreme ambient conditions. These may involve underground tunnels.

Establishing such capabilities will require both tools and construction materials. The tools will be brought from Earth while construction materials in most part will be found on Mars.

15.4.1 Exploration of Mars

Habitation of other planets will require thorough characterization of the environment in order to determine potential life threatening hazards. These hazards could be as simple as minute dust particles that can pass filters and seals and enter habitats and astronauts airways. The other part of the required planetary characterization involves resource exploration. This will probably take place in a similar manner as on Earth, except that instead of a human geologist doing it manually the task will be done by a roving robot driving around and performing the required reconnaissance of the surface and subsurface. The subsurface sampling will be enabled by different drill systems. Some drills will acquire intact cores for further analysis by astronauts, while others will drill and acquire small quantities of rock powder for analysis by onboard instruments. In addition, a logging tool with a number of instruments can be lowered into a borehole to perform a better observation of the subsurface. All these activities are essential to assessing the potential availability of resources in different regions on Mars.

Fig. 15.6 The Robonaut is remotely controlled to perform physical tasks by mirroring the actions of a human. The photo was provided through the courtesy of NASA Johnson Space Center, Huston, TX.



15.4.2 Robotics and Telepresence

Humans will need the assistance of robots that will operate either as autonomous machines or under telepresence control as humans' surrogates (Bar-Cohen and Hanson, 2009). The latter capability enables projecting the human intelligence into the function of the robots by directly and remotely operating it. For over a decade, roboticists at NASA Johnson Space Center have been developing a novel space robot called Robonaut (Fig. 15.6). This robot is capable of performing various tasks at remote sites and serving as a robotic astronaut on the International Space Station. Robonaut thereby provides a relatively fast response to situations that require astronauts attention outside the environment in which human can operate. The Robonaut is designed as an anthropomorphic robot, similar in size to a suited Extra Vehicular Activity (EVA) astronaut and having robotic arms that are capable of dexterous, human-like maneuvers to ensure safety and mission success. For mobility, two and four wheel designs have been developed allowing the Robonaut to reach desired remote locations.

15.5 Types of Drilling Equipment

Drills can be ideal tools for performing geological reconnaissance by collecting viable cores for analysis or for deploying scientific instruments in boreholes (Bar-Cohen and Zacny, 2009, Zacny et al. 2008b). In addition, drills can be used in a “drill and blast” mining system. This particular method involves drilling a borehole, lowering an explosive, blasting and in turn loosening the rock or soil for easy excavation by conventional bucket equipped mining equipment. The following Sections describe the principles of drilling and give examples of various planetary drill systems that are being developed with Mars in mind.

15.5.1 Principles of Drilling

In general, there are three basic approaches to breaking rocks: mechanical, thermal (thermal-spalling, thermal-melting and vaporization), and chemical approaches (Maurer 1968). The most widely used techniques are mechanical drills powered manually, electrically, hydraulically, or pneumatically. Their sizes vary from hand-held models to giant rigs such as those used in the oil industry.

Mechanical methods are the most widely used for rock penetration. They include the three key approaches: rotary, percussive and rotary-percussive (Bar-Cohen and Zacyn 2009). In rotary drilling, the most widely used technique consists of rotating a drill bit that cuts or crushes the drilled rock. In percussive techniques, repeated impacts produce finely crushed rock directly underneath the impacted location. A rotary-percussive system combines rotation and percussion and it is for many rock types the most efficient drilling method. The impact forces are applied at the contact points of the cutting teeth and if the stress is high enough the rock material under the teeth breaks. At high impact stress levels, cracks are initiated around the crushed zone, creating rock fragments (Maurer 1968).

Generally, under mechanical loading, most rocks behave as elasto-plastic materials, exhibiting strain hardening and softening (Han et al. 2009). A typical stress-strain curve of a rock under axial stress is shown in Fig. 15.7. Subjecting a rock to a stress level that is higher than its yield stress causes a permanent deformation. If the rock behaves as a brittle material it will fracture, and if it is more plastic, it will deform. Rocks can fail in fatigue too; cyclic loading can lead to the formation of miniature cracks that grow with time until they reach critical lengths, at which point the rock breaks.

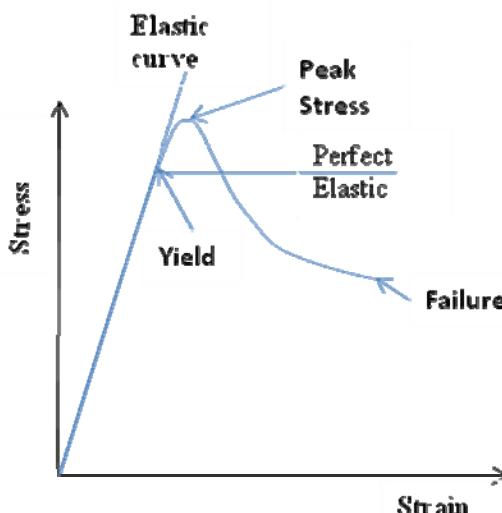


Fig. 15.7 Typical stress-strain curve of a rock under loading

Rocks can be subjected to tensile, shear and pure compressive stresses. Tension and shear stresses cause failure of the intergranular bonds, while pure compression stresses lead to crushing and breakage of grains and a shear failure. At different locations within the rock material that surrounds the drill bit, the drilling process may introduce all three types of loading. The relative levels of tension, shear, and compressive stresses are influenced not only by the loading mode (compressive impact and/or frictional shear loading), but also by the bit and borehole geometry as well as dynamic stress wave reflections that are generated by the drilled medium.

The challenges to drilling increase significantly with the depth of penetration into the subsurface. If the penetrated medium consists of soil, the wall of the formed borehole needs to be encased in order to avoid collapse. Pressure from the borehole wall grows as the drilling depth increases and the wall requires adequate attention during drilling. If the medium that is being drilled consists of ice, re-freezing of the melted water can freeze the drill in place.

15.5.2 Earth Based System

Small drills suitable for exploration purposes most often use a rotary drilling action and include both portable and fixture mountable types. Typical fixtures for small drills include those known as a drill press, a pedestal drill, a pillar drill, or a bench drill, and are mounted on a stand or bolted to the floor or workbench. Hammer drills use impact action at various frequencies to penetrate the material being drilled and are widely used to penetrate brittle materials, such as concrete. Large hammer drills, especially the pneumatically driven ones, are crude in their action and the energy delivered in each stroke is highly variable. The cheaper drills are used to smash the drilled object and they introduce significant vibrations into the surroundings that can cause collateral damage. Rotary-hammer drills are very effective in drilling hard concrete and masonry. This type of drill operates percussively onto objects and applies its weight to introduce impact forces while also rotating the fluted bit to remove the cuttings. The combined action significantly enhances the operation of the drill over the use of either rotation or hammering only. A rotary-hammering drill combining an actuator for rotary motion and a voice coil for percussion has been developed for use in the upcoming Mars Science Laboratory Mission expected to be launched in 2011.

If hard rocks are encountered, a high feed force (often referred to as Weight on Bit (WOB)) will be required. Hammering action can increase the effective weight on bit that is available from the drilling equipment. If the drilling mechanism applies a series of impacts of average force f and time of duration t , and it takes time T to re-cock the impact mechanism, where T is e.g., ten times t , the average reactive force that must be supplied by the drilling machine is $f t/T$, or only one tenth of f . In effect, therefore, the available WOB is multiplied by ten. Hammering is widely applied in drills that must be light in weight, and for which limited feed force is available (including hand-held devices), and also when the material to be penetrated is hard and brittle, viz. rocks and ice, because brittle materials can be broken and chipped by a crushing action. Another important consideration is that, provided that the teeth of a hammer drilling bit are resistant to shock loading, they generally suffer less wear than the teeth of a bit whose cutting action is one of scraping across the

surface to be cut. This is particularly true when drilling abrasive materials such as hard rocks.

Other types of excavation technique include explosive excavation that instantaneously unleashes significant energies. A detonator is used to trigger the explosion and the amount of explosives and their location determines the extent of damage. This method can be used to remove very hard materials but it is very difficult to control and failures can have severe consequences. The use of explosive excavation may be an effective tool for use in planetary applications; however their capability does not allow precise control and therefore they may be applicable to large scale excavation application only.

Given the availability of solar energy, the use of thermal systems may be a good alternative. Generally, there are two principal thermal methods: thermal-spalling and thermal-melting. The difference between these two techniques is that thermal-spalling occurs at lower temperatures of about 400°C to 600°C while thermal-melting followed by vaporization requires temperatures that are typically in the range of $1,100^{\circ}\text{C}$ to $2,200^{\circ}\text{C}$ (Maurer 1968).

Thermal stresses that are generated by applying heat are created by the mismatch in thermal expansion of the constituent minerals and grains within the structure of rocks (Just 1963). However, spalling is caused not only by thermal expansion coefficient mismatch but also by the different expansion between heated and unheated parts of the material being drilled. This process is called thermal spalling and it causes fracturing and degradation of rocks that result in flaking of rock fragments. Its effectiveness depends on the thermal gradients produced in the rock (Fig. 15.8). Exposing a rock to very high temperatures induces melting and vaporization of the rock material (Fig. 15.9). The disadvantage of this technique is the very high power that is required to melt a rock. This process, however, can be readily applied to melt ice. Ice melting drills are frequently used in the Polar

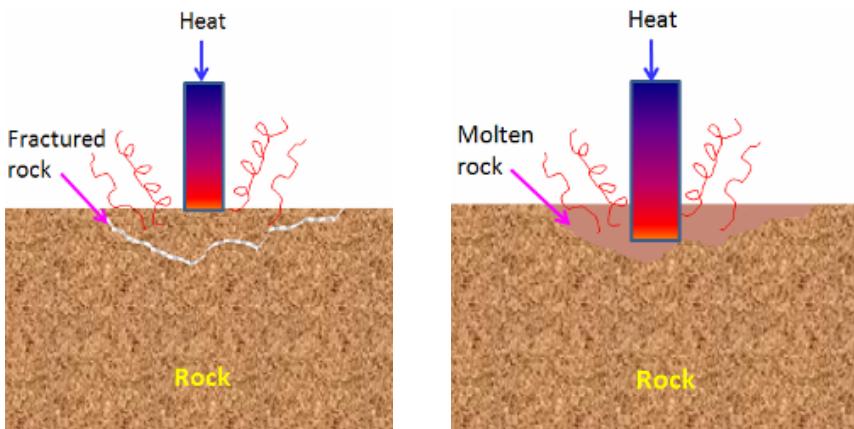


Fig. 15.8 Thermal fracturing of rock formation

Fig. 15.9 Melting and evaporation of rock formation

Regions to drill holes in ice sheets and ice shelves. Penetration rates as fast as one kilometer in ten hours is not uncommon. Ice melting drills have also been proposed for drilling on Europa, a moon of Jupiter, whose surface is composed of water-ice and on Mars, which has a water ice cap at its north pole and has considerable imbedded ice in the subsurface regolith at high latitudes (Bar-Cohen and Zacny 2009).

15.5.3 Drills Developed for Operation on Mars

Various subsurface drills and corers for planetary applications, including Mars, have been developed in recent years through NASA, ESA and JAXA funded projects. In most of these drills rotation is produced primarily by a rotary actuator (combined motor with a gearbox in front that is often coupled to an encoder at the back for precise positioning of the rotary shaft). Some drills, however, utilize rotary-percussive action in which the percussive action is generated by various actuation mechanisms driven at sonic and/or ultrasonic frequencies. The drivers of the sonic mechanism can be mechanical (using for example cam and spring loading, ratchet, etc.); pneumatic (pressurized gas via cam); electromagnetic (e.g., voice coil) and electromechanical (e.g. Ultrasonic/Sonic). The ultrasonic mechanisms are mostly driven by piezoelectric actuators.

A core provides much more information than a handful of cuttings. However, producing cores involves a much more complicated system than a solid face drill that produces cuttings only. A coring system sampler needs to be able to perform the following five key functions: coring, breaking off the core at its root, retaining the core, ejecting the core at the caching location and bit change-out allowing replacement of damaged bits (Bar-Cohen and Zacny 2009; Stroescu et al. 2008). A surface core drill that combines all five subsystems was built by Honeybee Robotics and successfully demonstrated coring from the MER rover (Fig. 15.10). The Corer Abrader Tool (CAT) is a dual purpose drill that not only acquires, captures and ejects 10 cm long cores, but also by changing from a drill bit to a grinding wheel, it can grind rock surfaces in the same manner as the Rock Abrasion Tool (RAT) on the Mars Exploration Rovers. The CAT drill weights only 4 kg and requires 10 Watts of power and 10 Newton of Weight on Bit (preload) to drill into limestone rocks. Cores of hard basalt, such as Saddleback basalt can be acquired in an hour.

In designing a coring mechanism, the designers also have to address operation issues that are critical to the sampling process and they include stabilization to avoid drill walk, drill feed that advances the bit as it drills, cuttings removal to avoid regrinding of cuttings, and sample encapsulation.

Generally, the drilling systems are made to reach various depths ranging from several centimeters to many meters and allow penetrating from soft soil and regolith to hard rocks, permafrost and other materials that pose challenges to penetrate (Bar-Cohen and Zacny 2009). For drilling to a depth of about 10 meters, Honeybee Robotics developed a number of drills, one of which is called the Mars Astrobiology Research and Technology Experiment (MARTE). The MARTE drill shown in Fig. 15.11 was designed for acquiring 25 cm long cores from depths of up to 10-m and the entire drilling, core analysis and borehole logging system was

Fig. 15.10 Honeybee Robotics Coring Abrading Tool autonomously acquires, captures, and ejects 10 cm long cores. It has the ability to change drill bits and can utilize a grinding system for grinding rock surfaces. The CAT shown in the picture was tested in Saddleback basalt from a MER rover at NASA JPL.



put together in 2005 by a team from NASA Ames Research Center, Honeybee Robotics, and Centro de Astrobiología (Spain), (Paulsen et al. 2006). This rotary coring drill has built-in functionality for autonomous core recovery and core delivery. To penetrate beyond depths of 1 meter, the system automatically adds drill tube segments to extend the length of the drill string. Although the MARTE drill uses flutes to transport cuttings away from the drill bit, it only requires flutes on the lead drill string since the cuttings are collected in the top section of the lead drill string and are emptied at the same time as when the core is ejected.



Fig. 15.11 The MARTE drill that can reach 10-meter deep. Courtesy Honeybee Robotics and NASA Ames.



Fig. 15.12 The ATK Modular Planetary Drill System (MPDS) during laboratory drilling of limestone. Courtesy J. Guerrero, ATK-Swales.

Penetrators that can reach depths of tens of meters were studied at ESA, NASA Johnson Space Center, Baker Hughes Incorporated and ATK Space (formerly Swales Aerospace). The Modular Planetary Drill System (MPDS) was developed by ATK Space to reach as deep as 50-meters. In developing this drill the challenges to removal of cuttings, axial preload and borehole maintenance were addressed (Fig. 15.12). Bottom Hole Assembly (BHA), and sample capture mechanisms that allowed drilling through ice, frozen soil and basalt were used. Cores 1.5cm diameter by 10cm length and cuttings were stored separately and brought to the surface to maintain a record of the borehole stratigraphy. The external surface of the drill segments acted as a casing to provide borehole stabilization until the sampling process was completed and the target depth reached. The MPDS was tested at Idaho National Laboratory (INL), Idaho Falls, Idaho, and penetrated ice-soil mixtures with 40W to 50W power. Drilling a frozen soil-mixture required 50W to 70W of power and drilling basalt required about 80W to 100W (Guerrero et al. 2005).

In-situ sampling can also be accomplished using a percussive drill. There are various percussive drills familiar to home handymen. However some of the ones that are available at local hardware stores require power in the range of hundreds of Watts, which will be a major drawback on Mars. As an alternative, one may consider the Ultrasonic/Sonic Driller/Corer (USDC) (Bao et al. 2003; Bar-Cohen et al. 2001, 2005). The drill requires minimal pre load, which is a tremendous advantage on Mars, with low gravity. Figure 15.13 shows the drill being operated while held from its power cord. The USDC drill consists of three components: actuator, free-mass and a bit. A schematic diagram of the USDC mechanism is shown in Fig. 15.14. The USDC operates as a frequency transformer converting 20-kHz ultrasonic waves to a 60-1000 Hz sonic hammering action (percussion). The USDC actuator consists of a stack of piezoelectric ceramics with a backing material that in essence reflects the emission of the acoustic energy forward, and a horn that amplifies the displacements generated by the stack. The tip of the ultrasonic horn impacts the free-mass creating a sonic resonance between the horn and the bit. The USDC has been demonstrated to drill rocks that range in hardness from basalt to soft sandstone and tuff. Other media that were drilled include soil, ice, diorite, and limestone. This novel drill is capable of high-speed drilling (2 to 20 mm/Watt·hr for a 2.85mm diameter bit) in basalt and Bishop Tuff using low axial preload (<10N) and low average power (<12 W). Current analytical models suggest that the USDC performance does not change significantly with changes in ambient gravity.

In order to perform large scale tasks, such as deep drilling, the USDC mechanism was also used to develop a “Gopher” that can acquire core samples using a bit diameter as large as 6.4 cm (Badescu et al. 2006). A schematic diagram of the Gopher and its operating principle are shown in Fig. 15.15. To demonstrate its capability, the Gopher was tested on a glacier at Mt. Hood, Oregon, and the lessons learned were implemented into the new design of the enhanced Gopher for Antarctic testing. The field test in Antarctica was conducted at Lake Vida and it provided an important opportunity to demonstrate the feasibility of this technology while determining the associated challenges and requirements to enhance its

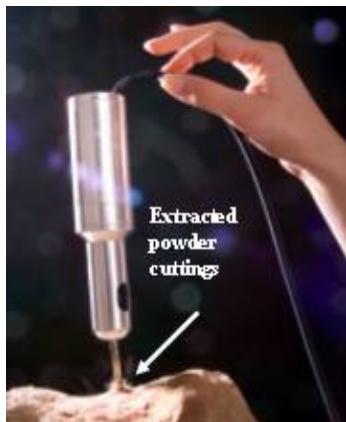


Fig. 15.13 The USDC is shown to require relatively small preload to drill a rock. The powder cuttings travel along the bit providing a removal mechanism for acquisition.

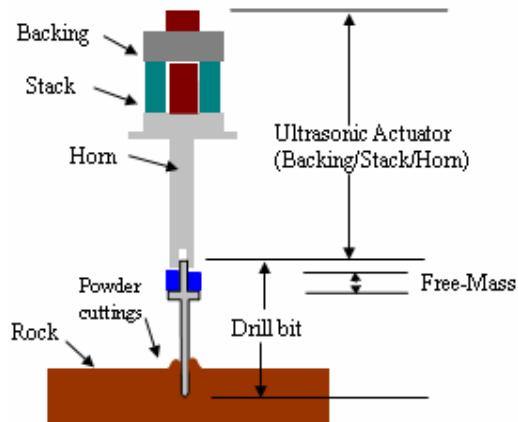


Fig. 15.14 A schematic view of the USDC components

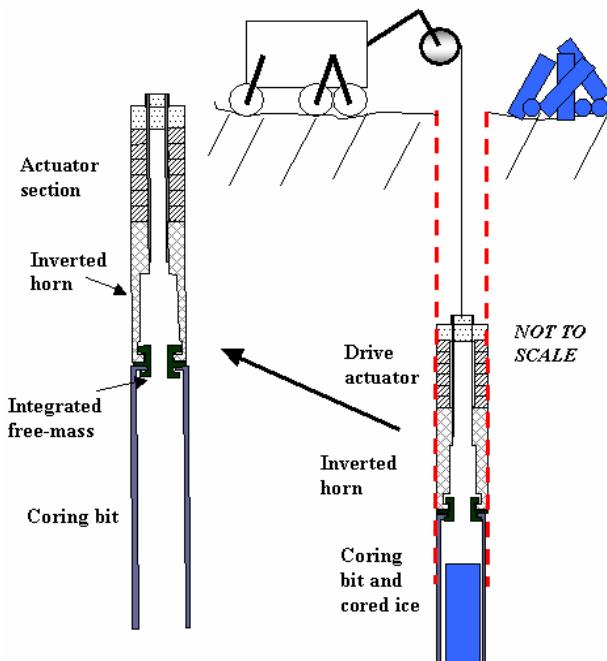


Fig. 15.15 Schematic view of the ultrasonic/Sonic Gopher operating inside the borehole

capability for future drilling objectives. The unit was successfully used to reach 1.76-m deep and it was a major milestone since it is significantly deeper than the length of the whole Gopher assembly.

15.6 Types of Construction Equipment

Mining and construction activities, as opposed to exploration drilling, require the movement of large quantities of material, and therefore the use of some sort of load-haul-dump (LHD) or other excavation equipment.

When dealing with extreme environments on Earth, the mining equipment has to be modified accordingly. For example, in the high Arctic, trucks use low freezing point greases, heaters to keep engines warm, low freezing point coolants, special rubber for tires etc. However, below a certain temperature, which is around -50°C, all operations are halted because everything just freezes and stops working. In very hot environments such as in the Australian outback, mining equipment is modified to deal with extremely high temperatures. Normally, the modification involves additional cooling systems. However, for both extreme environments, the modification of the conventional excavation equipment is not a major one and it is done without drastically altering or completely changing the system. Other terrestrial environments, such as underwater operation would require major system alteration or possibly development of a completely new type of an excavator. Here, the equipment has to be water proof, it cannot use a gasoline engine because of lack of oxygen (air), and it has to withstand very high hydrostatic pressures. For this reason, underwater mining is not as common as surface mining. Underwater mining would require development of new technologies and in turn would be very costly.

Excavation systems pictured for the Moon or Mars often involve some derivatives of terrestrial earthmoving systems. These include backhoes, bucket wheel excavators, bucket ladder excavators and many others. Very seldom one sees some novel approaches to excavation and mining. In this section, we present some conventional approaches that may work on Mars, and also introduce new approaches that are expected to work well only in environments such as found on Mars.

15.6.1 Mining with Sublimation

The key commodities that will be required to sustain human presence are water, oxygen, and fuel for the return journey. These can be obtained by mining water and using electrolysis for splitting water into oxygen and hydrogen. Oxygen and hydrogen may be used as raw ingredients for propellant production and energy storage. Oxygen and water are two ingredients that are required for life support.

As previously mentioned, Mars is a unique planet. It has vast amounts of water both in the northern and southern Polar Regions very close to the surface and deeper down in the subsurface at the equatorial region. The most recent Phoenix 2007 lander revealed that ice is within a few centimeters of the surface in the northern polar region (Fig. 15.16). In addition, Mars is unique in that its atmospheric pressure is very close (above and below) to the triple point of water. The atmospheric pressure in the southern polar region on Mars is always below the



Fig. 15.16 The landing thrusters aboard the Phoenix Mars Lander blew away three to six inches of Martian soil to reveal the shiny, slick face of a large ice patch. Credit: Kenneth Kremer, Marco Di Lorenzo NASA/JPL/ UA /Max Planck Institute/Aviation Week & Space Technology.

triple point, and therefore water-ice when warmed up turns directly into vapor and therefore no water in liquid form can be present. On the other hand, the atmospheric pressure in the northern polar region is always above the triple point of water, and therefore liquid water can exist there but in a very unstable form.

The water-ice sublimation phenomenon was first observed during drilling experiments conducted under Mars pressures (below the triple point as in the southern Polar Regions) in water-ice bearing formations (i.e. frozen rocks that were previously saturated with water) (Zacny et al. 2004). Frictional heat generated during the drilling process was absorbed by the drill bit and the rock. However, the heat that flowed into the rock was used to first warm up the water-ice in the rock and then to turn the ice into vapor. The volumetric expansion of ice as it turns into vapor is on the order of a hundred thousand times. These large volumes of vapor generated at the drill bit successfully blew the drilled chips out of the borehole. In essence, this was like conventional drilling using compressed gas for blowing chips out of the hole except that this compressed gas was being generated in the borehole by warming up water-ice below the drill bit (Fig. 15.17).

This particular phenomenon could also be used as a very effective water mining method (Mungas et al. 2006). In the northern polar region, where the atmospheric pressure is above the triple point, the same mining method will also work, but

Fig. 15.17 Rock cuttings being lifted out of the hole by sublimed in-situ water ice. The heat required for sublimation was generated by drilling friction (Zacny et al. 2004).

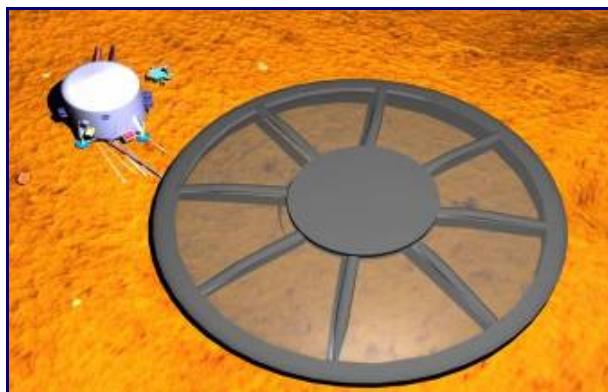
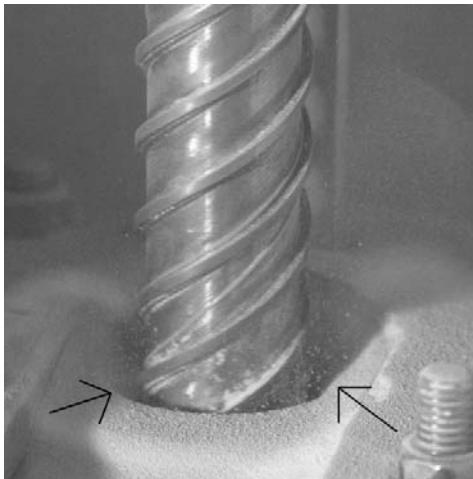


Fig. 15.18 Notional deployable cover for enhancing surface optical properties as well as providing a collection volume for collecting cold steam. (Mungas et al. 2006)

because of higher atmospheric pressure, it may not be as efficient as in the southern polar region. The concept is illustrated in Fig. 15.18. It consists of a green house like structure which traps heat and in turn warms up the ground layer. As the ground heats up, the trapped water-ice sublimes and escapes into the green house. The green house structure has eight strategically positioned trusses that act as cold fingers, i.e. where water vapor condenses. Each of the trusses is linked to a water collection chamber at the bottom. The collected water is then pumped to a close by water storage tank.

This is a very simple method that is entirely passive, i.e. does not require any energy input. The idea is borrowed from nature: the dew that condenses in the morning on the surface of leaves flows toward the plant stem.

15.6.2 Drill and Blast Method

The most conventional open pit, open cast or surface mining approach on Earth uses the so called drill and blast approach. It starts with drilling deep boreholes into a part of the pit known as a bench. The bench is an area that is football-pitch sized or smaller and many meters high. One side of the bench is free (it's like a step of an escalator, having its top and the front surface exposed). Holes are drilled in a certain pattern requested by a blasting engineer. The pattern of holes often looks like a chessboard with a hole at each corner of a square. The holes are a few meters apart with the exact distance being a function of the rock strength and its density of pre-existing fractures. For harder rocks, the separation of the holes is small, while for softer rocks, it is large.

The most modern drill rigs that can be used for drilling blast holes, such as the ones from Atlas Copco or Sandvik (Fig. 15.19) are equipped with many sensors allowing analysis of the drilled formation in terms of hardness or strength from drilling data including power and torque. The final output is information on formation strength, hardness or density as a function of depth. Knowing the strength, hardness or density it is possible to determine the type of rock and in turn the richness of the ore that is being mined. Apart from knowing how much resource lies beneath the surface, the information on the rock strength and the exact location below the surface, allows more accurate placement of explosives. Geologists also sample chips recovered from each hole during the drilling process. These chips or cuttings are analyzed to determine the fraction of the useful resource, such as iron in iron ore, or aluminum in bauxite.

After the holes are drilled, the charges are placed inside each hole. As before, the exact amount of explosive within each hole as well as the exact distribution of explosive within each hole (e.g. more at the bottom or at the top of the hole) is a



Fig. 15.19 Drill and Blast mining starts with drilling blast holes on the bench followed by blasting and loading onto the trucks. The photograph shows a surface drill rig for drilling blast holes.

function of where harder formations are located (if a harder layer is on top, more explosive is placed there). The blasting sequence is carefully choreographed, making sure there is a sufficient time delay between each line of holes. A good blasting sequence will result for example in the first layer breaking out, followed by a second layer, and so on, as in a domino effect.

After the blast, a front end loader loads the blasted rock onto the trucks that take it to a primary crusher. Rocks are first crushed into relatively coarse size, of the order of centimeters, and then through a secondary and tertiary crusher system into fines. The fines are then chemically processed in order to recover the valuable resource.

Resource recovery on Mars on a larger or smaller scale would most probably follow the same steps as described above. This is because using explosives to break up hard rock is one of the most efficient mining systems. An alternative approach would use a mechanical excavator such as a Surface Miner (SM) or a Continuous Miner (CM). For soft formations, that approach does work, but for harder rocks in excess of 80 MPa, the SM or CM becomes very ineffective and inefficient. In addition, considering the much lower Martian gravity (approximately three times lower than on Earth), the SM or CM would have to be three times more massive to exert the same effective forces. Thus, for harder formations, the above conventional drill and blast approach would still be most effective. Note that on the Moon or Mars, because of the reduced gravity, blasted rock would be thrown much further, and so the mining equipment would have to be moved much further back to be safe during blasting operations.

15.6.3 Surface Miner/Continuous Miner

A Surface Miner or Continuous Miner uses a rotating drum with tungsten carbide teeth or picks to break up and churn underlying rock or hard soils (Fig. 15.20). In a way, it operates in the same way as a conventional combined harvester. There is a number of advantages that SM or CM offers. Firstly, the formation is being mined and partially processed in an uninterrupted manner, that does not involve the cyclical movement of drilling, explosive charging and LHD equipment to and from the working area, and the final product is much finer than from drill and blast mining. The surface miner can also be deployed on any surface, smooth or rough and mining can be selective and this means that only areas or sections with high grade ore will be mined while the low grade area will be left behind. The big problem SM and CM suffers from is that the machines do not work well in hard formations. In very hard rocks, the machine vibrates excessively, tungsten carbide teeth wear out very quickly and have to be replaced, and the required power becomes exceedingly large.

If SM or CM is required for hard rocks, such formations would have to be pre-treated including possibly such pre-conditioning as microwave irradiation (Lindroth et al. 1992). It has been found that microwaves tend to induce cracks in different types of rocks and in turn make the rocks much weaker. Another solution may be the use of small scale blasting (Lin et al. 1994). Much smaller holes can be drilled and only small amounts of explosive can be put in with just enough quantity to induce cracks and fractures and allow SM or CM to perform selective mining.



Fig. 15.20 Surface miners such as Wirtgen Surface Miner 2200 SM can be used to continuously mine and dump the resources into the truck. Wirtgen GmbH, German.

15.6.4 Percussive/Vibratory Enhanced Excavation

Soil or ground excavators, such as back hoes, use brute force to dig up soil while for harder soils, a larger excavator is often used. However, on Mars where the gravity is three times lower than on Earth, bringing large excavators will be prohibitively expensive. An alternative to the use of brute force is to use vibrator or percussive systems. Percussive scoops can dig into the same soils with 95% less force (Zacny et al. 2008a). This translates into large reduction in the mass of the excavator (not 95%, because the excavator has to be three times the mass to exert the same excavation force). Shabo et al. (1998) also found that vibrating of bulldozer blades helps to reduce draft force up to 71%-93%. This means that Mars bulldozer can be less massive (Fig. 15.21). The disadvantage of a percussive or vibratory excavator is that it requires additional power to run the percussive motor or vibrator. However, given the fact that solar energy can be harvested essentially for free, in a tradeoff between lower mass or high power, lower mass will always win. Therefore, percussive or vibratory systems offer ideal capability for excavating on planets with low gravity, such as on Mars.

15.6.5 The Martian Mine of the Future on Earth?

The future mine on Mars will most probably in a way look like one of the futuristic mines on Earth. Currently, some mining companies, such as Rio Tinto are undertaking the tremendous task of fully automating all mining operations. These



Fig. 15.21 NASA Chariot rover fitted with a bulldozer blade. Photo: Sean Smith/NASA

operations include autonomous drilling, blasting, loading, hauling, crushing, in-situ resource assessment and many other activities. For this purpose, Rio Tinto is working with many equipment suppliers to achieve this goal. Moreover, Atlas Copco (an equipment manufacturer) has been developing drill rigs that can autonomously drill to the required depths, at pre-determined angles and that use a GPS system to pinpoint exact drilling locations. By acquiring drilling telemetry such as power and penetration rate, these rigs can give an indication of the hardness of rocks below the surface and via correlations, may even indicate the exact types of rocks and pin point the location of ores.

Truck companies such as Komatsu or Caterpillar are working on autonomous trammimg systems. These 200 ton capacity trucks are equipped with many sensors, collision avoidance systems, stereo cameras, GPS and highly elaborate software that computes the truck's every move and can position a truck within about 2.5 cm (1 inch) of a crusher or a loader.

The mine of the future will use few employees, and even those, will be housed many miles from the mine site, which often is in a desolate, hot or cold area. Lessons learned from developing and operating these mines will no doubt help in designing the future mines on Mars.

15.7 Conclusions

Drilling and excavation for construction and in-situ resource utilization (ISRU) require establishing capabilities that will allow humans to operate safely and comfortably in extended missions, and on a far away planet. Some of the required tools such as drills and excavators may be based on mature earth proven technologies but they will have to be adapted for the more extreme environment on Mars. New technologies, applicable only to Mars, may also have to be developed.

Terrestrial systems are becoming more efficient and are increasingly using effective analytical modeling approaches and capabilities to optimize their performance. The latest area of advancement is automation (automating all aspects) of drilling and excavation. On Earth, work in this area is driven by efforts to reduce the use of human labor, increase the profit and to remove human operators from extreme environments. On Mars, the automation will be driven by almost the same motivation: removing humans from dangerous environments and reducing cost, through automating mining operations and allowing future 'Martians' to do more important tasks.

Acknowledgement

Some of the research reported in this chapter was conducted at the Jet Propulsion Laboratory (JPL), California Institute of Technology, under a contract with National Aeronautics and Space Administration (NASA). The authors would like to acknowledge the personnel at the JPL's NDEAA Laboratory for their contribution to the development of the ultrasonic/sonic drilling technologies. In addition, we would like to acknowledge contributions from Honeybee Robotics Spacecraft Mechanisms Corporation.

References

- Bao, X., Bar-Cohen, Y., Chang, Z., Dolgin, B.P., Sherrit, S., Pal, D.S., Du, S., Peterson, T.: Modeling and Computer Simulation of Ultrasonic/Sonic Driller/Corer (USDC). *IEEE Transaction on Ultrasonics, Ferroelectrics and Frequency Control (UFFC)* 50, 1147-1160 (2003)
- Bar-Cohen, Y., Sherrit, S., Dolgin, B.P., Bridges, N., Bao, X., Chang, Z., Saunders, R.S., Pal, D., Kroh, J., Peterson, T.: Ultrasonic/Sonic Driller/Corer (USDC) as a sampler for planetary exploration. In: Proceedings of the 2001 IEEE Aerospace Conference on the topic of Missions, Systems, and Instruments for In Situ Sensing (Session 2.05), Big Sky, Montana, March 10-17 (2001)
- Bar-Cohen, Y., Chang, Z.S., Sherrit, S., Badescu, M., Bao, X.: The Ultrasonic/Sonic Driller/Corer (USDC) as a subsurface drill, sampler, and lab-on-a-drill for planetary exploration applications. In: Industrial and Commercial Applications of Smart Structures Technologies, SPIE Smart Structures and Materials Symposium, Paper #5762-21, San Diego, CA (2005)
- Bar-Cohen, Y., Zacny, K. (eds.): *Drilling in Extreme Environments - Penetration and Sampling on Earth and Other Planets*. Wiley – VCH, Hoboken (2009)
- Bar-Cohen, Y., Hanson, D.: *The Coming Robot Revolution - Expectations and Fears About Emerging Intelligent, Humanlike Machines*. Springer, New York (2009)
- Badescu, M., Sherrit, S., Olorunsola, A., Aldrich, J., Bao, X., Bar-Cohen, Y., Chang, Z., Doran, P.T., Fritszen, C.H., Kenig, F., McKay, C.P., Murray, A., Du, S., Peterson, T., Song, T.: Ultrasonic/sonic Gopher for subsurface ice and brine sampling: analysis and fabrication challenges, and testing results. In: SPIE Smart Structures and Materials Symposium, San Diego, CA, February 27 -March 2 (2006)

- Bruzel, A.: Analysis of Steel from the Titanic (2008), <http://dwb.unl.edu/Teacher/NSF/C10/C10Links/chemistry.about.com/library/weekly/aa022800a.htm> (accessed, 15 January 2008)
- Clancy, R.T., Sandor, B.J., Wolff, M.J., Christensen, P.R., Smith, M.D., Pearl, J.C., Conrath, B.J., Wilson, R.J.: An intercomparison of ground-based millimeter, MGS TES, and Viking atmospheric temperature measurements: seasonal and interannual variability of temperatures and dust loading in the global Mars atmosphere. *Science* 105, 9553–9572 (2000)
- Guerrero, J.L., Craig, J., Meyers, M., Reiter, J.W., Wang, G.Y., Wu, D.: Planetary Drill Technology and Applications to Future Space Missions. In: Proceeding Space Resources Roundtable VII: LEAG Conference on Lunar Exploration. Lunar and Planetary Institute, Houston (2005)
- Han, G., Dusseault, M.B., Detournay, E., Thomson, B.J., Zacny, K.: Principles of Drilling and Excavation. In: Bar-Cohen and Zacny, ch. 2 (2009)
- Lin, C., Goodings, D., Bernold, L., Dick, R., Fourney, W.: Model Studies of Effects on Lunar Soil of Chemical Explosions. *Journal of Geotechnical Engineering* 120, 1684–1703 (1994)
- Lindroth, D.P., Berglund, W., Morrell, R., Blair, J.: Microwave-assisted drilling in hard rock. *Mining Engineering* 292, 1159–1163 (1992)
- Martin, T., Bridges, N., Murphy, J.: Near-surface temperatures at proposed Mars Exploration Rover landing sites. *J. Geophys. Res.* 108, 301–307 (2003)
- Maurer, W.C.: Novel Drilling Techniques. Pergamon Press, New York (1968)
- Mellor, M.: Strength and deformability of rocks at low temperatures. *Cold Regions Research and Engineering Laboratory, Research Report* 294 (1971)
- Mungas, G., Rapp, D.R., Easter, R.W., Johnson, K.R., Wilson, T.: Sublimation Extraction of Mars H₂O for Future In-Situ Resource Utilization. In: Earth and Space Conference 2006, Houston, TX (2006)
- Paulsen, G., Zacny, K., Cannon, H., Glass, B., Chu, P., Mumm, E., Davis, K., Frader-Thompson, S., Petrich, K., Bartlett, P., Glaser, D.: Robotic Drill Systems for Planetary Exploration. In: Proceedings of AIAA Space San Jose, California (2006)
- Presley, M.A., Craddock, R.A.: Thermal conductivity measurements of particulate materials: 3. Natural samples and mixtures of particle sizes. *J. Geophys. Res.* 111, E09013 (2006)
- Smith, D., Zuber, M.: The relationship between MOLA northern hemisphere topography and the 6.1 mbar atmospheric pressure surface of Mars. *Geophys. Res. Lett.* 25, 4397–4400 (1998)
- Shabo, B., Barnes, F., Sture, S., Ko, H.: Effectiveness of vibrating bulldozer and plow blades on draft force reduction. *Transactions of the ASAE* 41, 283–290 (1998)
- Stroescu, S., Bar Cohen, Y., Badescu, M., Bao, X., Sherrit, S. (2008) Subsurface Sample Acquisition & Delivery Using the Ultrasonic/Sonic Driller/Corer (USDC) for Analysis by the Miniature X-Ray Diffraction (XRD) System. Honeybee Robotics final ASTEP report, HR Document # 138.RPT.001 (September 2008)
- Zacny, K.A., Quayle, M.C., Cooper, G.A.: Laboratory drilling under Martian conditions yields unexpected results. *J. Geophys. Res.* 109, E07S16 (2004)
- Zacny, K.A., Cooper, G.A.: Friction of drill bits under Martian pressure. *J. Geophys. Res.* 112, E03003 (2007)
- Zacny, K.A., Glaser, D., Bartlett, P., Davis, K., Gorevan, S.: Drilling Results in Ice-Bound Simulated Lunar Regolith. In: Space Technology and Applications International Forum 2007, Albuquerque, New Mexico, February 11–15 (2007)

Zacny, K., Craft, J., Wilson, J., Chu, P., Davis, K.: Percussive Digging Tool for Lunar Excavation and Mining Applications. Abstract 4046, LEAG-ICEUM-SRR, Cape Canaveral, FL (October 28-31, 2008a)

Zacny, K., Bar-Cohen, Y., Boucher, D., Brennan, M., Briggs, G., Cooper, G., Davis, K., Dolgin, B., Glaser, D., Glass, B., Gorevan, S., Guerrero, J., Paulsen, G., Stanley, S., Stoker, C.: Drilling Systems for Extraterrestrial Subsurface Exploration. *J. Astrobiology* 8, 665–706 (2008b)

Chapter 16

In Situ Utilization of Indigenous Resources

Donald Rapp

University of Southern California, Los Angeles, USA

16.1 Introduction

In situ resource utilization (ISRU) on the Moon or Mars is an approach for converting indigenous resources into various products that are needed for a space mission (Rapp 2007). By utilizing indigenous resources, the amount of materiel that must be brought from Earth may be reduced, thus reducing the *Initial Mass in Low Earth Orbit* (IMLEO.) IMLEO is typically used as a measure of the mission scope and cost. Mars mission planners deal extensively with IMLEO, and the problems involved in launching that materiel and sending it out of LEO on its way toward Mars.

ISRU has the greatest value when the following ratio is large:

$$R = [\text{mass of products supplied by ISRU to mission}] / [\text{mass of the ISRU system brought from Earth}]$$

Thus in order for ISRU to have net value, it is essential that the mass of the ISRU system (i.e. the sum of the masses of the ISRU plant, the power plant to drive it, and any feedstocks brought from Earth) must be less than the mass of products produced and used by the mission. If $R \gg 1$, then in comparing the IMLEO for two similar missions, one using ISRU, and the other not using ISRU, the IMLEO using ISRU will be lower. This comparison of IMLEO with and without ISRU will provide one measure of the "value" of ISRU. However, from a broader point of view, one should compare total investments (rather than IMLEO) with and without the use of ISRU. In this regard, the investment in ISRU includes the costs of (a) prospecting to locate and validate accessibility of indigenous resources, (b) developing and demonstrating capabilities to extract indigenous resources, (c) developing capabilities for processing indigenous resources to convert them to needed products, (d) additional operational costs in running or servicing ISRU systems in space, and (e) any ancillary requirements specifically dictated by use of ISRU (e.g. possibly a nuclear power system). The potential cost saving that might be realized by using ISRU is essentially equivalent to the investment that is eliminated by reducing IMLEO for as many launches as the ISRU system serves. If this potential cost saving is greater than the investment required to develop and implement ISRU, then ISRU has net value for a mission or a set of related missions. If one only compares ISRU system mass with ISRU product mass, one may in some cases conclude (incorrectly) that ISRU is beneficial, when in actuality, ISRU adds to the overall mission cost. Typically, NASA plans human missions

without the use of ISRU, and then considers tacking on ISRU rather late in the campaign as an embellishment. This limits the benefits of ISRU because all vehicles are sized without taking into consideration the benefits that ISRU could yield.

In the short run (next ~ 40 years) the main products that might be supplied to human missions by ISRU are:

- Propellants for: 1) ascent of return vehicles, and 2) intra-surface transportation, thus reducing the mass of ascent propellants that otherwise would need to be brought from Earth.
- Life support consumables (water and oxygen). It is possible that while on the surface, the usual life support cyclic system could be eliminated or downsized by using ISRU as an alternative to provide these consumables.

In addition, in principle, regolith can be piled up on top of a habitat for radiation shielding. However, it remains to be seen whether NASA designs for surface habitats allow for use of regolith as shielding from radiation. Existing designs for lunar and Mars habitats do not appear to be compatible with regolith radiation shielding.

In the longer run, it is possible that a wider range of products could be produced as the industrial and electronic revolution is transferred from Earth to extra-terrestrial bodies. Unfortunately, there does not seem to be a clear path leading from where we are now to such an ultimate utopia.

Although this book is primarily focused on Mars, it is worthwhile to also examine NASA plans for lunar ISRU because NASA seems to be intent on concentrating its resources and efforts on lunar ISRU for the next couple of decades, and lunar ISRU is viewed by many as a stepping-stone to Mars ISRU.

16.2 Potential Products

The NASA Human Exploration Initiative views lunar exploration as a stepping-stone to Mars. One of the important stepping-stones is demonstration of ISRU technology. The term "ISRU" occurs 106 times in the 2005-6 Exploration Systems Architecture Study (ESAS) Report, and the choice phrase "living off the land" occurs several times.

Most discussions of lunar ISRU seem to assume that resources are readily available, and they proceed to emphasize processing, while minimizing logistics (prospecting, excavating regolith, regolith transport, deposition into and removal of regolith from reactor, dumping waste regolith, etc.) However, the quantity and composition of end products provides the basis for considering the use of ISRU as well as for setting the requirements for ISRU systems. Therefore, we begin here with the potential end products.

16.2.1 Lunar Ascent Propellants

In the initial NASA ESAS architecture of 2005-6, the propulsion system for ascent from the Moon was based on $\text{CH}_4 + \text{O}_2$ propellants in order that ISRU-generated oxygen from the Moon could be utilized in place of oxygen brought from Earth.

Although methane would need to be brought from Earth, use of the $\text{CH}_4 + \text{O}_2$ propulsion system provided an implicit connection to Mars ISRU by using oxygen as the oxidizer for ascent. Later, when the realities of cost and schedule to develop the $\text{CH}_4 + \text{O}_2$ propulsion system became clearer, this ascent propulsion system was dropped in favor of space storable propellants that are incompatible with lunar ISRU. Yet, NASA continued to claim that ISRU was a major part of the lunar exploration program.

In the original 2005-6 architecture, the plan was to have two ascents per year from a polar outpost, each requiring about 4 metric tons (mT) of oxygen, for an annual need of roughly 8 mT. However, in a 2007 release (Connolly 2007) NASA persisted in specifying space storable propellants for ascent, and these propellants are incompatible with ISRU.

Since the "gear ratio" (required mass in LEO/mass delivered to surface) for lunar polar outposts is about 4:1, the potential mass saving in LEO from use of ISRU is ~ 16 mT per launch. However, because the launch vehicles were designed without using lunar ISRU, their designs will remain unaffected by later inclusion of lunar ISRU in the missions. Nevertheless, they would not need to carry 4 mT of oxygen if ISRU were employed. Hence they would be able to deliver an extra infrastructure payload (~ 4 mT) to the outpost with each launch (but rather late in the campaign - probably beginning in the late 2020s). (Even this minor benefit disappears if NASA persists in its present plan to use space storable propellants (NTO/MMH) for ascent, thus eliminating oxygen as an ascent propellant.) The "value" of the ~ 4 mT increase in payload delivery per launch using lunar ISRU can be estimated by noting that over a period of years, with continual infrastructure deliveries to the outpost, a cargo delivery launch might be eliminated once every several years with small incremental increases in mass delivered by each launch.

However, if the design policy "abort to orbit" remains a requirement for descent, then there is no possibility of providing ascent propellants using ISRU since the incoming Lunar Surface Access Module (LSAM) would have to possess ascent propellants as a safety measure.

16.2.2 Mars Ascent Propellants

A vital ISRU product of relevance to Mars missions is oxygen for use as a propellant in ascent from Mars. This oxygen would be stored as a cryogenic liquid in the Mars Ascent Vehicle (MAV) on the Mars surface. The amount of oxygen required will depend on several factors: (1) the mass of the capsule to hold the crew during ascent and rendezvous, (2) the number of crew-members, (3) the orbit in which rendezvous takes place, and (4) the fuel propellant used in conjunction with the oxygen. In a typical rocket using oxygen as the oxidizer (e.g. methane-oxygen rocket) the oxygen accounts for 75-80% of total propellant mass depending on the actual mixture ratio used. Thus if the methane (or hydrogen to produce methane from the CO_2 in the Mars atmosphere) is brought from Earth, and ISRU produces only oxygen from Mars feedstocks, ISRU would still provide 75-80% of ascent

propellant needs. Some forms of Mars ISRU produce not only oxygen, but methane as well. In that case, 100% of ascent propellant needs would be supplied by ISRU.

16.2.3 Life Support Consumables

Mars ISRU can also be used to produce life support consumables. The requirement for the surface phase of a human mission to Mars with a crew of six was given in Table 4.2 of Rapp (2007). The requirement is for about 100 mT of water and about 4 mT of oxygen. ISRU would supply these commodities. However, the mass benefit from using ISRU to supply these commodities depends on assumptions made regarding the efficiency and mass of competing recycling systems for these commodities. It is likely that the systems for air and water recycling in an Environmental Control and Life Support System (ECLSS) would have considerably lower mass than the 104 mT of commodities, but the longevity and reliability of such systems remains open to question.

Rapp (2007) estimated the reduction in IMLEO that would result from use of ISRU to produce methane and oxygen on the Martian surface from indigenous resources. The total propellant requirements for three operations were included (orbit insertion, ascent, and orbit departure). It was found that the benefit from ISRU was much greater if an elliptical orbit is employed for staging. The reason for this is that entry into an elliptical orbit is easier (i.e. requires less propellant) than entry into (and exit from) a circular orbit, but ascent to a circular orbit is much easier than ascent to an elliptical orbit. However the higher propellant requirement for ascent to an elliptical orbit can be furnished by ISRU. Based on a particular mission concept, the reduction in IMLEO from use of ISRU was estimated by Rapp (2007) to be about 130 metric tons (mT) using a circular orbit for staging, and about 370 mT using an elliptical orbit.

The mass saving due to ISRU propellant production is even greater if a mission architecture is used in which the Mars Ascent Vehicle (MAV) does not rendezvous with an Earth Return Vehicle (ERV) in Mars orbit, but instead, the MAV goes directly from the Mars surface all the way back to Earth. The Mars Direct (Zubrin et al. 1991; Zubrin 2000) and Mars Society (Hirata et al. 1999) missions used this approach, and the MIT study (Wooster 2005; Taylor et al. 2005; Wooster et al 2005) found significant benefits to this architecture. However it appears unlikely that NASA will consider this approach.

Oxygen requirements depend on crew activity but an average value is about 1 kg per crew-member (CM) per day. Water requirements have been estimated by JSC to be about 27 kg/CM-day. (See Table 4.1 of Rapp (2007)).

To support a crew of 6 for one year, we require $6 \times 1 \times 365 = 2,190$ kg ~ 2.2 mT of O₂, and $6 \times 27 \times 365 \sim 60$ mT of water.

It is likely that an ECLSS would be used to recycle these resources, thus greatly reducing consumable mass requirements. Therefore, the benefit of ISRU is not replacement of resources, but rather replacement of the mass of the ECLSS. JSC has estimated the mass of ECLSS systems. Using ISS experience as a basis, JSC estimated the mass and power requirements of ECLSS systems for a crew of six on Mars for 600 days as shown in Table 16.1.

Table 16.1 Mass and power requirements of ECLSS systems for a crew of six on Mars for 600 days

System	Physical plant mass (kg)	Backup cache mass (kg)	Total mass (kg)
Oxygen ECLSS	765	570	1,335
Water ECLSS	6,750	4050	10,800

For a typical Mars mission design, the stay on the surface would be about 1.5 years, so the total water/oxygen requirement would be about 93 mT for a crew of six, and the estimated mass of the ECLSS to supply these resources would be 18 mT. However, this estimate of the ECLSS mass is based on very sparse data and may represent an optimistic evaluation. Furthermore, if an ISRU system were used to supply water and oxygen, an ECLSS would still be required to maintain environmental requirements even if recycling is not used. However, such an ECLSS would be much less complex than one that recycles.

16.2.4 Propellants Delivered to LEO from the Moon

For a typical Mars-bound vehicle in LEO prior to trans-Mars injection, about 60% of the total mass in LEO consists of H₂ + O₂ propellants for trans-Mars injection. If Mars-bound vehicles could be fueled in LEO with H₂ and O₂ delivered from the Moon, then only the remaining 40% of the total vehicle wet mass would need to be delivered from Earth to LEO. The other 60% would be provided from lunar resources. For example, a Mars-bound vehicle that weighs say, 250 mT in LEO, would include about 150 mT of propellants for trans-Mars injection. If fueled by hydrogen and oxygen from the Moon, the mass that would have to be lifted from Earth to LEO would only be about 100 mT instead of 250 mT. This would have a huge beneficial impact on the feasibility of launching large Mars-bound vehicles.

In this regard, the question that we must deal with is: how feasible is it to transfer water (and then by electrolysis, produce H₂ + O₂) from the Moon to LEO? If this process is efficient, the scheme of supplying propellants to LEO from the Moon may be less costly than launching the propellants from Earth. If the transfer process is very inefficient, it is likely to be less costly to simply deliver propellants to LEO from Earth.

It is implicitly assumed here that accessible water ice can be exploited from polar areas on the Moon. If that is not the case, this entire concept becomes moot. Furthermore, the process may become untenable if the transfer vehicle masses are too high. If these vehicles are too heavy, all the water ice excavated on the Moon would be used to produce H₂ + O₂ to deliver the vehicles, and ultimately no net transfer of water to LEO would be feasible. Therefore, it is necessary to examine the details of the transfer process and estimate what percentage of water excavated on the Moon can be transferred to LEO. The percentage of water mined on the Moon that can be transferred from the Moon to LEO for fueling Mars-bound vehicles can be estimated as discussed in Sec. 5.4 of Rapp (2007). The figure of merit is the net percentage of water mined on the Moon that can be transported to LEO for use by Mars-bound vehicles. As this percentage increases, the cost of transporting

water to LEO from the Moon becomes more favorable. However, the best estimate is that most of the water excavated on the Moon is used up in transferring the tankers to LEO, and almost no net water is transferred to LEO. This estimate is based on standard spacecraft design principles. On the other hand, if these tanker vehicles can be made much less massive using some unspecified advanced technology, such transfer might one day become feasible.

16.2.5 Propellants Delivered to Lunar Orbit for Descent (and Ascent)

Whereas the amount of oxygen required for ascent from the Moon is a rather modest ~ 4 mT, the amount of oxygen required for descent is well over 20 mT. These propellant masses are based on 2005-6 data. In February 2007, a NASA release indicated that oxygen will not be used for ascent, and over 30 mT of LOX/LH₂ will be used for descent (Connolly 2007). If oxygen (and less importantly hydrogen as well) can be delivered to lunar orbit for fueling Moon-bound descent vehicles, the potential payoff from lunar ISRU would be much higher than if lunar ISRU were used only for ascent propellants. The gear ratio (required mass in LEO/mass delivered to lunar orbit) is roughly 2.5. Therefore generation of oxygen via lunar ISRU for use as a descent propellant would save > 55 mT in LEO. The combination of lunar ISRU-provided ascent and descent propellants (hydrogen + oxygen) would save more than 80 mT in LEO, and this mass saving is likely to increase if vehicles become more massive. The concept would then be as follows:

NASA would begin by establishing an outpost in a shadowed polar area of the Moon to excavate regolith, extract water, and to some extent, electrolyze water and store hydrogen and oxygen. This would have to be done robotically without crew participation. Is this possible? Who knows?

NASA would design and implement a tanker system for transferring water from the surface of the Moon to lunar orbit, and establish a filling station in lunar orbit to electrolyze water and fill tanks on incoming vehicles with hydrogen and oxygen. This tanker system would act as a shuttle to move back and forth between the lunar surface and lunar orbit, carrying full tanks on the way up and empty tanks on the way down. The percentage of water extracted on the lunar surface that can be delivered to lunar orbit (after providing propellants for descent of the empty tanker) was estimated in Sec. 5.5 of Rapp (2007).

Incoming LSAM vehicles on their way to the surface of the Moon would carry empty ascent and descent tanks, and would be fueled in lunar orbit prior to descent. In case of an unexpected problem, the crew could return in the CEV and never descend in the LSAM.

However, as in the case of lunar ISRU providing only ascent propellants, this > 80 mT reduction will not be realized in terms of reduced launch vehicle capability if lunar ISRU is adopted as an afterthought late in the campaign.

16.2.6 Regolith for Radiation Shielding

Use of regolith piled on top of habitats for radiation shielding is probably a legitimate potential use of in situ resources, but the requirements and benefits require

further study. Current habitat designs and plans for installing them on the lunar surface do not seem to be compatible with regolith shielding, nor is it known how to install the regolith.

16.2.7 Visionary Concepts

Visionaries and futurists have proposed a variety of ISRU applications to produce liquid fuels, metals, plastics, huge fields of solar cells, structural materials and electronics from indigenous resources on the Moon and Mars. It seems likely that such approaches will be relegated to generations beyond providing propellants and life support.

16.3 Lunar ISRU

16.3.1 Lunar Resources

There are basically four potential lunar resources:

- Silicates in regolith containing typically > 40 wt% oxygen.
- Regolith containing FeO for hydrogen reduction. FeO content may vary from 5 wt% to 14 wt% leading to recoverable oxygen content in the 1-3 wt% range.
- Imbedded atoms in regolith from solar wind (typically parts per million).
- Water ice in regolith pores in permanently shadowed craters near the poles (unknown percentage but possibly as high as a few percent in some locations).

16.3.2 Lunar ISRU Processing

16.3.2.1 Oxygen from FeO in Regolith

Hydrogen reduction of regolith depends on the reaction of hydrogen with FeO in the regolith to produce iron and steam $[FeO + H_2 = Fe^0 + H_2O]$. The non-FeO fraction of the regolith does not enter into the reaction. The water (steam) produced in the reactor (at ~ 1300 K) is subsequently separated by condensation and then electrolyzed. The oxygen is collected and saved while the hydrogen is recirculated. Some make-up hydrogen will be needed, as this process will not be 100% efficient. It is not clear how the regolith is fed into the reactor and withdrawn from the reactor. It is also not clear how one prevents "gunking up" within the reactor. Some heat recuperation can be accomplished by using heat from steam and perhaps spent regolith to pre-heat incoming regolith, but those measures will add complexity (and mass) to the process.

The expected recoverable oxygen (in wt%) from a given mass of lunar regolith varies with FeO content from 3% in Mare regions to about 1% in highlands. The projected power requirements to produce oxygen from lunar mare, assuming that solar power is used and that the duty cycle for the process is 40% (3500 hours of processing per year), are given in Table 16.2 (assuming 50% heat recovery and linear scalability).

Table 16.2 Projected power requirements to produce oxygen from lunar mare

Annual Oxygen Production Rate (mT) \Rightarrow	1	10	50	100
Annual regolith rate (mT)	34	336	1,681	3,361
1000s of kWh	5.1	51	255	510
Hours	3,500	3,500	3,500	3,500
kW to heat regolith	1.44	14.4	72	144

The technical and economic feasibility of this process has yet to be demonstrated, although recent pilot plant tests have shown that the reaction can be made to proceed.

16.3.2.2 Oxygen from Regolith Silicates

Lunar ISRU based on extraction of oxygen from regolith has two advantages compared to reduction of FeO:

1. Regolith is typically $> 40\%$ oxygen.
2. Regolith is available everywhere and solar energy may be feasible for processing.

Unfortunately, the oxygen in regolith is tied up in silicate bonds that are amongst the strongest chemical bonds that are known, and breaking these bonds inevitably requires very high temperatures and energy inputs.

JSC has investigated several processes for extracting oxygen from lunar silicates. One is the carbothermal process. This concept is based upon a high-temperature, direct energy processing technique to produce oxygen, silicon, iron, and ceramic materials from lunar regolith via carbonaceous high-temperature (carbothermal) reduction at ~ 2600 K. To prevent destruction of the container, they apply heat to a localized region of regolith and the surrounding regolith acts as an insulating barrier to protect the support structures. The plan is to use a set of solar parabolic dish concentrators to beam light directly onto the regolith in the carbothermal reduction cells. Methane gas is introduced into the reduction chamber to provide a source of carbon and hydrogen to scavenge oxygen. According to JSC:

"The lunar regolith will absorb the solar energy and form a small region of molten regolith. A layer of unmelted regolith underneath the molten region will insulate the processing tray from the solar energy. Methane gas in the reduction chamber will crack on the surface of the molten regolith producing carbon and hydrogen. The carbon will diffuse into the molten regolith and reduce the oxides in the melt while the hydrogen gas is released into the chamber. Some hydrogen may reduce the iron oxides in the regolith to form water, which will be recovered by the carbothermal system. A moveable solar concentrator will allow heating in the form of a concentrated beam on the regolith surface. A system of fiber optic cables will distribute the concentrated solar power to small cavities formed by reflector cups that concentrate and refocus any reflected energy. Solidified slag melts are removed from the regolith bed by a rake system. Slag waste and incoming fresh regolith are moved out or into the chamber through a double airlock system to minimize the loss of reactive gases."

This far-fetched scheme would be a nightmare to carry out on Earth. On the Moon, it would be far worse. Preliminary testing has not produced any encouraging results. In the extremely unlikely case that a high-temperature processor for oxygen from regolith on the Moon can be made into a practical unit, one would still be faced with the challenges (and costs) for development and demonstration of autonomous lunar ISRU systems for excavation of regolith, delivery of regolith to the high-temperature processor, operation of the high-temperature processor with free flow of regolith through it (without caking, agglomeration and "gunking up" of regolith), and removal of spent regolith from the high-temperature processor to a waste dump.

16.3.2.3 Extracting Volatiles

Analysis of lunar rocks from the Apollo missions indicated that heating of the lunar rocks evolved a variety of volatile materials. Hydrogen and nitrogen were reported to be present at the concentration of 10-20 ppm.

Based on this, JSC is seriously considering the prospect of extracting hydrogen for use as a propellant, and nitrogen for use as an oxygen diluent in breathing air – assuming "a best case scenario" that these putative volatiles are available at the 150 ppm level. It is assumed that the volatiles will be released when the regolith is heated to ~ 800 K. Assuming that the regolith starts at say, 200 K, this involves raising the temperature of the regolith by 600°C.

JSC has developed several fanciful concepts for implementing this process. One conceptual process uses "a large inflatable dome that has a center-driven scraper-wand similar to an agricultural silo top-unloading device." The scraper moves in a circular sweep and the regolith is directed by a sort of Rube Goldberg arrangement to a ramp where it is heated by IR or microwave heaters. Evolved volatiles are collected by means of either a cryocooler (for N₂) or a hydride bed (for H₂). However, hydride beds are notorious for being easily poisoned by impurities requiring extremely pure H₂ to operate. The need for nitrogen for a crew of four would be about 4,380 kg/year assuming 3 parts nitrogen to one part oxygen in breathing air. The need for hydrogen for ascent propulsion would be about 1,130 kg/year to go along with 7,350 kg of oxygen.

Assuming more realistic volatile concentrations of 20 ppm, it follows that in order to produce 4,380 kg of nitrogen per year, one must process 2.19×10^8 kg/yr of regolith. Assuming solar availability at 50%, leading to reactor duty cycle ~ 40%, the requirement is for heating 63,000 kg/hr of regolith from 200 K to 800 K, at a power level of 8.8 Megawatts. Such a power rate would seem to be so high as to make volatile recovery prohibitive.

16.3.2.4 Polar Ice Deposits

There is limited evidence that suggests that recoverable near-surface deposits of water ice may exist in permanently shaded craters near the lunar South Pole. This resource has the great advantage that removal of water from regolith is a physical (rather than a chemical) process and requires far less energy and much lower temperatures. However, on the negative side, it will take a considerable investment to

locate the best deposits of ground ice (if indeed they exist – which still remains to be proven conclusively – and if they are accessible); the percentage of water ice in the regolith is likely to be low, necessitating an extensive prospecting program to find the best, most accessible deposits, ultimately requiring processing a great deal of regolith; excavating ice-filled regolith may prove difficult; the logistics of autonomous regolith delivery, water extraction, and regolith removal from a reactor may prove difficult; and the water extraction process must be carried out in dark permanently shadowed craters, necessitating use of nuclear power or less likely, beamed solar power.

It is difficult to be sure how much, if any recoverable water ice is present, and how deeply buried it is below a putative layer of desiccated regolith. The Lunar Reconnaissance Orbiter (LRO) will use a neutron spectrometer (NS) to locate hydrogen signals at much higher spatial resolution than was possible with the Lunar Prospector.

JSC is considering a rather far-fetched scheme in which regolith is excavated from a dark region of a crater, and processed in the dark to remove water (estimated at 1.5% water ice content) from the regolith. The extracted water is carried by a rover to a solar energy system located on the rim of the crater where the water is electrolyzed to hydrogen and oxygen.

The spent regolith is dumped ~ 100 m distant, and the extracted water is transported ~ 8 km to an electrolysis plant located at a rise on the crater rim where sunlight is available at a putative 70% of the time. Within the crater, all power for excavation, regolith transport, and water extraction is claimed to be nuclear, but there are no plans for installing a reactor, there couldn't possibly be enough radioisotope thermal generators (RTGs) available to supply this power, and there certainly isn't enough plutonium available to enable such RTGs to be built. Driving water across a crater surface to an electrolysis plant appears to be a grossly inefficient process. The availability of solar energy on the crater rim will depend on the morphology of the surroundings. Whether 70% availability can be achieved is presently unknown.

Overall, the required investment to do prospecting and validation of resources, and development and demonstration of regolith excavation and transport, and operation of a water extraction system, appears to be many billions of dollars. The benefit/cost ratio remains uncertain but it may take many years to "break even" on the investment.

The power requirement to heat the regolith to drive off water was estimated by Rapp (2007) to be about 18 kW to produce 10 mT of oxygen per year. This does not include the power required for excavation and hauling 800 mT/year of regolith. This power is needed in the dark. The much greater power to electrolyze the water could presumably be supplied by solar energy.

16.3.3 The Campaign for Lunar ISRU

Unfortunately, NASA has not adequately defined the campaign for prospecting, demonstrating and implementing lunar ISRU. Note: in the present context "lunar ISRU" is restricted to oxygen (and possibly hydrogen) production, mainly for ascent propellants. While JSC has plans for manufacturing spare parts on the Moon,

producing silicon solar cells on the Moon from regolith, beaming power back to Earth, and extracting parts per million of solar-wind deposited atoms, such work is (fortunately) not yet funded even though it is described in JSC project plans.

Both JSC and ESAS appear to have simplistic notions regarding requirements to prospect for polar ice resources and demonstrate ISRU systems, which will not hold up to any serious scrutiny.

A campaign is an end-to-end sequence of missions and programs to accomplish a goal (Baker et al. 2006). My view of the first five steps of the required campaign for developing lunar ISRU based on polar ice is as follows:

1. The Lunar Reconnaissance Orbiter (LRO) will use a neutron spectrometer (NS) to locate hydrogen signals at high resolution.
2. Several long-distance rovers equipped with dynamic active neutron spectrometers must be sent to several craters identified by LRO to map out local water ice deposits and estimate vertical distribution.
3. From the paragraph 2 above we can see that rovers would be sent to the most promising site equipped with a drill and excavation equipment to: (a) map out the site with NS in great detail, (b) take subsurface samples to validate rover-mounted dynamic active NS measurements of water-equivalent content, (c) determine the actual form of hydrogen-containing compounds - which are almost surely dominated by water, (d) extract water from some samples and determine the water purity and the potential need for purification, and (5) determine the soil strength and requirements for excavation of the site.
4. Develop a ~1/10 scale lunar ISRU demonstration system for use at this site, deliver it with human oversight, get it started, and leave it to operate autonomously.

16.4 Mars ISRU

16.4.1 Timeline for ISRU on Mars

Launch opportunities to send vehicles to Mars are spaced at roughly 26-month intervals. The ISRU system would be launched ~ 26 months prior to departure of the crew from LEO. The cargo delivery will take about 9 months to get to Mars and perhaps a month to set up operations on the surface. Therefore, ISRU operations could begin ~ 10 months after launch. We would then have 16 months until the crew launches, and about 22 months until the crew arrives at Mars (assuming the crew transits via a fast ~ 6-month trajectory). The full mission timeline is shown in Fig. 16.1.

The ISRU system could be sized to fill the MAV tanks in 16 months to assure that they are full prior to crew departure from Earth.

The situation for life support is less certain. The ~ 100 mT of water needed on the surface has a volume ~ 100 m³. It may be possible to store this amount of water in an inflatable tank, and let it freeze. Alternatively, it may be permissible to only extract (from regolith) and store some fraction of this during the 16 months

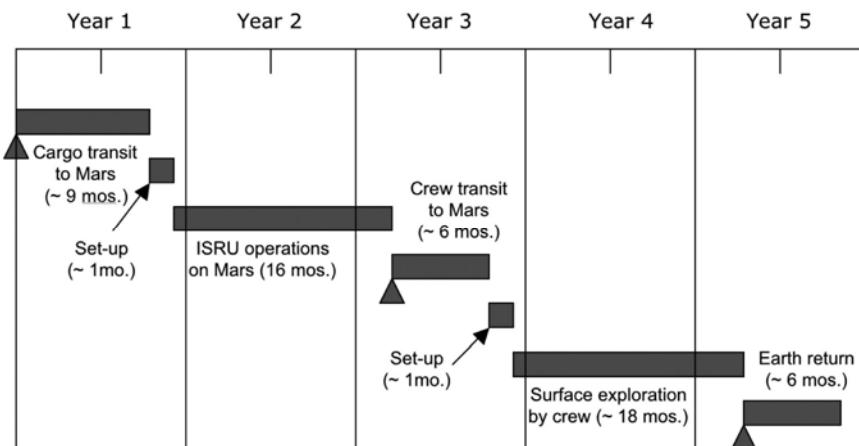


Fig. 16.1 Hypothetical timeline for a Mars mission utilizing two launch periods and ISRU. Triangles represent departure dates.

prior to crew launch. The requirements for oxygen and buffer gas (e.g. nitrogen) for a crew of 6 over 600 days are about 3.6 and 10.8 mT, respectively. One might not want to store all the needed buffer gas, and buffer gas might be recycled. As in the case of water, it will have to be decided whether to make all of the oxygen prior to human departure from Earth and then be faced with the problem of storing that large amount, or whether to be content with producing only part of the needed supply prior to crew arrival.

The decision regarding a strategy for production of propellants and consumables requires further thought regarding mass, volumetric and safety/risk considerations as well as power level. The most conservative approach would be to assume that the ISRU system would be sized so that all ascent propellants are produced in 16 months, so that the MAV tanks are full when the crew departs from Earth. The amount of life support consumables to be produced prior to crew departure from Earth remains open to question.

In regard to power, we first need to estimate power requirements for the ~18 month period during which the humans are on the Mars surface, assuming that ISRU operations stop or are significantly reduced when humans arrive. This sets the minimum power level needed for the mission. If this same power level can supply the power needed for ISRU during its 16-month operational period, then it is fair to claim that the mass of the power system is attributable only to human support, and no mass (or cost) attribution for power is made to the ISRU system. If ISRU processing continues while humans are on the surface, the additional mass (and cost) required to scale up the power system beyond human support requirements must be attributed to ISRU.

A likely scenario (though by no means the only one) is that during the 16-month period of intense ISRU operations prior to crew departure from Earth, the ISRU system will produce the requisite amount of methane and oxygen for ascent, and store these as cryogenic liquids in the tanks of the ascent vehicle. There will

be a gradual rate of boil-off of these propellants due to heat leaks into the tanks, unless zero boil-off techniques are used, with a consequent increase in mass, power and complexity. Therefore, it will be desirable to operate the ISRU plant in a greatly reduced mode while the crew is in transit to Mars (and after they land as well) to replenish boil-off. Alternatively, additional propellant tanks could be stored on Mars to hold extra propellants that could be piped to the Mars ascent vehicle prior to ascent to "top-off" the propellant tanks.

16.4.2 Mars Resources

16.4.2.1 *The Mars Atmosphere*

The atmosphere of Mars is ~ 95% CO₂, with the remainder made up mainly of Ar and N₂, and smaller amounts of CO and O₂. The CO₂ is a feedstock used to produce oxygen, and possibly hydrocarbon production if hydrogen is available. Atmospheric pressure varies with hour and season but is typically a bit less than 1/100 of the atmospheric pressure on Earth. Water has been detected in the top ~ 1 m of Mars subsurface by orbiting instruments, and is widely distributed on Mars, particularly at higher latitudes.

16.4.2.2 *Transporting Hydrogen to Mars*

Most of the studies and analyses of human missions to Mars were conducted prior to the discovery of widespread near-surface water deposits on Mars. Therefore, they depended on transporting liquid hydrogen from Earth to Mars as a feedstock to react with Martian CO₂ in ISRU processing. These included the JSC mission studies known as DRM-1 and DRM-3 and Bob Zubrin's Mars Direct.

Rapp (2007) presented a detailed review of various options for hydrogen storage. For transport to Mars, storage as a cryogenic liquid seems to be the only approach worth considering. The rate of boil-off from tanks insulated with multi-layer insulation (MLI) can be roughly estimated but this depends on the surroundings. A tank that would be earmarked for landing on Mars would undoubtedly be mounted inside an aeroshell en route to Mars. Configurations that are planned for landing on Mars would likely use a number of smaller tanks for packing density, rather than one large tank, and this would increase the ratio of area to volume compared to one large tank. In addition, the realistic insulating properties of MLI can differ significantly from those measured in the laboratory due to handling, connections and seams.

A number of papers on hydrogen storage were reviewed, and there is some diversity of opinion as to what level of boil-off may be achievable with a passive storage system. One can (at least in principle) keep adding layers of MLI, but there is bound to be a region of diminishing returns where penetrations, connections, seams, and handling factors lead to an asymptotic plateau for high numbers of layers. It appears that such a plateau may be in the range 2-7% per month boil-off in the real world, but it is difficult to fix the boil-off range any more precisely.

Because of the ~9-month trip time to get to Mars, and the fact that the hydrogen tank would be imbedded within an entry aeroshell, boil-off in transit to Mars is a

much more serious problem than for the Moon. Nevertheless, it may be possible to transport hydrogen to Mars, arriving with partly empty tanks. For a 9-month cargo transit, clearly there is large difference between 2 or 3% per month and 7% per month boil-off rates for passive storage systems. Alternatively, a zero boil-off (ZBO) system might be effective, but this would add complexity and risk. Depending on the rate of heat leak, the power requirement might be excessive.

But aside from the problem of transporting hydrogen to Mars, an even bigger problem is storing it on the surface of Mars where MLI is ineffective and other insulation systems are much less effective than MLI would be in a high vacuum. To avoid a lengthy storage time for hydrogen on Mars, Robert Zubrin has suggested an ISRU process that uses up the hydrogen rapidly via the Sabatier reaction (which is exothermic) and storing methane and water until the water can slowly be converted to oxygen via electrolysis. However, Zubrin's scheme, though ingenious, requires acquisition of a huge amount of CO₂ in a short time which requires significant power, and other volumetric and logistic challenges in which hydrogen, oxygen and methane would be moved around from tank to tank.

Acquisition of hydrogen from Martian water appears to be eminently more feasible than bringing hydrogen from Earth.

16.4.2.3 Water on Mars

Water on Mars is a broad topic that is reviewed in Appendix C of Rapp (2007). The conditions under which near-surface subsurface ice may exist in equilibrium with the atmosphere on Mars have been modeled by a number of prominent Mars scientists for forty years, and similar results have been obtained by all. The prediction is that subsurface ice is stable in the pores and interstices of Martian regolith at sufficiently high latitudes. Obviously, subsurface ice is stable at, and just below the surface in polar regions. At lower latitudes, an "ice table" forms in which a desiccated regolith covers an ice-filled layer with the depth of the ice table increasing with decreasing latitude. At some latitude near 55-60° (or perhaps as low as 45° depending on soil properties and slope), the ice table may be 1-3 meters down. At lower latitudes the depth of the ice table increases sharply and at latitudes less than typically ~ 55°, subsurface ice is not thermodynamically stable relative to sublimation to the atmosphere. These are equilibrium models and they do not preclude the possibility of non-equilibrium ice from previous epochs that is very slowly disappearing in regions where ice is not thermodynamically stable.

The Mars Odyssey neutron spectrometer has been used to scan the upper ~ 1 m layer of the Mars surface in elements 5°×5° latitude × longitude. These data support the predictions of models for latitudes > ~55°. High water concentrations are detected with apparent shallow ice tables approaching the surface toward the poles.

In the region of latitude from -45° to +45°, it is found that there is a residual water content that never drops below ~2%, probably representing chemically bound water in the minerals of the soil. In various localized areas within this region, the measured water content in the top 1 m can reach as high as 8 to 10%. Comparison of fast neutron data with epithermal neutron data suggests that there is an upper layer that is desiccated, with a higher water content layer below it. The thickness of the desiccated layer is suggested to be > 20-30 cm.

The localized equatorial regions with relatively high water content (8-10%) present an enigma. On the one hand, thermodynamic models predict that subsurface ice is not stable near the surface in the broad equatorial region. On the other hand, some aspects of the Odyssey data are suggestive of subsurface ice. It is possible that this is metastable subsurface ice left over from a previous epoch with higher obliquity. Alternatively, it could be soil heavily endowed with salts containing water of crystallization. The fact that these areas overlap somewhat with regions of high albedo and low thermal inertia suggest that it is indeed subsurface ice. Furthermore, the pixel size of Odyssey NS data is large, and the 8-10% water figure might represent small local pockets of higher water concentration (where surface properties and slopes are supportive) scattered within an arid background. Over the past million years, the obliquity, eccentricity and precession of the equinoxes of Mars has caused a variable solar input to the planet in which the relative solar input to high and low latitudes has varied considerably. It is almost certain that ground ice was transferred from polar areas to temperate areas during some of these epochs. It is possible that some of this ground ice remains today even though it is thermodynamically unstable in temperate areas. In order for remnant subsurface ice from past epochs to be a proper explanation, the process of ice deposition must be faster than the process of ice sublimation in the temperate areas over time periods of tens or hundreds of thousands of years.

We have one data point at the poles. We need ground measurements of ice content down to a few meters at latitudes in the 45-65° range to confirm the predicted ice table.

We need exploration in the 8-10% water equatorial regions to determine the state of the water in these areas because they are regions where ISRU processing is likely to be most efficient. The first step could be improved spatial resolution of orbital observations. Eventually, a landed mission is needed for ground-truth.

Deep within the interior, the temperature will rise to the point where liquid water could exist. Presently, there is no convincing evidence that it does. If liquid water exists deep within Mars (down several km), then the water vapor rising from this liquid water will pass through porous regolith at sub-freezing temperatures. Hence you cannot have liquid water at depth unless there is a huge thick layer of ice-filled regolith above it.

The crater record suggests that the interior of Mars down to several km is mainly filled with H₂O. The connection of this reservoir to near-surface H₂O has not yet been adequately investigated.

We may conclude that a near equatorial mission to Mars can chose a landing site where local water concentration in the upper meter of soil greater than 10%, and possibly much greater than 10%.

16.4.3 Mars ISRU Processes

16.4.3.1 CO₂ Acquisition from the Mars Atmosphere

All of the Mars ISRU systems that have been proposed require a supply of relatively pure, pressurized CO₂ from the atmosphere. Since the atmospheric pressure on Mars is typically about 6 torr (see Chap. 2), it is desirable to compress this by at least a

factor of ~ 100 to obtain reasonable throughput in small vessels. These ISRU systems therefore implicitly utilize a subsystem that sucks in dust-free atmosphere, separates the CO₂ from other atmospheric constituents, and compresses the CO₂. (In this process, a limited amount of Ar + N₂ may be obtained as a byproduct).

One approach for pressurizing atmospheric CO₂ is a sorption compressor that contains virtually no moving parts and achieves its compression by alternately cooling and heating a sorbent material that absorbs low pressure gas at low temperatures and drives off high pressure gas at higher temperatures (Rapp et al. 1997). By exposing the sorption compressor to the cold night environment of Mars (roughly 6 torr and 200 K at moderate latitudes), CO₂ is preferentially adsorbed from the Martian atmosphere by the sorbent material. During the day, when solar electrical power is available, the adsorbent is heated in a closed volume, thereby releasing almost pure CO₂ at significantly higher pressures for use as a feedstock in a reactor. A thermal switch isolates the sorbent bed from a radiator during the heating cycle. However, the energy required to heat up the sorbent is significant, and cooling down the sorbent overnight has been shown to be problematic. A large mass and volume of sorbent is needed.

An alternate approach for compression and purification of CO₂ was developed by a team led by Larry Clark at Lockheed-Martin that appears to be superior in that it requires less energy, less mass and less volume (Clark et al. 2001). This approach is a cyclic batch process in which the first cycle is freezing out solid CO₂ (using a mechanical cryocooler) on a cold surface while atmosphere is continuously blown over the surface. After a time, sufficient solid CO₂ builds up, and the chamber is closed off from the atmosphere. The chamber is then allowed to warm up passively, which causes the CO₂ to sublime, producing a high gas pressure in the chamber. This high-pressure CO₂ can then be vented to a larger accumulation chamber in which successive inputs of CO₂ will gradually build up the pressure. Because N₂ and Ar remain as gases at CO₂ solidification temperatures, and therefore pass out through the exit of the chamber during acquisition, relatively high-purity CO₂ is produced in this process. A Lockheed-Martin prototype test produced very encouraging results. Unfortunately, NASA does not seem to have funded any further development of this process since 2001.

16.4.3.2 Oxygen-Only Processes

Several schemes have been proposed for producing propellants from the Mars atmosphere. One approach utilizes only the CO₂ in the Mars atmosphere and produces only O₂ via the reaction



The two most developed concepts for utilizing Martian CO₂ are (1) Zirconia solid-oxide electrolysis process (SOE), and (2) the Reverse Water Gas Shift (RWGS) developed at Pioneer Astronautics.

Solid Oxide Electrolysis

Solid oxide electrolysis (SOE) is based on the very unusual and unique electrical properties of some ceramics that conduct electrical current using oxygen ions (O⁻)

as the charge carrier rather than electrons. Typically, a solid state yttria stabilized zirconia (YSZ) ion conductor is used. The doped crystal lattice contains "holes" allowing ions to move through the lattice when an electric field is applied across it. The electric field is generated by mounting porous platinum electrodes on each side of a zirconia wafer, and applying a difference in potential. In a zirconia cell, hot CO₂ is brought into contact with a catalyst on the cathode, thus causing some dissociation. Oxygen atoms in contact with the cathode pick up electrons to form O⁻ that are transported through the zirconia to form pure oxygen on the other side at the anode. YSZ has been under study for more than twenty years. The performance increases over the temperature range 800°C to 1000°C, so the materials of all cell components are critical, and sealing the edges is difficult, particularly when the cell must be repeatedly thermally cycled through many cycles. For a system that can produce, say, 4 kg/hr of O₂, the required ion current is 12,300 amps. Typical current densities on YSZ disks range from 0.3 to 0.5 amps/cm². For a value of say 0.4 amps/cm², the required area of zirconia wafer is about 30,750 cm². If a zirconia disk is as large as say, 5 cm x 5 cm square, its effective transport area is around 20 cm². This would imply that roughly 1,540 wafers of this size are needed for a full-scale unit. Thus a full-scale system will require many zirconia wafers connected in series in "stacks." Several investigators have built and tested single YSZ flat disk designs, but these cannot provide the required YSZ surface area in a small volume. A "stack" of YSZ disks is needed to produce a significant flow rate of oxygen.

The power requirement for the zirconia stack depends upon the voltage required to drive the ion current through the YSZ. It is found experimentally that as the temperature and voltage are raised, the current density (amps/cm²) increases. This allows use of less YSZ area, which leads to a more compact cell. However, as the temperature is increased, the problems of sealing and withstanding thermal cycling increase.

Relatively little work has been reported on use of YSZ stacks, but it appears that sealing problems are very challenging. Whether a workable robust multi-wafer cell can ever be produced remains doubtful.

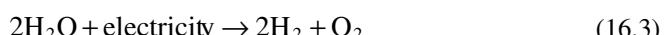
The Reverse Water Gas Shift Process

Robert Zubrin has made a number of innovations in ISRU technology. One of these is the development of the Reverse Water Gas Shift (RWGS) Process (Zubrin et al. 1997a, 1998).

The water-gas shift reaction is widely used by industry to convert relatively useless CO + H₂O into much more useful hydrogen. However, if reaction conditions are adjusted to reverse the reaction so it has the form:



then CO₂ can be converted to water. If that water is electrolyzed:



the net effect of the two reactions is conversion of CO₂ to O₂ (Reaction Eq. (16.1)). Ideally, all the hydrogen used in the first reaction is regenerated in the electrolysis reaction, so no net hydrogen is required. Actually, some hydrogen will be probably be lost if the first reaction does not go to completion, although use of a hydrogen recovery membrane can minimize this loss. The above two reactions in concert represent what is referred to as the "reverse water-gas shift" (RWGS) process.

Note that the reagents for the RWGS reaction are the same as for the S/E reaction (see next section). The main difference (aside from use of a different catalyst) is that the S/E process has a favorable equilibrium at lower temperatures (200 - 300°C) while the RWGS has a more favorable equilibrium at much higher temperatures (> 600°C). If one considers the combined equilibria where catalysts are present which allow both reactions to take place, the S/E process will be dominant below about 400°C, and the reaction products will be mainly CH₄ + 2H₂O. At temperatures above about 650°C, methane production falls off to nil and the RWGS products (CO + H₂O) are dominant. Between about 400°C and 650°C, a transition zone exists, where both reactions take place. In this zone, CO production rapidly rises as the temperature increases from 400°C to 650°C while methane production falls sharply over this temperature range. However, in the RWGS regime, no matter how high the temperature is raised, roughly half of the CO₂ and H₂ remain unreacted.

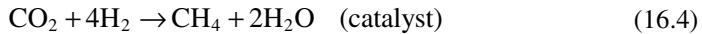
At lower temperatures (200-300°C), equilibrium would dictate that almost all the hydrogen is used up to produce methane and water, and the excess carbon dioxide is depleted by an equivalent amount. This is the S/E region. Almost no CO is formed. By contrast, at high temperatures (> 650°C) CO and H₂O are the principal products, and very little methane is formed, but roughly half of the initial carbon dioxide and hydrogen remains unreacted in the product stream. If one attempted to run the RWGS reactor at say, 400°C, Reaction Eq. (16.2) would only go about 25% to completion. Despite the unfavorable equilibrium, Zubrin and co-workers have suggested several methods to force Reaction Eq. (16.2) to the right, even at 400°C. These include:

- (i) Water condensation to water vapor pressure and recirculation of CO + CO₂ (water produced by the RWGS reaction is condensed out downstream of the reactor and resultant gases are recirculated with continuous mixing of a smaller flow of feed gases).
- (ii) Use of excess hydrogen (off-stoichiometric mixtures) to force the reaction to the right, with membrane recovery of unreacted hydrogen fed back into reactants.
- (iii) Increasing the reactor pressure.

Using these techniques, Zubrin reported high conversion efficiency in a breadboard system. It remains to be seen how efficient and practical this system will be when further developed. NASA does not seem to have funded further development of this process after about 1995. In the field of ISRU, the reward for good work seems to be a cutoff of funding.

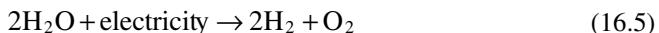
16.4.3.3 The Sabatier/Electrolysis Process

Another approach utilizes both CO₂ from the Mars atmosphere and hydrogen (brought from Earth or produced from Mars water deposits) – the Sabatier/Electrolysis (S/E) process. In the S/E process, hydrogen is reacted with compressed CO₂ in a heated chemical reactor:



The reactor is simply a tube filled with catalyst. Since the reaction is exothermic, no energy has to be supplied once the reaction is started.

The methane/water mixture is separated in a condenser, and the methane is dried, and stored for use as a propellant. The water is collected, deionized, and electrolyzed in an electrolysis cell:



The oxygen is stored for use as a propellant and the hydrogen is recirculated to the chemical reactor. Note that only 1/2 as much hydrogen is produced by reaction Eq. (16.5) as is needed for reaction Eq. (16.4), showing that an external source of hydrogen is necessary for this process to work.

It has been found experimentally that at a reactor pressure of the order of ~ 1 bar, if a mixture of CO₂ + 4H₂ enters a packed bed of catalyst, a temperature near ~300°C is high enough to approach equilibrium in a small reactor, and the equilibrium is far enough to the right that yields of over 90% CH₄ + 2H₂O are obtained. If the exit zone of the reactor is allowed to cool below 300°C, the yield can be > 95% (Clark 1997).

A problem with the S/E process when used with a hydrogen feedstock; is that it produces one molecule of methane for each molecule of oxygen, while the ideal mixture ratio for propulsion is roughly one molecule of methane for each 1.75 molecules of oxygen. Thus, there is an excess of methane for the amount of oxygen produced. This in itself is not so fundamentally bad, except that it requires that we must transport extra hydrogen to Mars to create this wasted methane. Several schemes have been proposed to recover hydrogen from excess methane, and in addition, other methods for reducing the required amount of hydrogen have been proposed by converting to higher hydrocarbons with higher C/H ratios than methane. These processes are undeveloped and are probably not needed since indigenous Mars water is likely to be used as the source of hydrogen. It has also been proposed that storage of cryogenic methane could be simplified by conversion to methyl alcohol (Zubrin et al. 1997b) but since cryogenic oxygen must be stored anyway, this appears to have only a minor benefit.

16.4.3.4 Mass and Power Requirements of a Mars ISRU System

A joint JPL-Lockheed-Martin Study in 2004-5 (Rapp et al. 2005) modeled an ISRU system that utilized the following elements:

- Water is obtained from near-surface regolith containing ~ 10% water by weight.

- CO_2 is acquired by a cryogenic freezing process.
- The Sabatier-Electrolysis process is used.
- ISRU is carried out in a 16-month cycle to produce 10 mT of CH_4 , 35 mT of propellant oxygen, 5 mT of consumable oxygen, and 108 mT of water.
- Propellant tank masses are not included in the ISRU system mass because they are needed as part of the MAV, with or without ISRU. Requirements for maintaining CH_4 and O_2 as liquids in the MAV are also attributed to the MAV, not to ISRU. However, the one-time requirement to liquefy the gases produced by ISRU is attributed to the ISRU system.

Rapp et al. (2005) estimated the mass and power requirements for various steps as shown in Table 16.3.

Table 16.3 Estimated gross mass and power requirements for various Mars ISRU process steps

Process Step	Total feedstock used in 16 mos.	Feed-stock rate (kg/hr)	Mass of unit (kg)	Power required (kW)
H_2O acquisition	22,500 kg H_2O	2.0	3000	24
CO_2 acquisition	27,500 kg CO_2	2.4	120	2.9
Sabatier conversion	27,500 kg CO_2	2.4	30	0.4
Water electrolysis	33,750 kg H_2O^*	2.9	33	7.0
Liquefying O_2	35,000 kg O_2	3.0	105	3.3
Liquefying CH_4	10,000 kg CH_4	0.87	74	2.6
TOTAL			3360	40.2

In addition to 22,500 kg of water feedstock, an additional 11,250 kg of recycled product water from the Sabatier process must be electrolyzed

Allowing for inefficiencies, we should probably increase the figures in this table by at least 15-20%.

It should be noted that we have assumed that the ISRU system functions 24/7 for sixteen months and thus a relatively small system produces a huge amount of product. For example, regolith excavation is only about 20 kg per hour (or 10 liters of volume per hour) to produce 2 kg/hr of water. Note that the ISRU system mass is dominated by excavation systems.

16.4.4 Summary

ISRU on Mars has significant advantages compared to lunar ISRU:

- The gear ratio for delivery of assets from LEO to the Mars surface is higher than for the lunar surface, thus requiring more IMLEO per unit mass delivered to the surface of Mars. This, in turn, makes mass replacement on Mars by ISRU more valuable than mass replacement by ISRU on the Moon.
- The Δv for ascent from the Mars surface to orbit is much greater than Δv for ascent in the lunar case, necessitating much greater propellant requirements for

ascent. This, in turn, makes propellant production on Mars by ISRU more valuable than propellant production by ISRU on the Moon.

- By placing the Earth Return Vehicle (ERV) in an elongated elliptical orbit (in the Mars case) one can increase the required Δv for ascent (and thereby the amount of ascent propellants supplied by ISRU) while decreasing the Δv requirements for orbit insertion of the ERV (propellants supplied from Earth) as well as for Earth return from Mars orbit. As it turns out, the mass savings by allowing the ERV to utilize an elliptical orbit are even greater than the mass saving due to reduced mass delivered to the Mars surface.
- The combination of the previous three points provides much greater mission impact (IMLEO reduction) for in situ production of ascent propellants on Mars than on the Moon.
- Because of the long round trip to Mars (~ 2.7 years), total consumption of life support consumables amounts to perhaps 200 mT. It is not clear whether an ECLSS system will have the longevity to provide fail-safe performance over that time period. Such an ECLSS system to provide air and water (if it is feasible) is likely to weigh > 30 mT.
- Unlike the Moon, Mars has a ready supply of carbon and oxygen in the easily acquired atmosphere.
- Unlike the Moon, Mars has significant near-surface deposits of water (believed to be mainly in the form of ground ice) widespread across much of the planet.
- The combination of atmospheric CO₂ and water from regolith provides feedstocks on Mars that enable proven, relatively simple Sabatier-electrolysis processing to produce methane and oxygen propellants, and water for life support.
- In conclusion, Mars ISRU is far more easily implemented and has far more mission impact than lunar ISRU.

The major unknowns regarding Mars ISRU are:

- What are the requirements for excavating water-bearing near-surface regolith and extracting water?
- In the case of equatorial water-bearing near-surface regolith, is the water in the form of ground ice or mineral hydrates?

Unfortunately, neither ESAS nor the Mars Exploration (Science) Program appear to have any specific plans to investigate these questions.

References

- Baker, E., Morse, E.L., Gray, A., Easter, R.: Architecting Space Exploration Campaigns: A Decision-Analytic Approach. In: IEEE Aerospace Conference, March 4-11, IEEEAC paper #1176 (2006)
- Clark, D.L.: In-Situ Propellant Production on Mars: A Sabatier/Electrolysis Demonstration Plant. In: Situ Resource Utilization (ISRU) Technical Interchange Meeting, February 4-5, AIAA-97-2764 (1997)
- Clark, D.L., Payne, K.S., Trevathan, J.R.: Carbon Dioxide Collection And Purification System For Mars. In: AIAA Space 2001, Conference and Exposition, Albuquerque, NM, August 28-30, AIAA Paper 2001-4660 (2001)

- Connolly, J.: NASA Watch, Kicking up Some Dust (2007), http://www.nasawatch.com/archives/2007/03/interesting_lun.html
- Hirata, C., Greenham, J., Brown, N., Shannon, D., Burke, J.D.: A New Plan for Sending Humans to Mars: The Mars Society Mission, Informal Report, Jet Propulsion Laboratory, California Institute of Technology (1999),
<http://www.lpi.usra.edu/publications/reports/CB-979/caltech99.pdf>
- Rapp, D.: Human Missions to Mars. Praxis Publishing, Chichester (2007)
- Rapp, D., Karlmann, P., Clark, D.L., Carr, C.M.: Adsorption Pump for Acquisition and Compression of Atmospheric CO₂ on Mars. In: Situ Resource Utilization (ISRU) Technical Interchange Meeting, February 4-5, 1997, AIAA 97-2763 (1997)
- Rapp, D., Andringa, J., Easter, R., Smith, J.H., Wilson, T., Clark, L., Payne, K.: Preliminary System Analysis of Mars ISRU Alternatives. In: IEEE Aerospace Conference, Huntley Lodge, Big Sky, Montana, 5-12 March (2005)
- Taylor, C., Broniatowski, D., Boas, R., Silver, M., Crawley, E., De Wec, O., Hoffman, J.: Paradigm Shift in Design for NASA's Space Exploration Initiative: Results from MIT's Spring 2004 Study. In: 1st Space Exploration Conference: Continuing the Voyage of Discovery, Orlando, Florida, January 30-1 (2005)
- Wooster, P.: From Value to Architecture: The Exploration System of Systems, Presentation at JPL, August 23 (2005)
- Wooster, P., Hofstetter, W.K., Nadir, W.D., Crawley, E.F.: The Mars-Back Approach: Affordable And Sustainable Exploration of the Moon, Mars, and Beyond Using Common Systems, International Astronautical Congress, October 17-21 (2005),
http://smartech.gatech.edu/dspace/bitstream/1853/8043/2/SSEC_SE1_ppt.pdf
- Zubrin, R.M.: The Mars Direct Plan, Scientific American (March 2000)
- Zubrin, R.M., Baker, D.A., Gwynne, O.: Mars Direct: A Simple, Robust, and Cost Effective Architecture for the Space Exploration Initiative, AIAA-91-0328 (1991),
http://www.marssociety.org/portal/TMS_Library/Zubrin_1991
- Zubrin, R., Frankie, B., Kito, T.: Mars In-Situ Resource Utilization Based on the Reverse Water Gas Shift. In: 33rd AIAA/ASME Joint Propulsion Conference, Seattle, WA, July 6-9, AIAA-97-2767 (1997a)
- Zubrin, R., Kito, T., Frankie, B.: Report on the Construction and Operation of a Mars Methanol in situ Propellant Production Unit Pioneer Astronautics Report, Pioneer Astronautics, Lakewood, CO (1997b)
- Zubrin, R., Frankie, B., Kito, T.: Report on the Construction and Operation of a Mars in situ propellant Production Unit Utilizing the Reverse Water Gas Shift. In: 34th AIAA/ASEE Joint Propulsion Conference, July 13-15, Cleveland Ohio, AIAA-98-3303 (1998)

Chapter 17

Available Resources and Energy Sources from Mars Rock and Soil

Yuuki Yazawa¹, Takashi Mikouchi², and Hiroshi Takeda^{1,2}

¹ Chiba Institute of Technology, Chiba, Japan

² University of Tokyo, Japan

17.1 Introduction

The surface of the planet Mars exhibits a world of reddish rock, soil, and dust as revealed by Mars landers (Fig. 17.1). Such an environment is quite distinct from the earth except for the desert areas and the big question is whether we are able to find or produce soils suitable for agriculture on Mars. If we send astronauts to Mars, they will have to stay for 18-24 months.

Therefore, considering agriculture is important to living off the land, recycle air, water, provide food etc as reviewed by Ming and Henninger (1989). In this chapter we review the information about the chemistry and mineralogy of the rocks and soils present on the surface of Mars and then discuss available resources and energy sources that can be utilized to perform Martian agriculture. There is also discussion of using released methane.

Fig. 17.1 Husband Hill in Gusev crater as imaged by the Spirit Mars Exploration Rover. Image by NASA (PIA09090).



17.2 Rocks and Soils on Mars

There are mainly two available sources to obtain information about rocks and soils on Mars. One is from direct analyses of meteorites originating from Mars that

were ejected by heavy impacts that eventually reached the earth. The other is from data collected by spacecraft exploration of Mars. Both sources have provided valuable insight into the geochemistry and mineralogy of Mars surface materials.

17.2.1 Martian Meteorites

17.2.1.1 Evidence for a Martian origin

At first we will introduce why we could recognize the presence of Martian meteorites by tracking the historical records related to Martian meteorites. The evidence for a Martian origin for certain group of meteorites is related to the characteristic natures of Mars (e.g., McSween 2002) and thus important for the following chapter discussing the possibility of performing Martian agriculture.

We have more than 50 individual meteorites at present that are believed to have originated from the planet Mars (Table 17.1). The total number of meteorites is now exceeding 40,000 and still growing due to recovery of many meteorites from Antarctica and hot deserts. However, Martian meteorites constitute only 0.1–0.2 % of the known meteorites (Meyer 2008). The recognition of Martian meteorites was mainly thanks to the discovery of many Antarctic meteorites starting in 1977.

Table 17.1 List of Martian meteorites (January 2009)

Name	Location	Year	Weight (g)	Class
Shergotty	India	1865	~5,000	Basaltic shergottite
Zagami	Nigeria	1962	~18,000	Basaltic shergottite
EETA 79001	Antarctica	1980	7,900	Basaltic shergottite
QUE 94201	Antarctica	1994	12	Basaltic shergottite
Los Angeles	USA	1999	698	Basaltic shergottite
Dhofar 378	Oman	2000	15	Basaltic shergottite
Northwest Africa 480/1460	Morocco	2000-2001	98	Basaltic shergottite
Northwest Africa 1669	Morocco	2001	36	Basaltic shergottite
Northwest Africa 856	Morocco	2001	320	Basaltic shergottite
Northwest Africa 5029	Morocco	2003	15	Basaltic shergottite
Northwest Africa 3171	Algeria	2004	506	Basaltic shergottite
Northwest Africa 2975	Morocco	2005	70	Basaltic shergottite
Northwest Africa 2800	Morocco	2007	686	Basaltic shergottite
Northwest Africa 4864	Northwest Africa	2007	94	Basaltic shergottite
Northwest Africa 5298	Northwest Africa	2008	445	Basaltic shergottite
ALH 77005	Antarctica	1977	482	Lherzolitic shergottite
Yamato 793605	Antarctica	1979	16	Lherzolitic shergottite
LEW 88516	Antarctica	1988	13	Lherzolitic shergottite
GRV 99027	Antarctica	2000	10	Lherzolitic shergottite
Yamato 000027/47/97	Antarctica	2000	38	Lherzolitic shergottite
Northwest Africa 1950	Morocco	2001	815	Lherzolitic shergottite

Table 17.1 (continued)

GRV020090	Antarctica	2000	10	Lherzolitic shergottite
Northwest Africa 2646	Algeria /Morocco	2005	9	Lherzolitic shergottite
Northwest Africa 4468	Western Sahara	2006	675	Lherzolitic shergottite
RBT04261/04262	Antarctica	2007	79	Lherzolitic shergottite
Northwest Africa 4797	Morocco	2007	15	Lherzolitic shergottite
Dar al Gani 476 + 6 others	Libya	1996-1999	10,414	Olivine-phyric shergottite
Yamato 980459	Antarctica	1998	82	Olivine-phyric shergottite
Sayh al Uhaymir 005 + 10 others	Oman	1999-2004	11,390	Olivine-phyric shergottite
Dhofar 019	Oman	2000	1,056	Olivine-phyric shergottite
Northwest Africa 1068 + 3 others	Morocco	2001-2004	797	Olivine-phyric shergottite
Northwest Africa 1195	Morocco	2002	315	Olivine-phyric shergottite
Northwest Africa 2046	Algeria	2003	63	Olivine-phyric shergottite
Northwest Africa 2626	Algeria	2004	31	Olivine-phyric shergottite
Northwest Africa 4480	Morocco	2006	13	Olivine-phyric shergottite
Northwest Africa 4527	Algeria	2006	10	Olivine-phyric shergottite
Northwest Africa 4222	Northwest Africa	2006	17	Olivine-phyric shergottite
LAR06319	Antarctica	2006	79	Olivine-phyric shergottite
Northwest Africa 4925	Morocco	2007	282	Olivine-phyric shergottite
Nakhla	Egypt	1911	~10,000	Nakhlite
Lafayette	USA	1931	~800	Nakhlite
Governador Valadares	Brazil	1958	158	Nakhlite
Yamato 000593/749/802	Antarctica	2000	15,020	Nakhlite
Northwest Africa 817	Morocco	2000	104	Nakhlite
Northwest Africa 998	Algeria / Morocco	2001	456	Nakhlite
MIL 03346	Antarctica	2003	715	Nakhlite
Chassigny	France	1815	~4,000	Chassignite
Northwest Africa 2737	Morocco	2000	611	Chassignite
ALH 84001	Antarctica	1984	1,940	Orthopyroxenite

Although it is common to use the term “Martian” meteorites these days, the name of “SNC” meteorites used to be more common. “SNC” was taken from “S”hergottites, “N”akhlites, and “C”hassignite. SNC meteorites are igneous rocks that crystallized from molten magma. They are distinct from other meteorite groups and share the same oxygen isotope compositions (Clayton and Mayeda 1983). Therefore, SNC meteorites were believed to share a common parent body.

The most unique characteristic of SNC meteorites was their young crystallization ages. Shergottites gave Rb-Sr and Sm-Nd ages of about 180 Ma, while nakhlites and chassignite gave 1.3 Ga (e.g., Nakamura et al. 1982). These ages are distinct from those of other basaltic achondrites (e.g., HED meteorites) that showed early differentiation and the formation of evolved planetesimals in the

early solar system (4.56 Ga) (e.g., Takeda 1997). These asteroids could not produce magmas since 4.4–4.5 Ga because they lost internal heat quickly because of their small sizes. The young crystallization ages thus show that the parent body of SNC meteorites sustained enough heat to produce igneous activity until recently, suggesting that the parent body is large.

SNC meteorites are also distinct from normal basaltic achondrites in the presence of aqueous alteration. Basaltic achondrites are free from aqueous alteration because their parent bodies did not keep water after the igneous activity. Although it is minor, we could find the presence of “iddingsite” or “rust” in nakhellites. These iddingsites show melting when they are in contact with fusion crust, thus showing that they are pre-terrestrial (e.g., Gooding et al. 1991). Because these alteration phases are present along fractures of igneous minerals, it is suggested that small amounts of liquid percolated into the fractures to deposit alteration products. Although amphibole is one of the most major rock-forming minerals on the earth, it rarely occurs in meteorites. This is mainly because meteorites were formed under low-pressure and anhydrous conditions, where amphibole is not stable. Kaersutite amphibole is present in shergottites and chassignites as magmatic inclusions, showing that their parent magmas contain water and formed at depth (Johnson et al. 1991).

SNC meteorites were known that they formed under relatively oxidizing conditions. SNC meteorites contain ulvöspinel and magnetite in which significant amount of Fe^{3+} is present. Kaersutite also contains Fe^{3+} . The oxygen fugacity of the nakhelite magma was estimated to be near the quartz-fayalite-magnetite (QFM) buffer (Reid and Bunch 1975). This is much more oxidizing than basaltic achondrites that contain metallic iron because they formed under relatively reducing conditions below the iron-wüstite (IW) buffer.

Another characteristic features of SNC meteorites include their heavy shock metamorphism such as maskelynitization of plagioclase and the presence of shock melts (e.g., Stöffler et al. 1986). If they came from a large parent body, the presence of strong shock metamorphism is consistent with the ejection by large impact events. SNC meteorites showed highly fractionated rare earth element (REE) patterns that imply a large parent body with substantial internal pressures (Nakamura et al. 1982).

Thus, by combining these observations it is suggested that SNC meteorites came from a large body with oxidizing and hydrous conditions. Mars is the best candidate explaining all these characteristics. However, these pieces of evidence are all indirect and it was not possible to conclude that the parent body was Mars in early 1980s because no meteorites from the Moon were recognized.

Since 1970s, the number of meteorites drastically increased by the discovery of Antarctic meteorites. EETA79001 was a new shergottite among them, which was important because it trapped unique gas in shock melt (Fig. 17.2). The trapped gas composition was turned out to be identical to the atmospheric gas analyzed by the Viking lander (Bogard and Johnson 1983). This is the direct evidence that connected SNC meteorites to Mars. The trapped gas also yields high D/H ratios of 2,500–3,000 ‰ characteristic of Martian atmosphere (Boctor et al. 2003). It was fortunate that a lunar highland breccia was discovered in Antarctica in 1983. This



Fig. 17.2 Cut surface of the EETA79001 Martian meteorite. Dark spots on the surface are shock melt pockets containing the Mars atmosphere. Image by NASA.

discovery of the first lunar meteorite showed that rocks could be ejected from large bodies and then reached the earth, which brought about reasonable consideration for the presence of meteorites from Mars. There has been discussion for the mechanism to eject rocks from the surface of Mars (e.g., Wasson and Whetherill 1979), but Melosh (1984) modeled that shock waves from large impacts could accelerate fragments near the surface to the Martian escape velocity (5 km/sec).

17.2.1.2 Martian Meteorites: Primary Igneous Mineralogy

The number of Martian meteorites is now over 50 (Table 17.1), which are generally represented by the three groups (shergottites, nakhlites and chassignites). However an additional group was added that is represented by ALH84001 which has much older crystallization age (4.55 Ga) than the SNC meteorites (e.g., Mittlefehldt 1994).

Shergottites

There are three distinct subgroups of shergottites. Basaltic shergottite are “classic” shergottites (e.g., Shergotty and Zagami) and are composed of pyroxene and maskelynite but with no olivine (e.g., Stolper and McSween 1979). The second subgroup of shergottite is Iherzolitic shergottite. ALH77005 is the first known sample of this subgroup (e.g., McSween et al. 1979). They show a poikilitic texture, and are mainly composed of pyroxene, olivine and maskelynite. The third group is

olivine-phyric shergottite that is mainly composed of pyroxene and maskelynite with large olivine phenocrysts reaching several millimeters (e.g., Goodrich 2002).

Although basaltic shergottites show simple mineralogy composed of pyroxene and maskelynite, their mineral compositions and crystallization ages show wide ranges (e.g., Mikouchi et al. 1999; Nyquist et al. 2001). Pyroxenes in basaltic shergottites are clinopyroxene. In Shergotty and Zagami, pyroxene is present as two separate grains of pigeonite and augite, each showing chemical zoning from Mg-rich cores to Fe-rich rims. In QUE94201, clinopyroxene shows complex chemical zoning from Mg-rich pigeonite cores to Fe-rich pigeonite rims via augite mantle (e.g., Mikouchi et al. 1999). Chemical zoning of pyroxenes in all known basaltic shergottites can be divided into these two types (Meyer 2008). Maskelynite in basaltic shergottites shows weak chemical zoning typically ranging An_{60-40} . These plagioclase compositions are clearly more Na-rich than those in basaltic achondrites and lunar rocks. Because of the basaltic texture and the presence of extensive chemical zoning, basaltic shergottites are likely to have crystallized near the Martian surface as lava flows or shallow intrusion (e.g., Mikouchi et al. 1999).

Lherzolitic shergottites are characterized by the presence of poikilitic textures (e.g., Mikouchi and Kurihara 2008). Large oikocrysts of pyroxene enclose cumulus Mg-rich olivine and chromite. Lherzolitic shergottites also contain a volumetrically minor non-poikilitic basaltic texture that is interstitial to oikocrysts. Olivine compositions are usually homogeneous in each sample, but show a small variation ranging Fo_{75-60} . Pyroxenes are present as both low-Ca pyroxene and high-Ca pyroxene. The core composition of the pyroxene oikocrysts is orthopyroxene ($\text{En}_{80}\text{Wo}_3$). Maskelynites are exclusively present in non-poikilitic areas, and their compositions are An_{55-40} . These rocks possibly crystallized from a melt parental to those from which basaltic shergottites had crystallized or they are peridotite type samples from which parent magmas of basaltic shergottites were derived by partial melting (e.g., McSween et al. 1979). Although the name of "lherzolite" suggests a plutonic origin, the chemical zoning of pyroxene shows that the final solidification occurred near the surface. Because almost all known lherzolitic shergottites show similar mineralogy and petrology with identical crystallization (1.8 Ga) and cosmic-ray exposure (3.5-4 Ma) ages, it is widely considered that they share the same original source on Mars, were probably ejected by the same impact, and fell on earth as separate falls (e.g., Mikouchi and Kurihara 2008).

Olivine-phyric shergottite possesses remarkable textures composed of olivine megacrysts set in the groundmass of pyroxene and maskelynite (Goodrich 2002). Olivine shows extensive chemical zoning in each sample and the composition covers a wide range from Fo_{85} to Fo_{25} . Pyroxene composition is also having a wide range although the most Mg-rich pyroxene has an orthopyroxene composition ($\text{En}_{80}\text{Wo}_2$). Maskelynite shows weak chemical zoning in each sample and its composition ranges An_{65-45} . Most olivine-phyric shergottites are considered that they directly crystallized from magmas without significant cumulate processes, and solidified near the Martian surface. For example, Y980459 shows a glassy texture and the cooling experiment using its bulk composition shows that it crystallized by ~ 1 °C/hour cooling rate (Mikouchi et al. 2004), probably in the lava flow erupted on the Martian surface. In some cases, olivine-phyric

shergottites contain large prismatic orthopyroxene crystals, and they are called olivine-orthopyroxene-phyric shergottites.

Although shergottites are divided into three subgroups based on petrography, recent studies have revealed that shergottites can be also divided into three subgroups (enriched, depleted and intermediate) based on geochemistry such as trace element and isotopic compositions (e.g., Borg and Draper 2003). There is no relationship between petrographic and chemical classifications, and both basaltic and olivine-phyric subgroups include samples divided into three chemically classified ones. In contrast, Iherzolitic shergottites show nearly identical mineralogy and ages within the subgroup, and only chemically intermediate samples are known (e.g., Mikouchi and Kurihara 2008). Thus, shergottites show a wide range of major and trace element compositions and initial Sr and Nd isotopic compositions that have been interpreted to reflect either geochemical variation inherited from their mantle source regions or varying degrees of assimilation of Martian crust (e.g., Borg and Draper 2003). The crystallization ages of shergottites range 180–440 Ma (e.g., Nyquist et al. 2001), suggesting that the igneous activity lasted (or still lasts) for a long period. Because the surface having such young volcanic activities are only found in Tharsis and Erisium areas, shergottites probably originated from these areas.

Nakhrites

Nakhrites are augite-rich cumulate rocks with variable amounts of olivine and groundmass plus minor Fe, Ti oxides (e.g., Treiman 2005). Pyroxene composition is nearly homogeneous ($\text{En}_{40}\text{Wo}_{40}$) except for the thin Fe-rich rims adjacent to the groundmass. Olivine shows chemical zoning (Fo_{45-5}) except for Lafayette and NWA998. The groundmass is mainly composed of feldspar (plagioclase: $\text{An}_{25}\text{Or}_5$, alkali feldspar: $\text{An}_5\text{Or}_{70}$) and titanomagnetite although some samples show nearly glassy nature (mesostasis). Nakhrites show correlated petrography and mineralogy that could be explained by different locations (burial depths) in a common cooling cumulate pile (e.g., Mikouchi et al. 2003) (Fig. 17.3).

Chemical zoning of nakhrite olivines is useful to estimate their cooling rates. Mikouchi et al. (2003) estimated cooling rates of olivine and demonstrated that burial depths ranged from 1 m for MIL03346 and NWA817 to >30 m for Lafayette and NWA998. This order is generally related to mineralogical characteristics of each sample. If we consider that gravitational settling of crystallizing grains played a significant role in the nakhrite igneous body, cumulus phases should have been more densely packed in deeper areas. This could produce loose cumulus framework for shallower samples, forming abundant groundmass or mesostasis areas. NWA817 and MIL03346 have higher mesostasis abundances and thus appear that they are shallow samples. In fact, the obtained burial depths for these two samples are shallowest among nakhrites, and glassy nature of mesostasis is also consistent with fast cooling of the intercumulus melt. In contrast, the other samples contain crystalline feldspar grains consistent with slow cooling at depth. Especially, the presence of large blocky feldspar in NWA998 is distinct from other samples. NWA998 is likely to represent the deepest sample among nakhrites, even deeper than Lafayette.

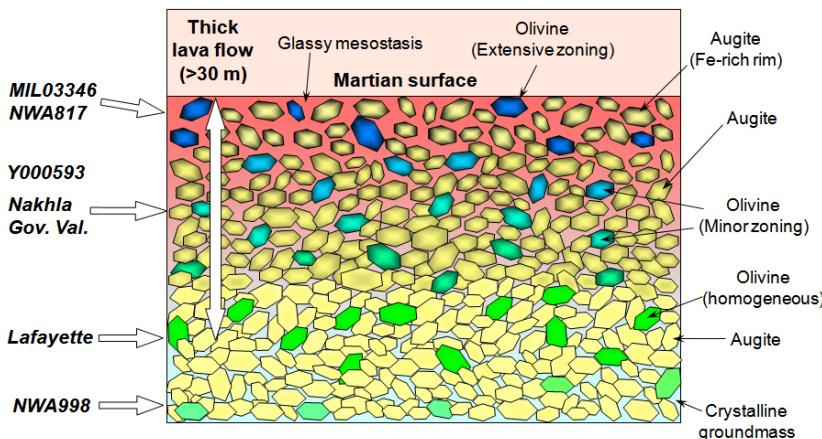


Fig. 17.3 Schematic illustration of the nakhelite igneous body at the time of final solidification.

Chassignites

There are two chassignites known, Chassigny and NWA2737 which have similar mineral compositions (olivine: Fo_{80-70}), with differing shock degrees. NWA2737 shows a higher shock degree, and olivine shows a dark color due to precipitation of Fe-Ni metal nano-particles (e.g., Van de Moortèle et al. 2007). Kaersutite is present in magmatic inclusions in olivine. Johnson et al. (1991) reported that prior to kaersutite crystallization, the melt must have contained at least 4 wt% dissolved water, suggesting that a minimum of 1.5 kbar total pressure. Although the initial crystallization of chassignites appeared to have taken place at depth, Ca in olivine shows chemical zoning that corresponds to the burial depth of ~15 m. Probably, chassignites record two-stage crystallization history, but their formation may be similar to nakhellites where cumulus phases were mostly olivine rather than augite.

ALH84001

ALH84001 is the only Martian meteorite with old initial crystallization age (4.55 Ga) and well known for evidence for possible biogenic activity (McKay et al. 1996), especially in the carbonates. The issue of biogenic activity is of great philosophical and scientific importance, and yet remains controversial. ALH84001 is an igneous orthopyroxene and chromite cumulate with interstitial plagioclase with evidence for multiple shock and alteration events (e.g., Treiman 2003). It is a brecciated rock, but orthopyroxene constitutes >90%. The pyroxene composition is homogenous ($\text{En}_{70}\text{Wo}_{3}$). Minor amounts of olivine (Fo_{35}) and maskelynite (An_{35}) occur.

17.2.1.3 Martian Meteorites: Secondary Alteration Mineralogy

As demonstrated above, all Martian meteorites are igneous rocks and there is little evidence for secondary alteration. However, careful examination reveals minor evidence for secondary alteration in almost all Martian meteorites. Secondary

alteration products generally formed by interactions of the igneous rocks with solutions percolating near the Martian surface. The solutions dissolved, leached and mobilized soluble mineral constituents in rocks and deposited the salt assemblages by evaporation and desiccation. These reactions occurred either *in situ* or at different locations depending on the fluid-rock ratio in the system. The presence of such alteration products is important to understand the surface condition of recent (or present) Martian environment that is crucial to consider the possibility of Martian agriculture.

Shergottites

Although the abundance is extremely low, secondary alteration minerals are reported from some shergottites that are likely to have formed by reaction with aqueous fluids. EETA79001 has been most extensively studied in terms of secondary alteration because it contains gas-rich impact melt (GRIM) glasses where secondary minerals are present (e.g., Gooding et al. 1988) (Fig. 17.2). The secondary mineral assemblages present in the GRIM glasses consist of salts such as calcite, gypsum, S-Cl-rich alumino-silicates, Fe-sulfate and Mg-phosphate that are mixed with other altered silicates. They were presumably trapped from porous regolith materials during glass formation by impact-melting near the Martian surface (e.g., Rao et al. 2008). The isotopic compositions of C and O extracted from carbonates in EETA79001 are distinct from terrestrial carbonates (Wright et al. 1988). The water extracted from shergottites has oxygen isotopic compositions that are distinct from those of igneous silicates, suggesting that the lithosphere and hydrosphere of Mars formed two distinct oxygen isotopic reservoirs that were isolated from each other because of the absence of plate tectonics (Karlsson et al. 1992). Sulfates, alumino-silicate clays and secondary silica were also found in QUE94201 (Wentworth and Gooding 1996). Na-, Ca-, and K-Fe-sulfates, low-Al silicate clay, silica, and Mg-Fe-phosphate are reported from Iherzolitic shergottites (Wentworth and Gooding 1993). However, the alteration may have occurred during the residence in Antarctica because some Antarctic eucrites contain secondary minerals such as Ca-sulfate, Na-sulfate, and K-Fe sulfate associated with other silicate minerals (Gooding 1986).

Nakhlites

Nakhlites contain more abundant alteration materials than shergottites, although the abundance is still low. Almost all nakhlites show evidence for secondary alteration as is suggested by brown color of olivine and mesostasis. “Iddingsite”, carbonates, halites, Mg-sulfate and clays occur filling fractures in olivines and mesostasis, suggesting that they were produced *in situ* by aqueous alteration that occurred on Mars (e.g., Treiman 2005). Among these alteration products, “iddingsite” is most abundant that are rusty brown veinlets and patches present along fractures and grain boundaries of olivine and mesostasis (Fig. 17.4).

Iddingsite is a mixture of poorly crystalline Fe-rich smectite and iron oxides including ferrihydrite and magnetite. The stable isotopic and noble gas studies showed that the low-temperature hydrous alteration products are of Martian origin (e.g., Karlsson et al. 1992). S and Cl abundances determined in the Nakhla

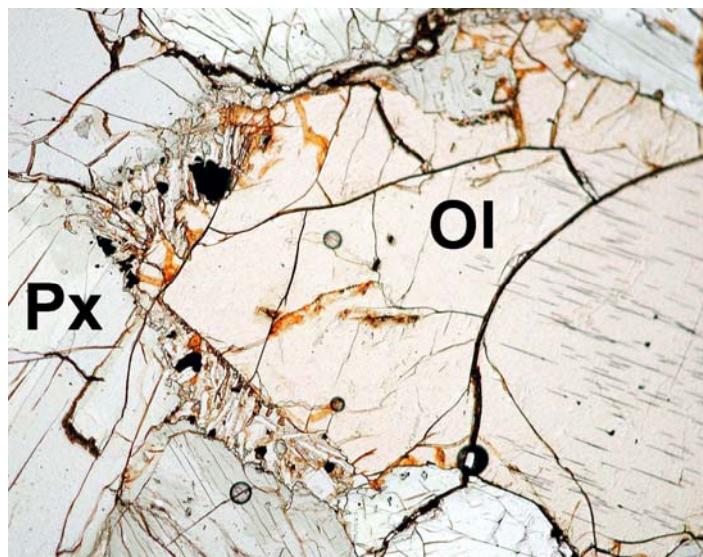


Fig. 17.4 Optical photomicrograph of the Y000593 nakhelite. Note the presence of brown alteration products (“iddingsite”) along fractures and the grain boundary of olivine. The width of the image is about 1 mm. Ol: olivine. Px: pyroxene.

salt-assemblages suggested that the salts in nakhellites are likely formed from relatively Cl-rich and sulfate-poor solutions on Mars. As these salts occur associated with carbonates, neutral to alkaline conditions seem to have prevailed during the salt deposition (Wentworth and Gooding 1996). In Y000593 nakhelite, poorly crystalline material and jarosite were found as alteration products of olivine (Noguchi et al. 2003). Misawa et al. (2005) obtained the Rb-Sr age of 650 Ma for these alteration products that are in good agreement with K-Ar and Rb-Sr ages of Lafayette iddingsite (Swindle et al. 2000). Because the crystallization ages of nakhellites are 1.3 Ga, these younger ages may suggest later aqueous alteration near the surface of Mars as recently as 650 Ma ago.

Chassignites

Chassignite shows only minor evidence for secondary alteration. Calcite, Mg-sulfate, and Mg-carbonate were found in Chassigny, but no clays or ferric oxides were reported (Wentworth and Gooding 1994).

ALH84001

ALH84001 is an orthopyroxenite of igneous origin that contains ~1% of carbonates in which evidence for possible biogenic activity has been reported (McKay et al. 1996). Subsequent to its igneous history, the sample suffered mechanical granulation along sheared zones (probably from shock), possible recrystallization of the granular material, and at least one shock event after the recrystallization. The plagioclase was melted and quenched to glass. At some point during these events, carbonates were introduced into the sample. Carbonates in ALH84001 are

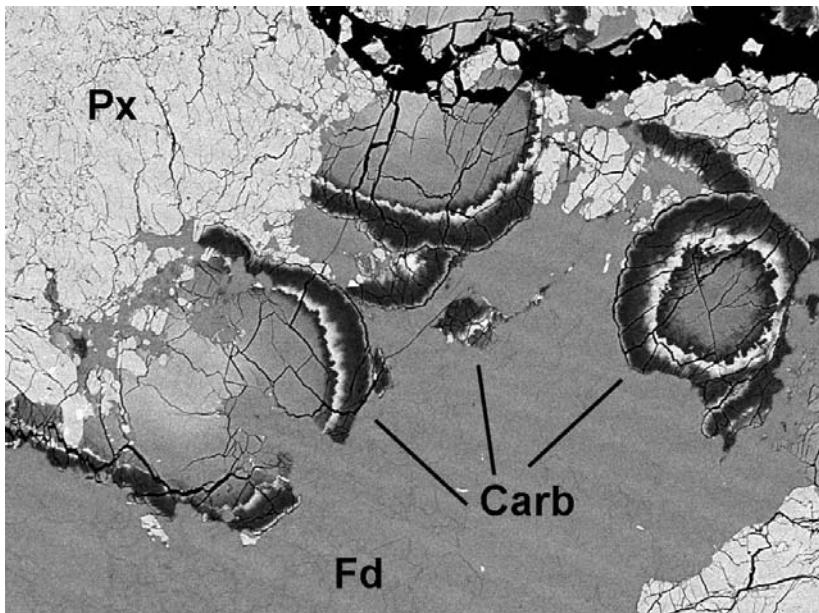


Fig. 17.5 Back-scattered electron image of ALH84001 Martian meteorite by scanning electron microscope. Carbonates show spherical shapes with complex chemical zoning. The width of the image is about 0.5 mm. Carb: carbonate. Px: pyroxene. Fd: feldspathic glass.

up to a few hundreds μm in size and exhibit complex chemical zoning from core to rim (Fig. 17.5).

Although they sometimes occur singly, globules are more often present as clusters or composite globules. The typical globule has an ankeritic dolomite core that changes rapidly but continuously to ferroan magnesite. The core is zoned towards pure magnesite and renewed growth is markedly Fe-rich in a zone including magnetite, a sulfide phase, and Fe/Mg carbonate. This zone ends abruptly in a transition back to nearly pure magnesite and quite thin, Fe-rich zone is present at the edge. The carbonates occur in association with some submicron magnetite grains, organic matter, and submicron objects shaped like bacteria that were interpreted as biogenic (McKay et al. 1996; Thomas-Keprta et al. 2000).

However, inorganic processes such as evaporite or shock and thermal metamorphism (e.g., Treiman 2003) may also explain these features. Golden et al. (2004) experimentally showed that zoned carbonate globules could be formed from carbonate-rich hydrothermal solutions and then carbonates could be decomposed to form magnetite by a later impact event. The ^{39}Ar - ^{40}Ar system shows complete resetting at 4.0 Ga (Bogard and Garrison 1999). Because the Rb-Sr age of the carbonates is indistinguishable from the 4.0-Ga ^{39}Ar - ^{40}Ar age (Borg et al. 1999), the carbonate formation occurred immediately after the intense shock event.

17.2.2 Mars Exploration

Mars meteorites have provided valuable insight into geochemical and mineralogical nature of Mars rock. Another information about the Mars surface science has been obtained by earth-based telescopic observations (e.g., Ehlman et al. 2008). However, it is obvious that the majority of the knowledge has been obtained through Mars explorations. For over 40 years, orbital and landed missions to Mars have provided useful information about the mineralogy and chemistry of the Martian surface material which enables a better understanding of the geological history from the wet past to the dry present.

17.2.2.1 Orbiters

NASA's Mars Global Surveyor (MGS) was the first successful orbiter mission that provided detailed information about the mineralogy and chemistry of the Martian surface. MGS was launched in 1996 and continued working until November 2006 when the orbiter stopped communicating with the earth. Thermal infrared spectra from the Thermal Emission Spectrometer (TES) on MGS were used to create first complete surface mineral maps. The distribution of olivine, pyroxene, feldspar minerals and some alteration products such as sheet silicates and hematite showed heterogeneous distribution although they were sometimes obscured by the dust covering the surface (e.g., Bandfield 2002). One of the most remarkable TES results is the discovery of crystalline gray hematite in Terra Meridiani (Christensen and Ruff 2004), which was selected for the landing site of one of the Mars Exploration Rovers. However, overall TES results were interpreted to indicate a predominance of volcanic compositions within Martian dust-free surfaces although minor Mg-carbonate was discovered (Bandfield et al. 2003).

NASA's 2001 Mars Odyssey was launched in 2001 to map chemical elements and minerals, especially to look for water by using Thermal Emission Imaging System (THEMIS) and Gamma Ray Spectrometer (GRS). The Odyssey orbiter is also supporting the lander missions to provide communications relay to the earth. THEMIS is detecting various rock types of igneous origin such as olivine-rich rocks in Nili Fossae (Hamilton and Christensen 2005). Granite and dacite were found in Syrtis Major (e.g., Bandfield et al. 2004) although these evolved igneous rocks are rare. THEMIS is also finding evidence for the presence of alteration materials. Deposits of chloride minerals were recently found in numerous places of the southern highlands (Osterloo et al. 2008). GRS is detecting abundant water ice in the polar regions and possible hydrated minerals around the equatorial regions and near Apollinaris and Arabia Terra (e.g., Boynton et al. 2007).

ESA's Mars Express carries visible and infrared mineralogical mapping spectrometer (OMEGA) to build mineral maps with a resolution of up to 100 m^2 . OMEGA started its normal science mission in 2004 and so far detected variable minerals on the surface including phyllosilicates in the oldest regions of Mars (e.g., Bibring et al. 2006). They suggest that phyllosilicates formed by aqueous alteration during long-term exposure to water very early in the planet's history. OMEGA also detected several different kinds of sulfates. Outcrops in Valles Marineris, Margaritifer Sinus, and Terra Meridiani show evidence for kieserite,

gypsum, and polyhydrated sulfates (Gendrin et al. 2005). Langevin et al. (2005) reported gypsum in northern circumpolar terrains. These sulfates were probably formed under an acidic environment.

NASA's Mars Reconnaissance Orbiter (MRO), launched in 2005, carries the Compact Reconnaissance Imaging Spectrometer for Mars (CRISM), which has been acquiring visible near infrared spectral data of the Martian surface. The increased spatial and spectral resolutions of CRISM have revealed a diversity of mineral types and compositions. The major discoveries are that aqueous alteration was more widespread and diverse than previously thought. Global mapping showed that phyllosilicates are widespread, but restricted to ancient terrains and a relatively narrow range of mineralogy (Fe/Mg and Al smectite clays) (e.g., Mustard et al. 2008). CRISM data indicates the presence of kaolinite, chlorite and illite or muscovite, and a new class of hydrated silicate (hydrated silica). Diverse Fe/Mg-OH phyllosilicates are present, and smectites such as nontronite and saponite are the most common. Ehlmann et al. (2009) reported serpentine in the Thaumasia and Nili Fossae regions. These observations indicate that phyllosilicate formation occurred during the Noachian and that the conditions to form phyllosilicate (moderate to high pH and high water activity) were specific to surface environments during the earliest era of Mars's history as suggested by the OMEGA results. A wide variety of mineralogies suggest variable environments and intensities of alteration of the early crust by liquid water. Interbedded sulfates in differing hydration states occur as layers as thin as tens of meters, sometimes in a repeating sequence (e.g., Noe Debra et al. 2008). CRISM identified opaline silica spread across large regions of Mars as thin layers extending for very long distances around the rim of Valles Marineris. This observation suggests that water may have existed as recently as 2 billion years ago. CRISM also found the presence of magnesium carbonate at a regional rock layer in the Nili Fossae region (Ehlmann et al. 2008). Carbonate may be closely associated with both phyllosilicate-bearing and olivine-rich rock units and probably formed during the Noachian or early Hesperian era from the alteration of olivine by either hydrothermal fluids or near-surface water. The preservation of carbonates indicates that subsequent sulfate-forming acidic aqueous environments were not sufficiently widespread to destroy all pre-existing carbonates.

17.2.2.2 Landers

NASA's Viking Landers that landed on Chryse Planitia and Utopia Planitia, respectively in 1976, performed the first direct measurement of the surface materials on Mars. Both Viking Landers possessed X-ray fluorescence spectrometers (XRFS) and gas chromatograph mass spectrometers that assessed the chemical compositions of the soils. The measured soil compositions at the two lander sites were almost indistinguishable and lower in Al, higher in Fe, and much higher in S and Cl concentrations compared to terrestrial soils or lunar regolith (e.g., Clark et al. 1982). The Viking soil chemistry data indicated the existence of a soil component that is derived by global dust deposition. The source materials for the regolith fines are mafic to ultramafic. Sulfur abundance is correlated with Cl and possibly Mg. A simple two-component model (mixture of silicates and salts) could explain

all pair-wise trends of the eight elements analyzed at the Viking Lander sites (Clark 1993).

In 1997, NASA's Mars Pathfinder successfully landed on Ares Vallis and the Alpha Proton X-ray Spectrometer (APXS) onboard the Sojourner mini-rover analyzed the rock and soil chemistry. The analyzed rocks showed S-enrichment due to the presence of the covering dust, but the chemistry and normative mineralogy of the sulfur-free rock resemble volcanic rocks, corresponding to andesite (e.g., McSween et al. 1999). Alternatively, the rock could be altered basalt or a rock of unknown chemistry covered with silica-rich dust. The Pathfinder rock and soil chemistry are distinct from the bulk chemistry of the Martian meteorites. The soil MgO and SO₃ chemistry have are positively correlated indicating the occurrence of MgSO₄ and are basically similar to the Viking results (e.g., McSween and Keil 2000). Basaltic volcanic rock and evaporitic salts were thus proposed to dominate the planet's surface.

In 2003 NASA sent the Mars Exploration Rovers (MER) to Gusev Crater and Meridiani Planum, respectively. The Athena science payloads onboard both rovers, Spirit and Opportunity have analytical instruments to analyze the mineralogy and chemistry of the surface materials in detail. The Miniature Thermal Emission Spectrometer (Mini-TES) has brought to the Martian surface which has essentially the same technology used for Mars Global Surveyor and 2001 Mars Odyssey. Mössbauer spectrometer (MB) was for the first time brought to Mars that specifically provides the detailed mineralogy of different kinds of iron-bearing rocks and soils. APXS was again used as the updated and more capable version used for the Mars Pathfinder Sojourner rover. MER is equipped with the rock abrasion tool (RAT) to expose the dust-free rock interior for chemical analysis. MERs are still actively working to collect data after 5 years since their arrival in early 2004.

Spirit has been continuing to explore in Gusev Crater and encountered more than a dozen types of rocks (e.g., McSween et al. 2008). These rocks have very distinct characteristics in surface morphology, mineralogy and chemistry, and cover a wide range of oxidation states (e.g., Ming et al. 2008). They are grouped into classes based upon unique differences in chemistry from minor sulfate enrichment of minimally altered picrobasalt to major silica enrichment by acid leaching of or aqueous precipitation on an unidentified host rock (e.g., Squyres et al. 2008). Adirondack Class basalts are relatively unaltered materials whereas Clovis and Watchtower Class rocks and outcrops are pervasively altered materials. High sulfur in Clovis and Peace Class rocks and high P and Ti in Wishstone Class rocks are unique (e.g., Gellert et al. 2006). Other findings are alkaline basalts and the possible occurrence of a smectite-like phase in the Independence Class rocks (e.g., Clark et al. 2007). Goethite was detected in Clovis Class rocks and hematite was found in Clovis, Wishstone, and Watchtower Class rocks (Morris et al. 2006). Opaline silica was recently found in Gertrude Weise soil and nodule-like outcrops (Squyres et al. 2008). The soils along the traverse can be subdivided into three types. They are surface dust with higher oxidation state, shallow basaltic soils exposed in rover wheel tracks with very little evidence of alteration and salty soils exposed by trenches made by Spirit (e.g., Yen et al. 2008). Light-toned soils were exposed and identified to be sulfate-rich or Si-rich. The shallow Gusev trenches

clearly show increase of Cl, SO₃, MgO, Fe₂O₃, Zn and Cr, and decrease of FeO are consistent with the high solubility of MgSO₄ and the interpretation that these salts have been concentrated in the trench substratum (e.g., Wang et al. 2008). Hydrous salty soils may be relatively common at similar geomorphic settings in Gusev Crater, which is suggested by the Mars Odyssey results that 6-7 wt% water content is present in the vicinity of Gusev Crater (Feldman et al. 2005). Buried sulfates with high hydration states (Mg-Fe-Al-sulfates), like some of the light-toned soil deposits could be responsible for such water abundance. Thus, water has played a significant role in the alteration of rocks and soils in the Columbia Hills, Gusev Crater. The pervasively altered materials in Husband Hill may have formed by aqueous alteration of basaltic rocks, volcanic clasts, and/or impact ejecta by solutions that were rich in acid volatile elements. Although chemical similarities exist among different landing sites, they also exhibit their own unique signatures and histories due to local differences in geology, chemical inputs, climatic conditions, and weathering.

In contrast to the first four landers, Opportunity landed on a broad plain comprised of a thick sequence of sedimentary rock including both eolian and aqueous depositional faces (e.g., Squyres et al. 2006). Opportunity's landing site was chosen because of the abundant distribution of hematite revealed by orbiters. Opportunity analyzed outcrops of the landing site, rocks and soils along the traverse and found many interesting features on Meridiani Planum, confirming what orbiters detected from orbit. The outcrop's layers of the landing site contain tiny spherules (few millimeters in diameter) nicknamed "blueberry" that show hematite-rich mineralogy (e.g., Calvin et al. 2008) (Fig. 17.6).

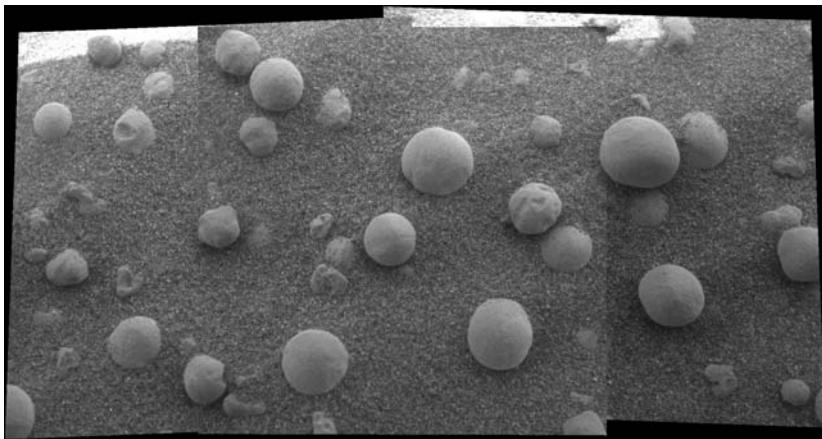


Fig. 17.6 "Blueberry" imaged by the Opportunity Mars Exploration Rover at Meridiani Planum. Image by NASA (PIA05273).

Opportunity conducted many observations to reveal the nature of these grains (e.g., Squyres et al. 2006). The APXS showed Fe-enrichment (e.g., Rieder et al. 2004) and the MB data showed the characteristic Fe³⁺ associated with hematite (e.g., Klingelhofer et al. 2004). More detailed analyses confirmed that hematites

are present in numerous soils and in small grains on the surface of dune crests. There are some explanations for the formation of these spherules. Chan et al. (2004) modeled that spherules formed as acidic groundwater percolated through porous sandstones. Morris et al. (2005) suggested that similar small hematite spherules formed during hydrothermal alteration of a cinder cone on Mauna Kea, Hawaii. Knauth et al. (2005) interpreted the spherules as accretionary lapilli during an impact base surge event. Some suggested that the spherules are interpreted to be post-depositional concretions formed during groundwater diagenesis (e.g., McLennan et al. 2005). Another important finding at Meridiani Planum is the high sulfur content of the bedrock. It contains typically 25% SO₃, suggesting that the sulfur is present as sulfate. Besides the high sulfur, mainly Mg, Fe and Ca are enriched in the remaining matrix, while Si and Al and most other elements are diluted compared to typical Martian materials like the soil. The Endurance Crater showed layering of likely very soluble minerals, Mg-sulfate and chlorides. The rock likely formed as briny water repeatedly flooded the site and then dried, building up layers dense with sulfate salts. Although localized subsurface soil spots of highly enriched ferric sulfates were found in Gusev Crater with near pure silica deposits, the high sulfur content of bedrock is constant over kilometers but dipping in deeper in Meridiani.

NASA's Phoenix landed on the polar region of Mars in 2008. The landing site and surroundings are covered with eroded ejecta deposits (e.g., Smith and the Phoenix Team 2009). Thermally Evolved Gas Analyzer (TEGA) and The Microscopy, Electrochemistry, and Conductivity Analyzer (MECA) onboard Phoenix showed the presence of Ca-carbonate (Boynton et al. 2009). The presence of perchlorates was reported in the soil surrounding the landing site (Hecht et al. 2009). The relative concentrations suggest that most of the ClO₄⁻ is likely associated with Mg²⁺, Na⁺, or both. Golden et al. (2009) compared the results with known geochemistry of the Martian polar regions and suggested that the most likely sulfur-bearing mineral is anhydrite.

17.3 Resource Utilization of Martian Materials

The first part of this chapter provided an overview of the chemistry and mineralogy of surface rock and soil by extensive studies of Martian meteorites and exploration by rovers and orbiting spacecrafsts. This valuable information can now be utilized to begin understanding how Mars soil and rock may be useful for sustaining human activities on Mars. Mineralogical and petrological characteristics of the materials on the surface of Mars, are more advanced than Moon in terms of our concept of "*Evolution of Planetary Materials*" (Takeda 2009), but the surface is not as complex as the earth because of the absence of oceans and atmosphere rich in oxygen and water. The dominant materials on Mars are igneous rocks slightly distinct from Martian meteorites (McSween et al. 2009), but recent explorations have revealed the wide distribution of weathered materials. Because Chap. 14 is devoted to the use of soil water ice as a resource for the astronauts, we will not discuss these topics in this chapter, except for the discovery of water ice on Mars in the previous section (Kargel and Furfaro 2009).

17.4 Problems of Soils for Agriculture on Mars

Wada et al. (2008) examined the possibility of Martian agriculture, for which air, water and regolith are brought into a pressurized dome at the first stage. At the second stage, they plan to revitalize air and water, in addition to supplying ordinary agricultural products such as foods, fibers and timber. Recycled use of materials is a major function of agriculture on Mars. Martian agriculture is also characterized by limited species in the ecosystem, contained inside a small dome. Information about Martian soils, acquired by landers and rovers of the Mars missions and analysis of Martian meteorites, is still insufficient in terms of resource evaluation for agriculture. For examples, textures including size distribution of soils and mineralogical species of clay minerals (phyllosilicates) have not been fully measured.

Because the main target of Martian exploration is astrobiology research to search for extraterrestrial life or biotic substances, Wada et al. (2008) pointed out that physical isolation barriers or a minimum distance should be secured between the site of scientific exploration and the dwelling site. Since it is important to keep microorganisms within a small space in the dome, we should not use microorganisms to make agricultural soils in the dome. Organic matter contributes to many important functions of the soil, to form and maintain suitable aggregate sizes in the fine regolith, and to store nutrients (bio-elements), especially Wada et al. (2008) pointed out that storage of the nutrients in organic matter can be achieved by applying compost and excreta of animals.

The important functions of microorganisms inhabiting the soil on the earth, are decomposition of various organic waste, release of nutrients stored in the organic matter, nitrification, and N-fixation. Wada et al. (2008) proposed to use peat moss with the desired microorganisms to endow the soils with these functions, but it is to be avoided to rely upon microorganisms by the above reasons. Desalination and neutralization are required to grow plants in Martian soil. Martian soil usually contains water-soluble salts. Removal of the water-soluble salts can be achieved by leaching the salt-rich soil with non-saline water. Common plants are damaged by soil with too low or too high pH. The pH of the soil should be adjusted for both too acidic and too alkaline soils. Improvement of physical properties of soil is also difficult if we do not know where we are going to live on Mars. Because we have many samples of Martian rocks, it is possible to produce better soil on the earth from them, before we will immigrate to Mars. Simulants of Martian soils "Mars-1" have been prepared at NASA Johnson Space Center (Hudson et al. 2007).

17.5 Layer Silicates in Regolithes on Mars

Clay minerals on the earth have crystal structures of layer silicates or phyllosilicates. Clay minerals are main components of terrestrial soils. Clays have been assumed to be present on Mars since the first inorganic chemical analyses of the

Viking landers (Toulmin et al. 1977). Arguments based on terrestrial analogues, have suggested that various clays of the smectite group are present on Mars, including nontronite $\text{Na}_{0.3}\text{Fe}^{3+}_2(\text{Si},\text{Al})_4\text{O}_{10}(\text{OH})_2 \cdot n\text{H}_2\text{O}$ and montmorillonite $(\text{Na,Ca})_{0.33}(\text{Al, Mg})_2\text{Si}_4\text{O}_{10}(\text{OH})_2 \cdot n\text{H}_2\text{O}$, and saponite $(1/2\text{Ca,Na})_{0.33}(\text{Mg,Fe}^{2+})_3(\text{Si},\text{Al})_4\text{O}_{10}(\text{OH})_2 \cdot 4\text{H}_2\text{O}$. These mineral phases typically result from alteration of rock-forming silicates in the presence of water. Following the results of Viking landers, Boslough et al. (1986) and others have also suggested that shocked or heated nontronite could explain the red color and magnetic properties of the Martian fines. Gooding (1978) suggested that the major phyllosilicates expected to be present on Mars are the iron-rich members: minnesotaite $(\text{Fe}^{2+},\text{Mg})_3\text{Si}_4\text{O}_{10}(\text{OH})_2$ (talc group). This mineral is a product of fayalite (olivine) weathering under Martian conditions. Greenalite $(\text{Fe}^{2+},\text{Fe}^{3+})_{2-3}\text{Si}_2\text{O}_5(\text{OH})_4$ (serpentine group) was presumed to be present in the primitive alteration assemblages depending on the CO_2 partial pressure. This mineral is a silicate phase in equilibrium with siderite. None of these phases has been observed on Mars, but recent reports using MRO-CRISM detected both serpentine and Mg-carbonate in Nili Fossae regions (Ehlmann et al. 2008, 2009). Ehlmann et al. (2009) pointed out that an olivine-serpentine-carbonate assemblage in eastern Nili Fossae would be typical of weathered, serpentized ultramafic rocks on the earth and would indicate a hydrothermal alteration process which could have served as an energy source for microorganisms and led to production of methane (Mumma et al. 2009).

The presence of clays on Mars has also been inferred from the observation of complex associations of clays and iron oxides (i.e. “iddingsite”) in Martian meteorites that have undergone preterrestrial alteration. Chevrier and Mathé (2007) summarized recent findings of clays by Mars expedition (see Sect. 17.2.2.1). Clays are also present in Mini-TES spectra of several outcrops in Meridiani Planum sedimentary deposits, and may be present in some rocks of Gusev Crater. More recently, clay minerals have been detected on the surface of Mars by Mars Express OMEGA infrared spectrometer. These clays are mainly localized on the south Noachian crust. The best matches of spectral features are Fe-rich (nontronite), Mg-rich (saponite) and Al-rich (montmorillonite) type clays. The presence of clays on the southern hemisphere old crust may indicate specific climatic conditions during the Noachian, with more abundant liquid water than required for the formation of sulfates in the conditions of the Hesperian. These clays may not be available near the dome for Martian agriculture. We recommend producing clays from common rock-forming silicates abundant on Martian surfaces.

17.6 Alteration of Rock-Forming Silicates

There seems to be a consensus that human presence is necessary to explore Mars for the purpose of clarifying the origin of life. Microorganisms should not be used for this goal. The goal of space agriculture is to create and maintain optimum living environment on extraterrestrial planet for human life. Although some scientists promoting an idea to utilize ecological system of the hyperthermophilic aerobic composting bacteria for producing soil from human and animal waste and inedible biomass, astrobiological exploration of Mars may encounter problems for detecting

signs of life on Mars if living terrestrial bacteria were introduced to Mars. We are also trying to introduce vegetables to support our life on Mars, which have been evolved on soils and atmosphere of the earth for billions of years. We have to produce soils from Martian rocks for terrestrial vegetables keeping microorganisms as small as possible.

Global scale soil formation on the earth, depends on latitude, but at a local scale, the nature of the country rock is a predominant factor that governs soil formation at the local scale. The primary Martian rocks are essentially basaltic. Many studies emphasize the importance of the primary mineralogy in addition to the physicochemical conditions of the alteration processes. Dissolution of main rock-forming silicates by inorganic and organic acids is the basic process for producing clay minerals either on the earth or Mars. Organic acids produced by decomposition or fermentation of plant leaves and roots are important reagents for natural weathering in fields or forests on the earth. We will focus on the dissolution rates of main rock-forming silicates by organic acids.

Table 17.2 Dissolution rate of major mineral (25°C, pH = 5) (Lasaga et al. 1994)

Mineral	Dissolution rate (mol/m ² /s)	Mol. Vol. (cm ³ /mol)	Dissolution time of 1 mm φ particle (y)
Quartz	4.07 x 10 ⁻¹⁴	22.688	34,000,000
Kaolinite	5.25 x 10 ⁻¹⁴	99.52	6,000,000
Muscovite	8.51 x 10 ⁻¹⁴	140.71	2,600,000
Epidote	2.45 x 10 ⁻¹³	139.2	923,000
Microcline	3.16 x 10 ⁻¹³	108.741	921,000
Prehnite	3.89 x 10 ⁻¹³	140.33	579,000
Albite	5.50 x 10 ⁻¹³	100.07	575,000
Sanidine	1.00 x 10 ⁻¹²	109.008	291,000
Gibbsite	3.55 x 10 ⁻¹²	31.956	276,000
Enstatite	1.00 x 10 ⁻¹⁰	31.276	10,100
Diopside	7.08 x 10 ⁻¹⁰	66.09	6,800
Forsterite	3.16 x 10 ⁻¹⁰	43.79	2,300
Nepheline	2.82 x 10 ⁻⁹	55.16	211
Anorthite	2.82 x 10 ⁻⁹	100.79	112
Wollastonite	1.00 x 10 ⁻⁸	39.93	79

Dissolution rates of major minerals for soil production under relatively equilibrium conditions are listed in Table 17.2 (Lasaga et al. 1994). Olivines (forsterite) and pyroxenes (enstatite and diopside) are more easily dissolved, but they do not contain much silicon and aluminum to synthesize layer silicates. Feldspars are good candidates for clay mineral formation. Uliman and Welch (2002) studied organic ligands and feldspar dissolution and pointed out that organic compounds can both enhance and inhibit the rates of feldspar dissolution in the laboratory and presumably in field settings where organic matter is abundant. Organic acids can directly enhance dissolution by either a proton- or a ligand-mediated mechanism.

Organic polymers can either inhibit dissolution by irreversibly binding to the mineral surface and reducing the effective reactive surface area (i.e. the number of reactive sites) of the mineral or enhance dissolution by a ligand-mediated mechanism. The experimental determinations of the impact of a wide variety of organic ligands on feldspar dissolution rates provide evidence for multiple mechanisms of mineral ligand interaction that can impact the dissolution process.

Dissolution experiments done by these authors are for a few organic acids and limited compositional ranges of feldspars. For example, Huang and Kiang (1972) dissolved anorthite and five feldspars by acetate, salicylate, aspartate, and citrate, and found that the net release of feldspar into solutions of organic acids is proportional to the anorthite (An) content. Yazawa et al. (2007) performed more systematic studies comparing lactic acid, oxalic acid, garlic acid, fulvic acid for dissolving Ca-rich plagioclase. They found that fulvic acid could dissolve Ca-rich plagioclase by a process similar to natural weathering involving plants. Because dissolution rates of more Na-rich plagioclases found in Martian rocks are lower than the Ca-rich members, clay formation of Na-rich members will be slower than those of the Ca-rich ones (Blum and Stillings 1995). Concentrations of ions in the solutions by these dissolution experiments will provide us with useful information for plant fertilization, described in detail in the next section.

17.7 Production of Soils by Fulvic Acids for Martian Agriculture

This section has been designed to determine if it is possible to grow plants that will provide a significant portion of the NASA-defined human nutritional requirements utilizing Martian regolith as the growing medium. Photosynthetic plants require N, P, K, Ca, Mg and S in considerable quantities and Fe, Mn, Zn, Cu, B and Mo in trace quantities. The Martian regolith contains P, Ca, Mg, S, Fe, Zn, Cu and Mo in sufficient quantities. Nitrogen is available in the atmosphere, but it is not expected to provide the full needs of plant growth. Potassium is available in small amounts in the regolith – this will also need to be supplemented. There is no evidence of B on the planet Mars. Mn is present in Martian meteorites in amounts of less than 0.5% by weight (Stoker et al. 1993). Most of the elements that are required are available, but the specific detailed mineralogy and the ease with which the plants could use them in their present form is unknown. Fertilization will be needed to provide some essential elements.

Other elements provide special concern due to the specific nature of the Martian regolith. Aluminum is present in a smaller concentration than is common in earth soils, but due to the acidic conditions on Mars, it may be a toxic level for plants. In soil solution culture experiments, Yazawa et al. (2000) found that addition of humified natural organic matter reduced the amount of monomeric Al present in solution and alleviated the toxic effect of Al on root growth of wheat under acidic condition. Consequently, terrestrial humic matters including fulvic acid can not only control to buffer for nutrient and toxic element but also accelerate to dissolve elements and synthesize clay from Martian regolith with chemical weathering.

We investigated a process of soil production from Ca-plagioclase by geochemical weathering with fulvic acid in comparison with normal acid (hydrochloric, oxalic, lactic and garlic acids). Fulvic acid is a natural organic matter produced in humified soils. The fulvic acid can be extracted from many natural materials, such as weathered coal, Canadian peat, Andsol, and forest soils. Our recent studies of fossil brine water in the Kazusa formation in the Mobara area in the Southern Kanto gas field, Chiba, Japan, showed high concentration of fulvic acid. The Kazusa formation is a sedimentary formation of interstratified shale and sand layers of the Quaternary, Pleistocene to Tertiary, Pliocene age (between 3 million and 400,000 years ago) distributed at the Boso peninsula, Chiba, Japan. The brine water in this formation includes soluble pure-methane, iodine (100 ppm), fulvic acid (40~60 ppm) and a similar chemical composition to seawater.

Fulvic acid is a complex organic acid with high molecular weight from about 680 to 920. The fulvic acid can be extracted by the procedures by the standard method suggested by the International Humic Substance Society (IHSS).

The molecular structures of fulvic acid (Fig. 17.7), extracted from waste brine water, was determined by Yazawa et al. (2005). The average chemical structure of fulvic acid had practically aliphatic (C 48.5, H 5.2, O 43.5) and high content of nitrogen (N 2.8). Now, annual 50 million tons of the brine water after production of methane and iodine around Mobara, Chiba is being discharged to the sea or returned into the gas field as drainage. Large amounts of fulvic acid will be available as industry product. We can make use of this acid for the Martian agriculture.

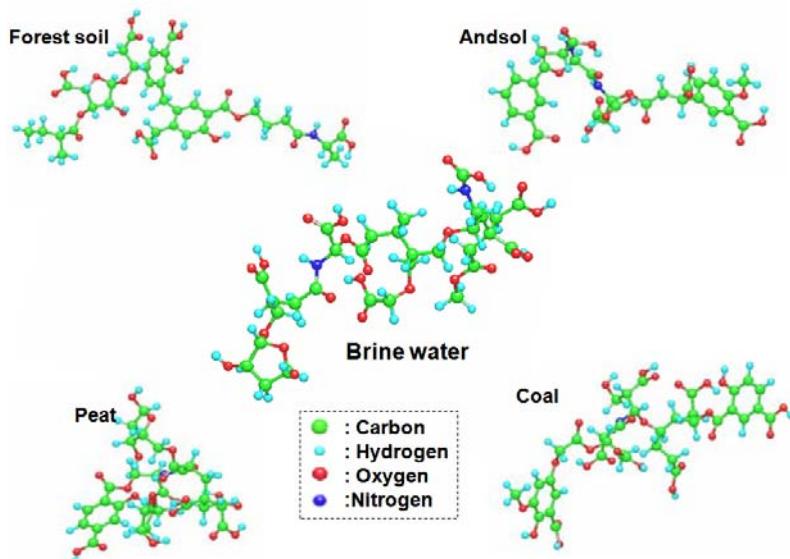


Fig. 17.7 The “average” chemical structure of fulvic acid extracted from brine water and other resources

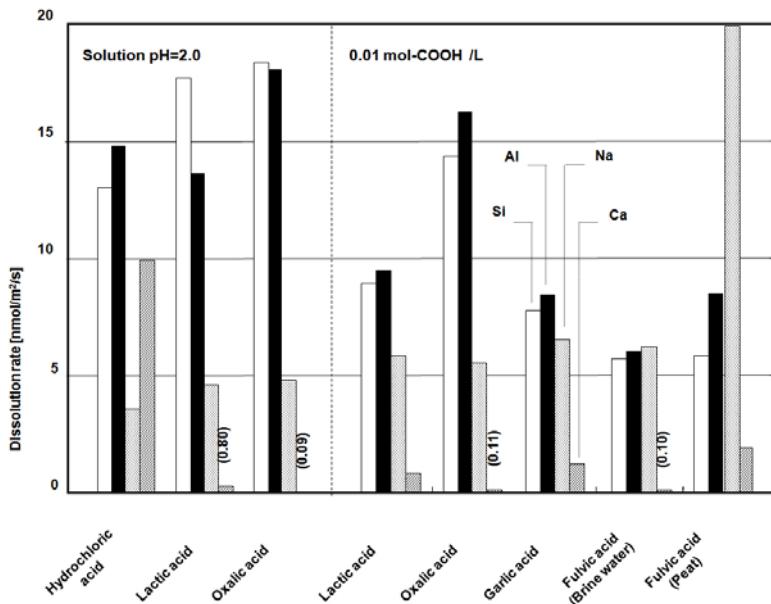


Fig. 17.8 Comparison of dissolution rate of plagioclase by fulvic and other acidic solutions

We performed experiments to produce clay minerals from plagioclase by fulvic acid. We employed Ca-plagioclase from Crystal Bay, Minnesota, with An₇₆, which is in between lunar plagioclases and Martian ones. At a carboxylic concentration of 0.01 mol-COOH/L in acid solutions, the relative effectiveness of these ligands on promoting dissolution rate of Ca-plagioclase was fulvic acid (5.7 nmol/m²/s) with the standard of Si element (Fig. 17.8). This value is only little lower than oxalic (14 nmol/m²/s), lactic (9.0), and garlic (7.8) acids. The strengths of organic acids to promote dissolution is related to the initial pK_a 2.58 (oxalic), 3.52 (lactic), 4.37 (garlic) and 4.16 (fulvic), respectively, but it is related to strength of Ca-ligand complex formed in solution (formation of secondary mineral; clay) or on the mineral surface (development of soil aggregation). Formation of clay minerals is expected after adjusting condition of pH and Al/Si ratios to bring them in the stability field of clay minerals.

Fulvic acid and its complex with iron had a strong accelerative for physiological activities of root growth in rice and chlorophyll production in phytoplankton (Fig. 17.9). Confirmation of dissolution of the main rock-forming silicates by fulvic acid will give us better understanding of the basic process of weathering in nature on the earth, as well as of application for Martian agriculture (Yazawa and Takeda 2008). Pandeya et al. (1998) performed experiments to study influence of fulvic acid on transport of iron in soils and uptake by paddy seedlings, and found that the uptake of Fe by the crop and the percentage of tissue iron content derived from fertilizer were higher in the case of Fe fulvic acid complex in comparison with FeCl₃, indicating the superiority of organically complex Fe fertilizers over inorganic salts.

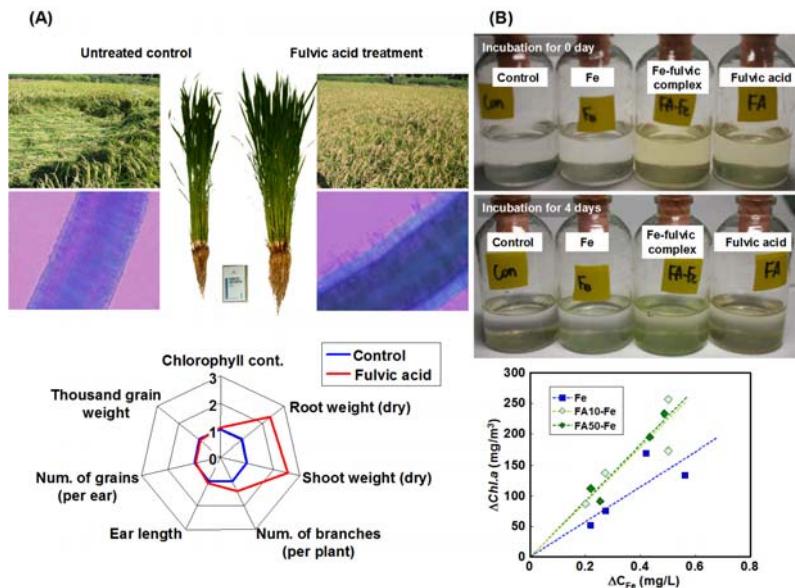


Fig. 17.9 Influence of fulvic acid or its iron complex on physiological activities of rice plant (A) and sea phytoplankton (B) under solar energy

Some nutrients may be supplied to the mixtures of rock powders, clay minerals and fulvic acid complex. Most of the plant nutrients, except N and P, are expected to be available on Mars, because they may take forms accessible to plant roots after they are released from minerals by weathering. Application of some N fertilizer (NH_3 , NO_3^- , urea etc.) or cultivation of some green manure crop plants capable of N-fixation is inevitable. P is actually contained in the minerals of the Martian regolith such as Ca phosphates, and may be released by weathering in a similar way as other nutrients. However, Wada et al. (2008) suggest that the released P is considered to be unavailable for plant roots, because of its strong adsorption (P-fixation) on CaSO_4 , CaCO_3 and iron oxides.

Organic wastes accumulated by human activities and agriculture products in the dome, have been proposed to be used as composting. However, recycling of composted materials in Martian agriculture directly applied to soils may have a chance of releasing microorganisms in the Martian ecosystem. Composting should be done in a closed vessel and their products should be sterilized by high temperature steam before applying to the farming soils in the dome.

17.8 Energy Resources of Mars

Using high dispersion infrared spectrometers at three ground-based telescopes, Mumma et al. (2009) measured methane and water vapor simultaneously on Mars over several longitude intervals, and found that the methane was released from discrete northern regions (e.g. Nili Fossae) in mid-summer. The atmosphere of

Mars is strongly oxidized, composed primarily of carbon dioxide (CO₂, 95 %) with remainder consisting of nitrogen, carbon monoxide, oxygen, water vapor, and radiogenic argon. Reduced gases such as methane (CH₄) are rare and thus should be easy to detect when their concentration increase. The principal plume contained ~19,000 metric tons of methane and the estimated source strength was comparable to that of the massive hydrocarbon seep at Coal Oil Point (Santa Barbara, CA, USA). In Mobera, Chiba, Japan, methane is released in a pond, from the Kazusa formation, as was introduced in the previous section. Both geochemical and biological origins for Martian methane have been explored, but no consensus has emerged. Most theses draw upon known terrestrial analogues such as production in magma or production by psychrophilic methanogenic biota in Mars-analogue, or cryoregimes such as permafrost. Gases accumulated in some zones might be released to the atmosphere if pores or fissures open seasonally, connecting these deep zones to the atmosphere at scarps, crater walls, or canyons.

Methane can be used as fuel directly or as fuel to generate electricity by a steam engine. A small electric generator operated by methane released from the Shimaiziri formation in the southern Okinawa Island, Japan, similar to that of Mobera, Chiba has been developed and used to supply electricity of an entire hotel (Yamaguchi et al. personal comm. 2007). Such small generator can be brought to Mars, if the base will be built near the place where methane is released. Methane contained in methane hydrate has been found in the deep sea flowers off shore of the Japan Islands (Murayama et al. 2001), and is expected to be important energy resources in future to prevent the greenhouse effect of carbon dioxide. Development of energy technology for Mars can be used for solution of the energy problems of the earth in future.

Another important materials for energy resources of Mars is silicon to be used for solar power stations. A method of producing silicon for solar panels has been introduced for dissolution of Ca-rich plagioclase. If such silicon panel factory is difficult to be constructed on Mars, silicon panel can be produced on the Moon from pure anorthosites, and may be brought to Mars. Lunar resource utilization studies in Japan have developed a method to use lunar anorthosite as an important resource. Plagioclase is the most abundant mineral in the lunar highlands and nearly pure anorthosite samples can be found close to the lunar base. Plagioclase is also available on Mars in many volcanic lavas. One problem on Mars is that calcium contents of plagioclase is lower than those of lunar volcanic rocks, and sodic plagioclase is not as easily decomposed in comparison with calcic members. It is difficult to separate resource minerals by some mechanical processes used on the earth, because they are too heavy to carry them to other planets. Pure anorthosite can be used on the Moon without separation processes. On Mars, sodium polytungstate Na₆(H₂W₁₂O₄₀)· H₂O with specific gravity 3.11 dissolved in water can be used as heavy liquid with low viscosity for separation in small scale.

Utsunomiya and Takeda (1996) used plagioclase (An₇₈) to produce silicon, aluminum, calcium and oxygen by dissolution. Plagioclase is dissolved in hydrochloric acid and after filtration of residual plagioclase, the solution is evaporated to dryness and leached by hydrochloric acid to dissolved aluminum chloride. Silica powders are obtained by filtration (Fig. 17.10). This material can be used to

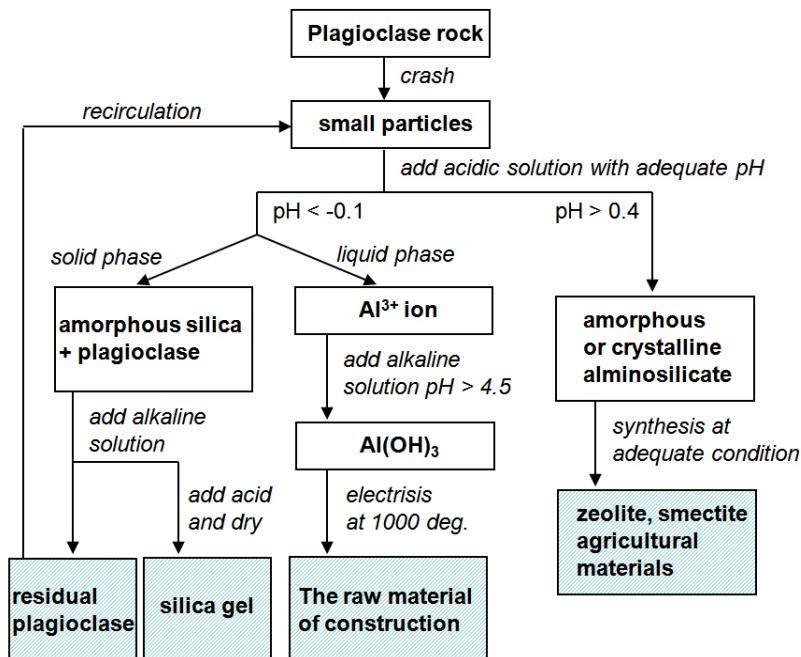


Fig. 17.10 Hydrothermal treatments on powders of plagioclase for utilization of lunar and Martian resources

produce silicon solar cells. Aluminum hydro-oxide can be precipitated by adding sodium hydroxide solution. Aluminum metal can be obtained by the electric fusion of aluminum hydro-oxide with catalytic compound, cryolite (Na_3AlF_6). Aluminum is a good construction material for low gravity on the Moon. Electric power to reduce aluminum oxides is available on the Moon by silicon solar panels. We use graphite (carbon) electrode to reduce fused materials on the earth, but carbon dioxide bubbles are generated by combining C with O_2 . If we replace graphite by platinum electrode on the Moon, we can generate oxygen bubbles to be used for breathing in the base. If we extract hydrogen from the regolithes, which were brought by solar winds, we can produce water or use it for hydrogen batteries. Other investigator used fluoride gas to produce oxygen from lunar minerals by making mixtures of CaF_2 (fluorite) and aluminum fluoride.

Calcium ion left in the solution after precipitation of aluminum, can be used to fix carbon dioxide generated in a dome of the base. Carbon dioxide is in the atmosphere of Mars, but it is not permitted to release it on the Moon. Calcium carbonate thus produced can be used as construction materials of the base such as limestone or marble, or used as cement-like materials to join bricks to built extension buildings of the base. This technique to produce calcium carbonate from plagioclase should be also developed on the earth to fix carbon dioxide to reduce the greenhouse effects.

Opportunity in Meridiani Planum confirmed the presence of crystalline hematite concretions, “blueberries”, identified by Mössbauer and mini-TES (Thermal Emission Spectrometer) with sedimentary formations of evaporitic type in Meridiani Planum (See Sect. 17.2.2.1) (Fig. 17.6). Hematite has also been found in the alteration assemblages studied by MER Spirit in Columbia Hills. These iron oxides can be used as iron ores to extract iron metal if such technology is developed. Jarosite ($KFe_3(OH)_6(SO_4)_2$ or $NaFe_3(OH)_6(SO_4)_2$) and calcium sulfates and halides are found, but their distribution are localized. Although the presence of ocean in the past has been proposed (e.g., Head et al. 1998), a thick layer of limestone and evaporates have not been found as yet. Effective use of these materials as resources, have not been developed at this stage. Other minerals can be used to extract some precious rare metals from the KREEP-rich rocks if present, although their quantities are small.

17.9 Some Problems for the Proposed Plans

Any time the soil and rock are mined for resources (e.g., water, Si, etc), attention should be paid as to what will happen to the by-products of that extraction procedure. The use of HCl on Mars will likely have significant environmental consequences. This is the reason why we select organic acids, especially fulvic acid. Waste generated from the use of fulvic acid can be added to soils without any environmental problems. On the earth, mining operations just pile the rock off to the side in “tailing piles”. The implications of doing this on Mars are different from the surface of the earth. Piles of the rock can be used as protection materials for high energy particles from space. Mixtures of fulvic acid and clay minerals are good garden soils for agriculture. Also, every time the soil and rock are disturbed for resource utilization, crucial geological and maybe biological information are potentially lost. Care should be taken to avoid this, for example, by limiting mining actions only within one crater.

Obtaining water from the icy soil at locations as indicated by the GRS maps will probably be the first and most easily obtainable resource on Mars. We are dealing with only mineralogical materials in this chapter, and this kind of discussion will be a main topic of other chapters. Water will be used for drinking and oxygen production and of course watering of any plants. Water is not the scope of mineralogy on the earth.

More detailed plans of agricultures on Mars, are described in Chap. 18, including items such as why we would want to grow crops on Mars in the first place, and some discussion as to why growing plants in an enclosed life support system would be good for human exploration. Plants provide food, recycle air, produce water (transpiration), and are nice for human psychology etc. The edited text for additional insight into the use of plants in enclosed life support systems is given by Ming and Henninger (1989).

We did not discuss the potential soil fertility on Mars, since we assumed soils produced from rocks by fulvic acid are essentially the same as those on the earth. Many inorganic nutrients are likely to be easily available by dissolving rock-forming minerals. What nutrients will need to be brought from the earth except for

fulvic acid? They are probably only nitrogen and a few others. What about toxicity of elements derived from the Martian rocks? Ni and Cr toxicity could be a problem with lunar soils, and may be so for soils derived from rocks containing pyroxenes and spinel minerals of Mars. Soil pH issues and salinity depend on the conditions of decomposing rocks by fulvic acid. There is a wealth of soil chemical and mineralogical information that can to first order provide at least an idea of what may be expected relative to the 13 essential *mineral* derived nutrients N, P, K, Ca, Mg, S, B, Cl, Fe, Mn, Cu, Zn, and Mo. Elements like B, Cu, and Mo have not been measured on Mars but these are micronutrients and if not present on Mars (unlikely) could easily be supplied. We may have the opportunity to probably be the first to discuss Martian soil fertility with the most current Mars soil chemical data, but it also depends on the style of agriculture we will adopt, as are discussed in Chap. 18.

We may have to perform more experiments to see what nutrients are potentially limiting and how they can be more easily extracted. Our dissolution experiments by organic acids, showed that especially fulvic acid will help with Fe and other micro-nutrients, by forming organo-metallic complex. The point in discussing chemical weathering and phyllosilicate formation from Martian rocks is that the chemical processes proposed here are the same as producing soils in forests on the earth.

The issue for microbial contamination from an advanced life support system could potentially be a problem for detecting life on Mars. It should be remembered, just having astronauts on Mars will pose just as much as a problem for microbial contamination as would plants. However, this problem is not within the scope of our chapter. Once we have discussed how Mars soils can be used for growing crops, then we can discuss the problems associated with growing microbial mediated plant growth. What we are proposing in this chapter is that we should limit the use of microbial means as small as possible. In a large dome for growing plants, we propose not to use products of microbiology. Composts produced from human activities should be kept in small containers, and the produced fertilizers will be applied for soils in a large dome after killing bacteria. Potential microbial contamination that will cause false positive for the detection of life should be kept as small as possible.

References

- Bandfield, J.L.: Global mineral distributions on Mars. *J. Geophys. Res.* 107, E6:5042 (2002)
- Bandfield, J.L., Glotch, T.D., Christensen, P.R.: Spectroscopic identification of carbonate minerals in the Martian dust. *Science* 301, 1084–1087 (2003)
- Bandfield, J.L., Hamilton, V.E., Christensen, P.R., McSween Jr., H.Y.: Identification of quartzofeldspathic materials on Mars. *J. Geophys. Res.* 109, 10009 (2004)
- Bibring, J.-P., Langevin, Y., Mustard, J.F., Poulet, F., Arvidson, R., Gendrin, A., Gondet, B., Mangold, N., Pinet, P., Forget, F., The OMEGA team: Global mineralogical and aqueous Mars history derived from OMEGA/Mars Express data. *Science* 312, 400–404 (2006)

- Blum, A.E., Stillings, L.L.: Feldspar dissolution kinetics. In: White, A., Brantley, S.L. (eds.) *Reviews in Mineralogy 31: Chemical Weathering Rates of Silicate Minerals*, Mineral. Soc. America, pp. 291–352 (1995)
- Boctor, N.Z., Alexander, C.M., Wang, J., Hauri, E.Y.: The sources of water in Martian meteorites: Clues from hydrogen isotopes. *Geochim Cosmochim Acta* 67, 3971–3989 (2003)
- Bogard, D.D., Johnson, P.: Martian gases in an Antarctic meteorite. *Science* 221, 651–654 (1983)
- Bogard, D.D., Garrison, D.H.: ^{39}Ar - ^{40}Ar ‘ages’ and trapped Ar in Martian shergottites, Chassigny, and ALH84001. *Meteorit Planet Sci.* 34, 451–473 (1999)
- Borg, L.E., Connolly, J.N., Nyquist, L.E., Shih, C.-Y., Wiesmann, H., Reese, Y.: The age of the carbonates in martian meteorite ALH84001. *Science* 286, 90–94 (1999)
- Borg, L.E., Draper, D.A.: A petrogenetic model for the origin and compositional variation of the martian basaltic meteorites. *Meteorit Planet Sci.* 38, 1713–1731 (2003)
- Boslough, M.B., Venturini, E.L., Graham, R.A., Williamson, D.L.: Physical properties of shocked and thermally altered nontronite: Implications for the martian surface. *J. Geophys. Res.* 91, E207–E214 (1986)
- Boynton, W.V., Taylor, G.J., Evans, L.G., Reedy, R.C., Starr, R., Jances, D.M., Kerry, K.E., Drake, D.M., Kim, K.J., Williams, R.M.S., Crombie, M.K., Dohm, J.M., Baker, V., Metzger, A.E., Karunatillake, S., Keller, J.M., Newsom, H.E., Arnold, J.R., Bruckner, J., Englert, P.A., Gasnault, O., Sprague, A.L., Mitrofanov, I., Squyres, S.W., Trombka, J.I., d'Uston, L., Wänke, H., Hamara, D.K.: Concentration of H, Si, Cl, K, Fe, and Th in the low- and mid-latitude regions of Mars. *J. Geophys. Res.* 112, E12S99 (2007)
- Boynton, W.V., Ming, D.W., Sutter, B., Arvidson, R.E., Hoffman, J., Niles, P.B., Smith, P., The Phoenix Science Team: Evidence for calcium carbonate at the Phoenix landing site. *Lunar Planet Sci XXXX, Abst#2434, Lunar Planet Inst, Houston (CD-ROM)* (2009)
- Calvin, W.M., Shoffner, J.D., Johnson, J.R., Knoll, A.H., Pocock, J.M., Squyres, S.W., Weitz, C.M., Arvidson, R.E., Bell, J.F., Christensen, P.R., de Souza, P.A., Farrand, W.H., Glotch, T.D., Herkenhoff, K.E., Jolliff, B.L., Knudson, A.T., McLennan, S.M., Rogers, A.D., Thompson, S.D.: Hematite spherules at Meridiani: Results from MI, Mini-TES, and Pancam. *J. Geophys. Res.* 113, E12S37 (2008)
- Chan, M.A., Bejtler, B., Parry, W.T., Ormo, J., Komatsu, G.: A possible terrestrial analogue for haematite concretions on Mars. *Nature* 429, 731–733 (2004)
- Chevrier, V., Mathé, P.E.: Mineralogy and evolution of the surface of Mars: A review. *Planet Space Sci.* 55, 289–314 (2007)
- Christensen, P.R., Ruff, S.W.: Formation of the hematite-bearing unit in Meridiani Planum: Evidence for deposition in standing Mars. *J. Geophys. Res.* 109, E08003 (2004)
- Clark, B.C., Baird, A.K., Weldon, R.J., Tsusaki, D.M., Schnabel, L., Candelaria, M.P.: Chemical composition of Martian fines. *J. Geophys. Res.* 87, 10059–10067 (1982)
- Clark, B.C.: Geochemical components in Martian soil. *Geochim Cosmochim Acta* 57, 4575–4581 (1993)
- Clark, B.C., Arvidson, R.E., Gellert, R., Morris, R.V., Ming, D.W., Richter, L., Ruff, S.W., Michalski, J.R., Farrand, W.H., Yen, A., Herkenhoff, K.E., Li, R., Squyres, S.W., Schroder, C., Klingerhofer, G., Bell, J.F.: Evidence for montmorillonite or its compositional equivalent in Columbia Hills. *Mars. J. Geophys Res.* 112, E06S01 (2007)
- Clayton, R.N., Mayeda, T.K.: Oxygen isotopes in eucrites, shergottites, nakhlites, and chassignites. *Earth Planet Sci. Lett.* 62, 1–6 (1983)

- Clayton, R.N., Mayeda, T.K.: Oxygen isotope studies of achondrites. *Geochim Cosmochim Acta* 60, 1999–2017 (1996)
- Ehlmann, B.L., Mustard, J.F., Murchie, S.L., Poulet, F., Bishop, J.L., Brown, A.J., Calvin, W.M., Clark, R.N., Des Marais, D.J., Milliken, R.E., Roach, L.H., Roush, T.L., Swayze, G.A., Wray, J.J.: Orbital identification of carbonate-bearing rocks on Mars. *Science* 322, 1828–1832 (2008)
- Ehlmann, B.L., Mustard, J.F., Murchie, S.L.: Detection of serpentine on Mars by MRO-CRISM and possible relationship with olivine and magnesium carbonate in Nili Fossae. *Lunar Planet Sci.* XXXX, Abst#1787, Lunar Planet Inst, Houston (CD-ROM) (2009)
- Feldman, W.C., Prettyman, T.H., Maurice, S., Nelli, S., Elphic, R., Funsten, H.O., Gasnault, O., Lawrence, D.J., Murphy, J.R., Tokar, R.L., Vaniman, D.T.: Topographic control of hydrogen deposits at low latitudes to midlatitudes of Mars. *J. Geophys. Res.* 110, E11009 (2005)
- Gellert, R., Rieder, R., Bruckner, J., Clark, B.C., Dreibus, G., Klingelhofer, G., Lugmair, G., Ming, D.W., Wänke, H., Yen, A., Zipfel, J., Squyres, S.W.: Alpha Particle X-Ray Spectrometer (APXS): Results from Gusev crater and calibration report. *J. Geophys. Res.* 111, E02S05 (2006)
- Gendrin, A., Mangold, N., Bibring, J.-P., Langevin, Y., Gondet, B., Poulet, F., Bonello, G., Quantin, C., Mustard, J., Arvidson, R., LeMouelic, S.: Sulfates in Martian layered terrains: The OMEGA/Mars Express view. *Science* 307, 1587–1591 (2005)
- Golden, D.C., Ming, D.W., Morris, R.V., Brearley, A.J., Lauer, H.V., Treiman, A.H., Zolensky, M.E., Schwandt, C.S., Lofgren, G.E., McKay, G.A.: Evidence for exclusively inorganic formation of magnetite in Martian meteorite ALH84001. *Amer Mineral* 89, 681–695 (2004)
- Golden, D.C., Ming, D.W., Sutter, B., Clark, B.C., Morris, R.V., Boynton, W.V., Hecht, M.H., Kounaves, S.P.: Sulfur mineralogy at the Mars Phoenix landing site. *Lunar Planet Sci.* XXXX, Abst#2319, Lunar Planet Inst, Houston (CD-ROM) (2009)
- Gooding, J.L.: Chemical weathering on Mars. Thermodynamic stabilities of primary minerals (and their alteration products) from mafic igneous rocks. *Icarus* 33, 483–513 (1978)
- Gooding, J.L.: Clay-mineraloid weathering products in Antarctic meteorites. *Geochim Cosmochim Acta* 50, 2215–2223 (1986)
- Gooding, J.L., Wentworth, S.J., Zolensky, M.E.: Calcium carbonate and sulfate of possible extraterrestrial origin in the EETA 79001 meteorite. *Geochim Cosmochim Acta* 52, 909–915 (1988)
- Gooding, J.L., Wentworth, S.J., Zolensky, M.E.: Aqueous alteration of the Nakhla meteorite. *Meteoritics* 26, 135–143 (1991)
- Goodrich, C.A.: Olivine-phyric martian basalts: A new type of shergottite. *Meteorit Planet Sci.* 37(suppl.), B31–B34 (2002)
- Hamilton, V.E., Christensen, P.R.: Evidence for extensive olivine-rich bedrock in Nili Fossae, Mars. *Geology* 33, 433–436 (2005)
- Head, J.W., Kreslavsky, M., Hiesinger, H., Ivanov, M., Pratt, S., Seibert, N., Smith, D.E., Zuber, M.T.: Oceans in the past history of mars: Tests for their presence using Mars Orbiter Laser Altimeter (MOLA) data. *Geophys Res. Lett.* 25, 4401–4404 (1998)
- Hecht, M.H., Catling, D.C., Clark, B.C., DeFlores, L., Gospodinova, K., Kapit, J., Kounaves, S.P., Ming, D.W., Quinn, R.C., West, S.J., Young, S.M.M.: Perchlorate in Martian soil: Evidence and implications. *Lunar Planet Sci.* XXXX, Abst#2420, Lunar Planet Inst, Houston (CD-ROM) (2009)

- Huang, W.H., Kiang, W.C.: Laboratory dissolution of plagioclase in water and organic acids at room temperature. *Amer. Mineral.* 57, 1849–1859 (1972)
- Hudson, T.L., Aharonson, O., Schorghofer, N., Farmer, C.B., Hecht, M.H., Bridges, N.H.: Water vapor diffusion in Mars subsurface environments. *J. Geophys. Res.* 112, E5 (2007)
- Johnson, M.C., Rutherford, M.J., Hess, P.C.: Chassigny petrogenesis: Melt compositions, intensive parameters, and water contents of Martian (?) magmas. *Geochim Cosmochim Acta* 55, 349–366 (1991)
- Karlsson, H.R., Clayton, R.N., Gibson Jr., E.K., Mayeda, T.K.: Water in SNC meteorites - Evidence for a Martian hydrosphere. *Science* 255, 1409–1411 (1992)
- Kargel, J.S., Furfaro, R.: Mars: Prospective Energy and Material Resources. In: Badescu, V. (ed.) *Martian water: Science, resource exploration, and base siting*. Springer, Heidelberg (2009)
- Klingerhofer, G., Morris, R.V., Bernhardt, B., Schroder, C., Rodionov, D.S., de Souza Jr., P.A., Yen, A., Gellert, R., Eylanov, E.N., Zubkov, B., Foh, J., Bonnes, U., Kankeleit, E., Gutlich, P., Ming, D.W., Renz, F., Wdowiak, T., Squyres, S.W., Arvidson, R.E.: Jarosite and hematite at Meridiani Planum from Opportunity's Mössbauer spectrometer. *Science* 306, 1746–1749 (2004)
- Knauth, L.P., Burt, D.M., Wohletz, K.H.: Impact origin of sediments at the Opportunity landing site on Mars. *Nature* 438, 1123–1128 (2005)
- Langevin, Y., Poulet, F., Bibring, J.-P., Schmitt, B., Doute, S., Gondet, B.: Sulfates in the north polar region of Mars detected by OMEGA/Mars Express. *Science* 307, 1584–1586 (2005)
- Lasaga, A.C., Soler, J.M., Ganor, J., Burch, T.E., Nagy, K.L.: Chemical weathering rate laws and global geochemical cycles. *Geochim Cosmochim Acta* 58, 2361–2386 (1994)
- McKay, D.S., Gibson Jr., E., Thomas-Keprta, K.L., Vali, H., Romanek, C.S., Clemett, S.J., Chillier, X.D.F., Maechling, C.R., Zare, R.N.: Search for past life on Mars: Possible relic biogenic activity in martian meteorite ALH84001. *Science* 273, 924–930 (1996)
- McLennan, S.M., Bell III, J.F., Calvin, W.M., Calvin, W.M., Christensen, P.R., Clark, B.C., de Souza, P.A., Farmer, J., Farrand, W.H., Fike, D.A., Gellert, R., Ghosh, A., Glotch, T.D., Grotzinger, J.P., Hahn, B., Herkenhoff, K.E., Hurowitz, J.A., Johnson, J.R., Johnson, S.S., Jolliff, B., Klingelhofer, G., Knoll, A.H., Learner, Z., Malin, M.C., McSween Jr., H.Y., Pocock, J., Ruff, S.W., Soderblom, L.A., Squyres, S.W., Tosca, N.J., Watters, W.A., Wyatt, M.B., Yen, A.: Provenance and diagenesis of the evaporite-bearing Burns formation, Meridiani Planum, Mars. *Earth Planet Sci. Lett.* 240, 95–121 (2005)
- McSween, H.Y., Stolper, E.M., Taylor, L.A., Muntean, R.A., O'Kelley, G.D., Eldridge, J.S., Biswas, S., Ngo, H.T., Lipschutz, M.E.: Petrogenetic relationship between Allan Hills 77005 and other achondrites. *Earth Planet Sci. Lett.* 45, 275–284 (1979)
- McSween, H.Y., Murchie, S.L., Britt, D.T., Brückner, J., Dreibus, G., Economou, T., Ghosh, A., Golombek, M.P., Greenwood, J.P., Johnson, J.R., Moore, H.J., Parker, T.J., Rieder, R., Singer, R., Wänke, H.: Chemical, multispectral, and textural constraints on the composition and origin of rocks at the Mars Pathfinder landing site. *J. Geophys. Res.* 104(E4), 8679–8716 (1999)
- McSween Jr., H., Keil, K.: Mixing relationships in the Martian regolith and the composition of globally homogeneous dust. *Geochim Cosmochim Acta* 64, 2155–2166 (2000)
- McSween Jr., H.Y.: The rock of Mars, from far and near. *Meteorit Planet Sci.* 37, 7–25 (2002)

- McSween, H.Y., Ruff, S.W., Morris, R.V., Gellert, R., Klingelhofer, G., Christensen, P.R., McCoy, T.J., Ghosh, A., Moersch, J.M., Cohen, B.A., Rogers, A.D., Schroder, C., Squyres, S.W., Crisp, J., Yen, A.: Mineralogy of volcanic rocks in Gusev Crater, Mars: Reconciling Mössbauer, Alpha Particle X-Ray Spectrometer, and Miniature Thermal Emission Spectrometer spectra. *J. Geophys. Res.* 113, E06S04 (2004)
- McSween Jr., H.Y., Taylor, G.J., Wyatt, M.B.: Elemental composition of the Martian crust. *Science* 324, 736–739 (2009)
- Melosh, H.J.: Impact ejection, spallation, and the origin of meteorites. *Icarus* 59, 234–260 (1984)
- Meyer, C.: The Mars Meteorite Compendium. NASA Johnson Space Center. Houston, Texas, USA (2008),
<http://curator.jsc.nasa.gov/antmet/mmc/index.cfm>
- Mikouchi, T., Miyamoto, M., McKay, G.: The role of undercooling in producing igneous zoning trends in pyroxenes and maskelynites among basaltic Martian meteorites. *Earth Planet Sci. Lett.* 173, 235–256 (1999)
- Mikouchi, T., Koizumi, E., Monkawa, A., Ueda, Y., Miyamoto, M.: Mineralogy and petrology of Yamato 000593: Comparison with other Martian nakhelite meteorites. *Antarct Meteorite Res.* 16, 34–57 (2003)
- Mikouchi, T., Koizumi, E., McKay, G., Monkawa, A., Chokai, J., Ueda, Y., Miyamoto, M.: Yamato 980459: Mineralogy and petrology of a new shergottite-related rock from Antarctica. *Antarct Meteorite Res.* 17, 13–34 (2004)
- Mikouchi, T., Kurihara, T.: Mineralogy and petrology of paired lherzolitic shergottites Yamato 000027, Yamato 000047 and Yamato 000097: Another fragments from a Martian “lherzolitic” block. *Polar Sci.* 2, 175–194 (2008)
- Ming, D.W., Henninger, D.L.: Lunar base agriculture: Soils for plant growth. ASA, CSSA, and SSSA, Madison, WI (1989)
- Ming, D.W., Gellert, R., Morris, R.V., Arvidson, R.E., Bruckner, J., Clark, B.C., Cohen, B.A., d'Uston, C., Economou, T., Fleischer, I., Klingelhofer, G., McCoy, T.J., Mittlefehldt, D.W., Schmidt, M.E., Schroder, C., Squyres, S.W., Treguier, E., Yen, A.S., Zipfel, J.: Geochemical properties of rocks and soils in Gusev Crater, Mars: Results of the Alpha Particle X-Ray Spectrometer from Cumberland Ridge to Home Plate. *J. Geophys Res.* 113, E12 (2008)
- Misawa, K., Shih, C.-Y., Wiesmann, H., Garrison, D., Nyquist, L.E., Bogard, D.D.: Rb-Sr, Sm-Nd and Ar-Ar isotopic systematics of Antarctic nakhelite Yamato 000593. *Antarct Meteorite Res.* 18, 133–151 (2005)
- Mittlefehldt, D.W.: ALH84001, a cumulate orthopyroxenite member of the martian meteorite clan. *Meteoritics* 29, 214–221 (1994)
- Morris, R.V., Ming, D.W., Graff, T.G., Arvidson, R.E., Bell III, J.F., Squyres, S.W., Mertzman, S.A., Gruener, J.E., Golden, D.C., Le, L., Robinson, G.A.: Hematite spherules in basaltic tephra altered under aqueous, acid-sulfate conditions on Mauna Kea volcano, Hawaii: Possible clues for the occurrence of hematite-rich spherules in the Burns formation at Meridiani Planum, Mars. *Earth Planet Sci. Lett.* 240, 168–178 (2005)
- Morris, R.V., Klingerhofer, G., Schroder, C., Rodionov, D.S., Yen, A., Ming, D.W., de Souza, P.A., Fleischer, I., Wdowiak, T., Gellert, R., Bernhardt, B., Evlanov, E.N., Zubkov, B., Foh, J., Bonnes, U., Kankeleit, E., Gutlich, P., Renz, F., Squyres, S.W., Arvidson, R.E.: Mössbauer mineralogy of rock, soil, and dust at Gusev crater, Mars: Spirit's journey through weakly altered olivine basalt on the plains and pervasively altered basalt in the Columbia Hills. *J. Geophys Res.* 111, E02S13 (2006)

- Mumma, M.J., Villanueva, G.L., Novak, R.E., Hewagama, T., Bonev, B.P., DiSanti, M.A., Mandell, A.M., Smith, M.D.: Strong Release of Methane on Mars in Northern Summer 2003. *Science* 323, 1041–1045 (2009)
- Murayama, Y., Fehn, U., Yoshida, S.: Recycling of iodine in Fore-arc areas: Evidence from the Iodine brines in Chiba, Japan. *Earth Planet Sci. Lett.* 192, 583–593 (2001)
- Mustard, J.F., Murchie, S.L., Pelkey, S.M., Ehlmann, B.L., Milliken, R.E., Grant, J.A., Bring, J.-P., Poulet, F., Bishop, J., Dobre, E.N., Roach, L., Seelos, F., Arvidson, R.E., Wiseman, S., Green, R., Hash, C., Humm, D., Malaret, E., McGovern, J.A., Seelos, K., Clancy, T., Clark, R., Marais, D.D., Izenberg, N., Knudson, A., Langevin, Y., Martin, T., McGuire, P., Morris, R., Robinson, M., Roush, T., Smith, M., Swayze, G., Taylor, H., Titus, T., Wolff, M.: Hydrated silicate minerals on Mars observed by the Mars Reconnaissance Orbiter CRISM instrument. *Nature* 454, 305–309 (2008)
- Nakamura, N., Unruh, D.M., Tatsumoto, M., Hutchison, R.: Origin and evolution of the Nakhlite meteorite inferred from the Sm-Nd and U-Pb systematics and REE, Ba, Sr, RB and K abundances. *Geochim Cosmochim Acta* 46, 1555–1573 (1982)
- Noe Dobrea, E.Z., Poulet, F., Malin, M.C.: Correlations between hematite and sulfates in the chaotic terrain east of Valles Marineris. *Icarus* 193, 516–534 (2008)
- Noguchi, T., Imae, N., Misawa, K., Nakamura, T.: Mineralogy of “iddingsite” and symplectite in Y000593 and Y000749: Implications for their post-crystallization and aqueous alteration. In: *Internat. Symp-Evolution of Solar System: A New Perspective from Antarctic Meteorites*, Natl. Inst. Polar Res., Tokyo, pp. 105–106 (2003)
- Nyquist, L.E., Bogard, D.D., Shih, C.-Y., Greshake, A., Stöffler, D., Eugster, O.: Ages and geologic histories of Martian meteorites. *Space Sci. Rev.* 96, 105–164 (2001)
- Osterloo, M.M., Hamilton, V.E., Bandfield, J.L., Glotch, T.D., Baldridge, A.M., Christensen, P.R., Tornabene, L.L., Anderson, F.S.: Chloride-bearing materials in the southern highlands of Mars. *Science* 319, 1651–1654 (2008)
- Pandeya, S.B., Singh, A.K., Dhar, P.: Influence of fulvic acid on transport of iron in soils and uptake by paddy seedlings. *Plant Soil.* 198, 117–125 (1998)
- Rao, M.N., Nyquist, L.E., Wentworth, S.J., Sutton, S.R., Garrison, D.H.: The nature of Martian fluids based on mobile element studies in salt-assemblages from Martian meteorites. *J. Geophys. Res.* 113, E06002 (2008)
- Reid, A.M., Bunch, T.E.: The nakhrites part II: Where, when and how. *Meteoritics* 10, 317–324 (1975)
- Rieder, R., Gellert, R., Anderson, R.C., Bruckner, J., Clark, B.C., Dreibus, G., Economou, T., Klingelhofer, G., Lugmair, G.W., Ming, D.W., Squyres, S.W., d'Uston, C., Wänke, H., Yen, A., Zipfel, J.: Chemistry of rocks and soils at Meridiani Planum from the alpha particle X-ray spectrometer. *Science* 306, 1746–1749 (2004)
- Smith, P.H., The Phoenix Science Team: Water at the Phoenix landing site. *Lunar Planet Sci XXXX*, Abst#1329, Lunar Planet Inst, Houston (CD-ROM) (2009)
- Squyres, S.W., Knoll, A.H., Arvidson, R.E., Clark, B.C., Grotzinger, J.P., Jolliff, B.L., McLennan, S.M., Tosca, N., Bell, J.F., Calvin, W.M., Farrand, W.H., Glotch, T.D., Golombek, M.P., Herkenhoff, K.E., Johnson, J.R., Klingelhofer, G., McSween, H.Y., Yen, A.S.: Two years at Meridiani Planum: Results from the Opportunity Rover. *Science* 313, 1403–1407 (2006)
- Squyres, S.W., Arvidson, R.E., Ruff, S., Gellert, R., Morris, R.V., Ming, D.W., Crumpler, L., Farmer, J.D., Des Marais, D.J., Yen, A., McLennan, S.M., Calvin, W., Bell, J.F., Clark, B.C., Wang, A., McCoy, T.J., Schmidt, M.E., de Souza, P.A.: Detection of silica-rich deposits on Mars. *Science* 320, 1063–1067 (2008)

- Stöffler, D., Ostertag, R., Jammes, C., Pfannschmidt, G., Sen Gupta, P.R., Simon, S.B., Papike, J.J., Beauchamp, R.H.: Shock metamorphism and petrography of the Shergotty achondrite. *Geochim Cosmochim Acta* 50, 889–903 (1986)
- Stoker, C.R., Gooding, J.L., Roush, T., Banin, A., Burt, D., Clark, B.C., Flynn, G., Gwynne, O.: The physical and chemical properties and resource potential of Martian surface soils. In: Lewis, J., Matthews, M.S., Guerrieri, M.L. (eds.) *Resources of Near-Earth Space*, pp. 659–707. The University of Arizona Press (1993)
- Stolper, E., McSween Jr., H.Y.: Petrology and origin of the shergottite meteorites. *Geochim Cosmochim Acta* 43, 1475–1498 (1979)
- Swindle, T.D., Treiman, A.H., Lindstrom, D.J., Burkland, M.K., Cohen, B.A., Grier, J.A., Li, B., Olson, E.K.: Noble gases in iddingsite from the Lafayette meteorite: Evidence for liquid water on Mars in the last few hundred million years. *Meteorit Planet Sci.* 35, 107–115 (2000)
- Takeda, H.: Mineralogical records of early planetary processes on the HED parent body with reference to Vesta. *Meteorit Planet Sci.* 32, 841–853 (1997)
- Takeda, H.: Evolution of Planetary Materials (in Japanese), Gendaitosyo, Sagamihara, pp. 69–110 (2009)
- Thomas-Keprrta, K.L., Bazylinski, D.A., Kirschvink, J.L., Clemett, S.J., McKay, D.S., Wentworth, S.J., Vali, H., Gibson, E.K., Romanek, C.S.: Truncated hexa-octahedral magnetite crystals in ALH84001: Presumptive biosignatures. *Geochim Cosmochim Acta* 64, 4049–4081 (2000)
- Toulmin, P., Baird, A.K., Clark, B.C., Keil, K., Rose Jr., H.J., Christian, R.P., Evans, P.H., Kellher, W.C.: Geochemical and mineralogical interpretation of the Viking inorganic chemical results. *J. Geophys Res.* 82, 4625–4634 (1977)
- Treiman, A.H.: Submicron magnetite grains and carbon compounds in Martian meteorite ALH84001: Inorganic, abiotic formation by shock and thermal metamorphism. *Astrobiol.* 3, 369–392 (2003)
- Treiman, A.H.: The nakhlite meteorites: Augite-rich igneous rocks from Mars. *Chemie Erde* 65, 203–270 (2005)
- Ullman, W.J., Welch, S.A.: Organic ligands and feldspar dissolution, water-rock interactions, ore deposits, and environmental geochemistry: A tribute to David A. Crerar. The Geochemical Society, Special Publication No. 7 (2002)
- Utsunomiya, S., Takeda, H.: Hydrothermal treatments of powders of plagioclase –Lunar resource utilization. In: Proc. 29th ISAS Lunar Planet Symp., Inst. Space Astro. Sci., pp. 44–47 (1996)
- Van de Moortèle, B., Reynard, B., Rochette, P., Jackson, M., Beck, P., Gillet, P., McMillan, P.F., McCommon, C.A.: Shock-induced metallic iron nanoparticles in olivine-rich Martian meteorites. *Earth Planet Sci. Lett.* 262, 37–49 (2007)
- Wada, H., Yamashita, M., Katayama, N., Mitsuhashi, J., Takeda, H., Hashimoto, H.: Agriculture on Earth and on Mars. In: Denis, J.H., Aldridge, P.D. (eds.) *Space Exploration Research*. Nova Science Publishers (in press) (2009)
- Wang, A., Bell, J.F., Li, R., Johnson, J.R., Farrand, W.H., Cloutis, E.A., Arvidson, R.E., Crumpler, L., Squyres, S.W., McLennan, S.M., Herkenhoff, K.E., Ruff, S.W., Knudson, A.T., Chen, W., Greenberger, R.: Light-toned salty soils and coexisting Si-rich species discovered by the Mars Exploration Rover Spirit in Columbia Hills. *J. Geophys Res.* 113, E12 (2008)
- Wasson, J.T., Whetherill, G.W.: Dynamical chemical and isotopic evidence regarding the formation locations of asteroids and meteorites. In: Asteroids, G.T. (ed.), pp. 926–974. Univ. Arizona Press (1979)

- Wentworth, S.J., Gooding, J.L.: Weathering features and secondary minerals in Antarctic shergottites ALHA77005 and LEW88516 (Abstract). *Lunar Planet Sci.* XXIV, 1507–1508, Lunar Planet Inst., Houston (1993)
- Wentworth, S.J., Gooding, J.L.: Carbonates and sulfates in the Chassigny meteorite: Further evidence for aqueous chemistry on the SNC parent planet. *Meteoritics* 29, 860–863 (1994)
- Wentworth, S.J., Gooding, J.L.: Water-based alteration of the Martian meteorite, QUE 94201 by sulfate-dominated solutions (Abstract). *Lunar Planet Sci* XXVII, 1421–1422, Lunar Planet. Inst., Houston (1996)
- Wright, I.P., Grady, M.M., Pillinger, C.T.: Carbon, Carbon, oxygen and nitrogen isotopic compositions of possible Martian weathering products in EETA 79001. *Geochim Cosmochim Acta* 52, 917–924 (1988)
- Yazawa, Y., Wong, M.T.F., Gilkes, R.J., Yamaguchi, T.: Effect of additions of brown coal and peat on soil solution composition and root growth in acid soil from wheatbelt of Western Australia. *Commun. in Soil. Sci. Plant Anal.* 31, 743–758 (2000)
- Yazawa, Y., Takarada, T., Irisawa, A., Yamaguchi, T.: An “average” structure proposed for fulvic acid isolated from iodine-brine water in Chiba (in Japanese). *Rep. Chiba Inst. Tech.* 52, 29–36 (2005)
- Yazawa, Y., Saito, M., Takeda, H.: Variation between terrestrial and lunar or martian desert-Surface materials on Mars and its possible use for agriculture (in Japanese). *Biol. Sci. Space* 21, 129–134 (2007)
- Yazawa, Y., Takeda, H.: Roles of terrestrial fulvic acid in producing agricultural soils from regolithes of habitable planets for space agriculture. In: *Soils 2008, -The Living Skin of Planet Earth* for the Joint Conference of the Australia and New Zealand Societies of Soil Science in conjunction with the International Year of Planet Earth, Abstract p.113, Massey University, NZ (2008)
- Yen, A.S., Morris, R.V., Clark, B.C., Gellert, R., Knudson, A.T., Squyres, S.W., Mittlefehldt, D.W., Ming, D.W., Arvidson, R.E., McCoy, T.J., Schmidt, M., Hurowitz, J.A., Li, R., Johnson, J.R.: Hydrothermal processes at Gusev Crater: An evaluation of Paso Robles class soils. *J. Geophys Res.* 113, E6 (2008)

Chapter 18

On-Site Resources Availability for Space Agriculture on Mars

Masamichi Yamashita¹, Hirofumi Hashimoto¹, and Hidenori Wada²

¹ Institute of Space and Astronautical Science, Japan

² Space Agriculture Task Force, Japan

18.1 Overview of Space Agriculture on Mars

Mars is the second target of our manned space flight next to the Moon, and possibly the most distant extraterrestrial body to which we could travel, land and explore within the next half century. The requirements and design of life support for a Mars mission are quite different from those being operated on near Earth orbit or prepared for a lunar mission, because of the long mission duration. A Mars mission must include at least 2.5 years for round trip travel, and a restricted opening of the launch window, both for forward and return flights once every two years. Precursor manned mission to Mars might be conducted with a small number of crew and a conservative life support system on the space ship. Once the scale of the manned mission is enlarged, an advanced bio-regenerative life support system provides an “economical” advantage over the open loop life support, based on cost comparison between the cumulative sum of consumables with the open loop system versus the initial investment for a recycling system. We further propose use of on-site resources to supplement loss of component materials in the recycling process. Reproducing recycling materials on an expanded scale is another advantage of the use of on-site resources for space agriculture.

In addition to producing foods, the function of biological and ecological systems in space agriculture can revitalize air and refresh water (Alling et al. 2005). An argument has been made for 100 % closure in the materials recycle loop, given the advantages of closed ecological life support over conventional life support in space. However, without having huge stock and sink in the loop, strong control is required to achieve full closure or such a system will be inevitably unstable. From an economical standpoint, investment or efforts to attain a higher degree of closure shall be compared to the amount of supplementary materials required to replace the dropped-out materials over the mission period. Space agriculture is

based on “more than 100 % closure” by utilizing on-site resources. Full use of on-site resources may largely reduce required mass that must be transported from Earth, particularly for expansion of life support capability to increase number of crew for extended exploration on Mars.

In order to reduce the penalty of the heaviest mass of consumption, water is partially recycled even on the International Space Station orbiting Earth. Oxygen for breathing is generated by electrolysis of water. Carbon dioxide, a metabolic product, is removed by regenerative adsorbent chemicals. Food is supplied from Earth, and feces are dumped. The distance to Mars makes it costly and difficult to send all these consumables, so space agriculture should regenerate water, oxygen and foods from metabolic waste. Space agriculture will recycle most materials, but supplement deficient components if it is economical. We aim to enlarge the scale of agriculture on Mars through a reproductive process employing on-site resources.

Prior to habitation on Mars, a small expedition crew will need to transport all required materials from Earth. There would be several phases of exploration and habitation on Mars. Early phases will consist of accumulating system components and materials resources. The space agriculture phase will begin after more than 20 years of system operation, and will need about 100 people. Instead of describing the course of development from the beginning, space agriculture at the stage of 100 people and 20 years is explained first, followed by a scenario of phased development. We will conclude with the expected requirements for on-site resources, both energy and materials, to configure and operate space agriculture for sustainable habitation on Mars.

18.2 Building a Space Agro-Ecosystem on Mars

In an ecological system, there are three members, i.e., producer, consumer and decomposer. Humans are the top consumer in agro-ecosystems. Supplying food, oxygen, and water for humans is the primary issue in engineering space habitation (Yokota et al. 2006). The main producers are photosynthetic plants, which convert solar energy to a chemical form of energy fixed in their biomass. Plants in space agriculture also act as water distillers. Processed waste water gets transported through the plant body and evaporates from leaves. Decomposers are the bridge between the consumers and the producers, and drive materials recycling. Human waste and inedible biomass are composted by bacterial action, and used to fertilize the soil with nutrients for farm plants. If no substance or element drops out from the recycling loop, the same amount of food is regenerated with oxygen in a stoichiometric relation to the biomass revitalized. Plant transpiration ratio is the ratio between amount of water transpiration and photosynthetic fixation of energy, dry mass. This ratio can estimate quantity of water that could be recovered from air in the farming space. If food and oxygen are produced by plants to fill the requirements of people, the amount of water recovered from the air in the system must exceed 200 liters per person, the typical consumption by a person on the ground.

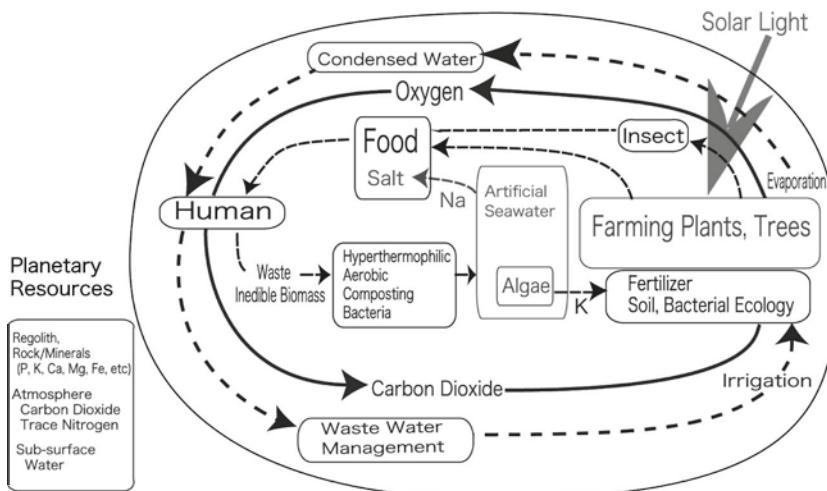


Fig. 18.1 Concept of space agriculture for habitation on Mars. On site resources are employed to make the system for more than 100% recycle possible (Yamashita et al. 2005b)

Three loops of water, oxygen, and food recycling are shown in Fig. 18.1 with major members of the agro-ecosystem on Mars. Planting trees is intended to help achieve the goal of “more than 100% closure,” since ordinary crop plants produce returns to carbon dioxide and water after metabolic consumption. By keeping biomass not oxidized, oxygen is left as excess. Wooden lumber can also be used for housing interiors and other uses in the living compartment. Leakage of air from the pressurized structure of the habitat and farming sections can be replaced with excess oxygen. Such extra production of biomass is enabled with use of the on-site resource of carbon dioxide and water available on Mars. Another by-product of having trees on Mars is the ability to rear insects on tree leaves. The addition of insects to the human diet is recommended because the placement of the consumer in the food web determines efficiency of use of original energy resource fixed by photosynthesis. Given that only 10 % of biomass energy is converted with each to that of one step up in the ladder, the higher in the ladder, the less efficient the usage of energy. Insect as part of the human diet is chosen in this context given the desire to maximize energy-efficiency of conversion.

Though elimination of sodium in processed compost is somewhat peculiar to space agriculture, terrestrial agriculture is facing salination problems in many areas now. Space agriculture provides an extreme case of salination, because of high dominance of humans in its materials loop. We propose adding marine algae and salinity resistant plants to solve this problem.

Environmental control, including building pressurized greenhouses, is essential for operating agriculture on Mars, where atmospheric pressure and temperature are both incomparably low for farming plants and animals. Figure 18.2 summarizes our design of environmental factors in the pressurized Martian greenhouse

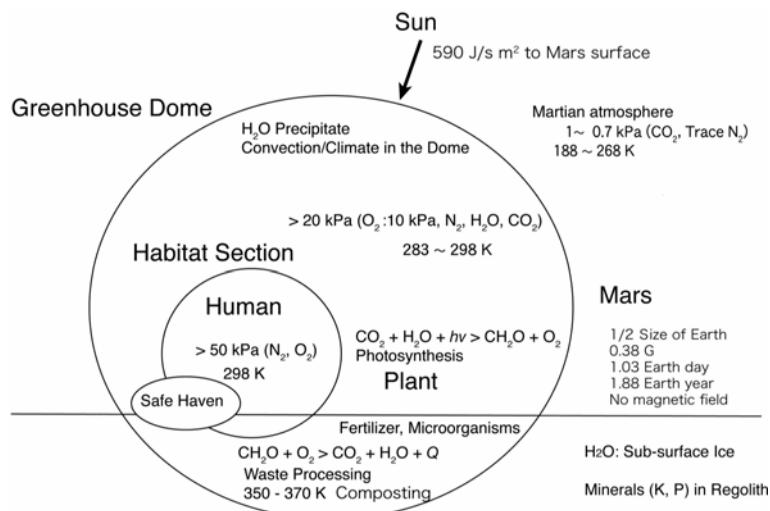


Fig. 18.2 Scheme of pressurized Martian greenhouse for space agriculture (Kanazawa et al. 2008)

structure. Pressure and the composition of the atmosphere inside the greenhouse are determined by the physiological requirements of plants. Goto et al. (2003) established a lower limit at 10 kPa of oxygen partial pressure for plant cultivation. They also showed enhanced permeation of oxygen through the seed sheath at reduced total pressure. By this enhancement, maturation of seeds was accelerated. Hinokuchi et al. (2005) and Levine et al. (2008) studied the lower limit of pressure for germination of plant seeds, and found most seeds germinate under 10 kPa of oxygen as well. For the fire safety concerns, total pressure is set to 20 kPa, balanced by inert nitrogen, together with carbon dioxide and water vapor at minor level (Yamashita et al. 2007).

Environmental conditions in the human living section were designed to meet the physiological and medical requirements for humans. A human living compartment is sectioned in the greenhouse and pressurized at a level higher than the rest of the greenhouse. Pressure in the compartment will be maintained higher than one half of Earth atmospheric pressure determined by the physiological requirements for humans. This cabin pressure would be set to higher value, if human productivity and amenity can be shown to improve with the higher pressure. Another consideration could be a pre-breathing protocol where an extra-living-section activity suit with lower inner pressure could be worn.

The highest altitude of ordinary living on Earth is around 4,000 m, where atmospheric pressure indicates 60 kPa. Many climbers require an oxygen cylinder at 8,000 m (35 kPa). In space agriculture on Mars, oxygen will be selectively transferred from the plant compartment to the sectioned human compartment using a physico-chemical membrane principle.

18.3 What to Eat on Mars

Even though the task of space agriculture is not limited to production of food, but includes the revitalization of air and water, the choice of farming species is the start of designing agriculture on Mars. The species should be selected on the basis of their nutritional content that will meet human requirements for a healthy life. Secondly, crop efficiency should be considered. Since both the area and the volume available for agricultural production on Mars are limited, the yield of crop species per unit area or per unit volume, per unit time, needs to be considered so that oxygen and food can be produced at a sufficient rate. Robustness of crop production might be further assessed, as survivability of agricultural plants is the top priority for life support on a distant planet. Several crop plant species were compared in this sense. Harvest of energy and protein is compared among species listed in Table 18.1. As an important variable for the comparison, required farming area was estimated from yield of crop, period of culture, and energy or protein content in crop. Amino acid score was an additional factor to be considered for the protein source.

Table 18.1 Evaluation of crop plant species. Modified from Katayama (2005).

Crop Plant	Yield kg/1000m ²	Period of Culture month	Energy Content kcal/100g	Required Farming Area for Energy kcal/(daym ²)	Protein Content g/100g	Required Farming Area for Protein (m ²)	Amino Acid Score	Mode of Reproduction
Rice	526	4	356	130	6.8	204	64	Wind pollination
Wheat	280	7	337	451	10.6	430	42	Wind pollination
Soybean	367	3.5	417	139	35.3	49	86	Insect pollination
Buckwheat	106	2.5	364	394	12.0	359	100	Insect pollination
Quinua	178	3	403	254	13.4	230	85	Wind pollination
Potato	3000	3	76	80	1.6	114	73	Vegetative propagation
Sweet potato	3180	5	150	64	0.9	319	83	Vegetative propagation
Cassava	7000	11	160	60	1.4	205	52	Vegetative propagation

Among major cereal crop species, rice (*Oryza sativa*) was found greatly superior over wheat (*Triticum aestivum*) and others. Modern agriculture improved yield of rice by developing proper use of fertilizer and agro-chemicals, and breeding high productive strains. Soybean (*Glycine max*) is an excellent species for providing protein, though its amino acid score is not perfect. Sweet potato (*Ipomoea batatas*) shows a higher efficiency for energy supply than potato (*Solanum tuberosum*). It also provides a sweet taste, and rich dietary fiber necessary for maintaining healthy digestive organs. Champion harvest data for Cassava (*Manihot esculenta*) exceeds sweet potato. Cassava is a quite tolerant species grown on less fertile soil.

Thus, in a space-based agro-ecosystem, the best combination of plant species to fill needs of metabolic energy, dietary fiber, protein and lipid are rice, soybean, and sweet potato (and cassava in part). In order to provide vitamins, and trace

Fig. 18.3 Model food materials to meet the nutritional requirements for one person a day. Rice 300g, soybean 100g, sweet potato 200g, green-yellow vegetable (Komatsuna) 300g, silkworm pupa 50g, loach fish 120g, and sodium salt 3g (Yamashita et al. 2007)



elements such as iron and calcium, a green-yellow vegetable such as Komatsuna (*Brassica campestris* var. *peruviridis*) or equivalent is added to the original selection. The best composition of these four, in terms of nutrition, was determined by iterative process of computation to be 300g rice, 100g soybean, 300g green-yellow vegetable, and 200g sweet potato (fresh weight) per person per day. Criteria used for this selection were: total energy intake, amino acid score (Schaafsma, 2000), adequacy of supply of each nutrient, and the energy ratio between proteins, lipids, and carbohydrates. This combination of plant products does not, however, completely meet all nutritional requirements, as it is low in sodium and lacks animal-origin vitamins and fat such as B₁₂ and cholesterol, and its amino acid content is not balanced. This is a common feature of plant-based diets. To overcome these deficiencies, sodium can be supplied in mineral form and the other problems can be addressed through supplying animal diet. Insects and fish are appropriate animal food sources in a space agro-ecosystem, given the limited area available for their rearing, and for efficient use of other resources to fill the nutritional requirements. A promising combination of the selected food substances is shown in Fig. 18.3 at an amount for one person per day.

18.4 Planting Trees and Eating Insects

Harvesting excess oxygen is the major purpose of growing trees on Mars. Wooden lumber may provide a more amenable interior for the living cabin compared to those made of metals and plastics. In order to make efficient use of bio-resources produced in space agriculture, it is essential not to downgrade bio-substances, such as combustion of bio-molecules of high entropy to produce heat, and oxidized molecules of carbon dioxide and water. If organic materials can be upgraded to substances for a use of higher priority, total efficiency of the materials loop can be improved. Converting inedible biomass to edible biomass is one “improvement” of the use of resource in space agriculture. For example, leaves of the mulberry tree are edible only at a young stage but are hard to eat after that. Mulberry

leaves contain high protein compared to other plant species. Furthermore, silk-worm larvae eat mulberry leaves, and their pupae can be eaten by humans. The advantage of eating insects is that food of animal origin can be provided with less competition with other major agricultural products.

18.4.1 Trees for Space Agriculture

Although trees have been quite common in our terrestrial life and used in many applications, wooden materials have never been employed for space crafts. The reason might be the high priority of safety and materials performance, rather than amenity for crew. Once mission duration exceeds a year, the choice of comfortable materials becomes essential for the refreshment of human spirit and provision of a less stressful environment. Thus wood production on-site will be beneficial for productive exploration and habitation on Mars, even though it requires extra area or space in the greenhouse.

A pioneering study by Nagatomo (2003) assessed the use of wooden materials for space ship or cabin based on fundamental properties of wooden materials. He also verified, with experimental evidence, the possibility of tree growth under hypobaric conditions, even as low as one-tenth of surface pressure on Earth (Nagatomo 2005). The use of wooden material can be expanded beyond providing excess oxygen, wooden cabin interiors, and insect rearing habitat. Wooden biomass can be converted to edible biomass, i.e., insect meat, by termites (*Macrotermes subhyalinus*). Culturing of wood-degrading fungi, such as Japanese mushroom (*Lentinula edodes*) or Jew's ear fungus (*Auricularia auricula*), is common in terrestrial agriculture. Vitamin D, a component deficit in the plant-based diet, is rich in these mushrooms, especially after the drying process and irradiation by ultraviolet light.

Trees are the largest living terrestrial organisms on Earth. Tree height is limited by gravity (Koch et al. 2004). Under reduced gravity on Mars, about 1/3 magnitude of terrestrial gravity, trees might grow taller. A similar phenomenon is the height of a Martian volcano, which is three times higher than Everest. Wooden plants may exhibit a physiological response to hypogravity, and reduce the production of hard components in the cell walls. Plants typically resist gravity by hardening cell walls to support their own weight. Resources allocated to hardening cell walls under normal gravity are directed to other functions when plants are exposed to reduced gravity (Hoson et al. 2002).

18.4.2 Eating Insects

Insects are a successful animal group, and about 70-75% of all animal species living on Earth are insects. They play an important role in recycling materials in the terrestrial biosphere. Furthermore, insects have been a good source of food for humans since early era, as evidenced by fossilized human feces. Because of wide availability of insects, insects have been eaten almost everywhere on Earth. Phylogenetically, insects are closely related to shrimp, lobster, and crab which are commonly eaten (Mitsuhashi 1997; Katayama et al. 2008) and which have a taste and texture similar to that of insect meat. The great diversity of insects originates in

their co-evolution with flowering plants. Inter-species interactions among insects and plants can be readily found. Thus, in many instances, the leaves of a certain plant are only eaten by one insect species, which invented the capacity to overcome the plant's defenses. Similarly, a particular plant may depend on a specific insect for pollination, at the cost of providing floral nectar and pollen. Because of such ecological and evolutional features, the design of space agriculture will depend heavily on the natural interactions between components so that the life cycle of each species will be performed, the food relationships will work out and the recycling loops can be closed. Thus, for the engineering of ecosystems we should have a good understanding of the web of interaction among species.

For space agriculture purposes, we will examine several insect species: the silkworm (*Bombyx mori*), the hawkmoth (hornworm; *Agrius convolvuli*), the drug-store beetle (*Stegobium paniceum*), the termite (*Macrotermes subhyalinus*), and fly (Diptera). Among their many advantages, these insects do not compete with humans for food resources, but convert inedible biomass or waste into edible food for humans.

18.4.2.1 Silkworm

The silkworm has been domesticated for 5,000 years in China, and probably in India too, and as a result, has lost its ability to fly. Obviously, this is advantageous for ease of rearing them. Rearing methods, including automated feeding machinery, are well established with many strains derived for producing different kind of silk fiber or attaining resistance against microbial disorders. Since the larvae feed exclusively on the leaves of mulberry (Akai and Kurabayashi 1990), the horticultural production of mulberry trees is really part of this agricultural system, and various mulberry strains are available to adapt to different farming conditions. About 40% of the leaf is digested, with the remaining 60% being excreted as feces. Final biomass of silkworm pupa is about 10% of the food consumed during its larval stage. Together with other constituents such as the larval casts, silkworm feces can be utilized in many ways, including as feed for fish. It can also be composted to increase soil fertility. Although the main reason for raising silkworms on Earth is the production of silk which can be woven into a high-quality cloth, both the silkworm pupae and the moth are quite widely accepted as snack foods in East Asia. The Kaneman Co. Ltd. sells canned silkworm pupae and moth cooked with soy sauce and sugar (www.kaneman1915.com).

18.4.2.2 Hawkmoth

Because the breeding of silkworms has been aimed at making high yield and quality of silk fiber, a large portion (65%) of amino acids goes to silk fiber, and less to the pupa meat. Silk fiber can be edible after its decomposition to small oligomer or monomer of amino acid. A wild moth species that keeps a larger portion of protein for the pupa body, and hawkmoth (hornworm) might be a good candidate for insect food. Its larvae feed on the leaves of sweet potato and also other plants. This species is a model insect for scientific studies (Kiguchi and Shimoda 1994) and its rearing technology has been well developed. The pupa is two or three times

larger than that of the silkworm, and it is very tasty when fried. Since the hawkmoth does not spin a cocoon, most of the nitrogen absorbed from plant leaves is used for the synthesis of meat protein. In this sense, the efficiency of biomass conversion from plant leaves to insect biomass is higher than for the silkworm. Even if sweet potato leaf is considered edible by humans, the use of a fraction of the leaves available for animal protein production would be advantageous because of the high dietary value of the insects. One issue that needs to be considered with this species is that they need to be airborne for normal mating to occur. It should be studied whether the adult hawkmoth is able to fly under conditions of reduced gravity and atmospheric pressure such as are likely to occur in a Martian greenhouse. Flying performance of insects in helium replaced air (Roberts et al. 2004) mimics that under hypobaric condition.

18.4.2.3 Drugstore Beetle, Termite, and Fly

Both the drugstore beetle and the termite are able to convert cellulose to animal biomass. Like ruminants, they accomplish this by having symbiotic protozoa in their gut which break down cellulose into sugars which can be utilized by the insects (Brune and Friedrich 2000). Nitrogen fixation in the termite is also accomplished by the symbiotic microbial community in the gut (Noda et al. 1999). Some work has been done on the mass production of the drugstore beetle (Kok 1983). Fly larvae grow on rotten biomass including dead animals. Fly larvae can then be fed to fish. This food web feature can upgrade biomass to edible food. Mitsuhashi (2007) suggested the use of fly larva in funeral service protocol on Mars for deceased human, a tradition in ancient China.

18.4.2.4 Model Insect Diet and Its Nutritional Value

The nutritional value of insects was examined to establish whether they could form a viable alternative to vertebrate meat and dairy products. Comparison between the protein content and amino acid composition of silkworm pupae and mulberry leaves indicates how much “vegetable” biomass can be converted to “animal” food by insects. Although mulberry leaf has relatively high protein content as compared to other plants, there is a considerable upgrading when the protein is passed through the insect. Thus, the protein content is improved and several critical amino acids are enriched. The lipid content of the silkworm pupa is another index of “animal” food composition, and is eight times that of the mulberry leaf (Katayama et al. 2005).

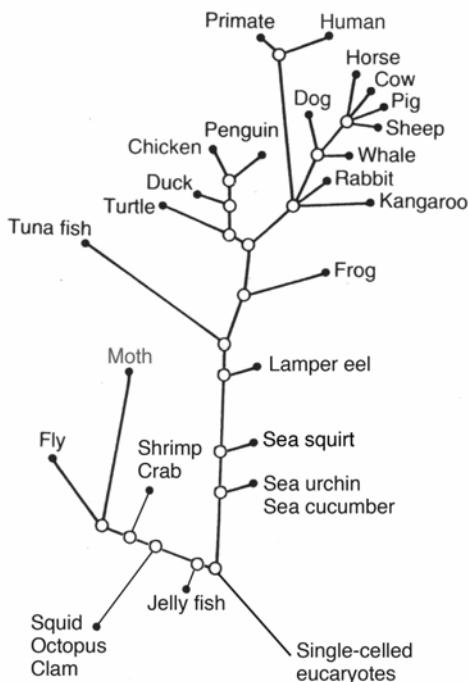
In our model diet we have added 50g of silkworm or other insect to the core composition of the four plants described earlier, i.e., rice, soybean, sweet potato, and green-yellow vegetable. In order to supply this quantity of silkworm every day, the area required for mulberry farming is estimated to be 64m²/person. For farming the four core plants, 200m²/person in addition to the mulberry farming is expected to be needed (Yamashita et al. 2006). Although insects can not provide a full replacement of nutrients obtained from vertebrate meat or avian egg, they can supply the majority of those required nutrients. Remaining nutrient requirements, vitamin D, B₁₂ and cholesterol, can then be met in different ways described in the

next section. The Japanese mushroom is included mainly as a supply of vitamin D. The amount of mushroom required in the diet can be greatly decreased if it is finely cut or powdered and then irradiated with ultraviolet light to induce the conversion of a precursor substance to vitamin D. Other trace nutrients could be fed as supplements in the form of food additives or tablets, at a minor penalty of their openness of the materials flow.

18.5 Tri-Culture of Rice, *Azolla* and Loach Fish in Rice Paddies

In order to fill the nutritional needs provided by animal origin substances, it is better to choose animal species phylogenetically close to human because of higher commonality in bio-materials. Insects are phylogenetically quite far from humans or vertebrate animals as shown in Fig. 18.4. Fish may be an alternative source of animal-origin nutrients. Aquaculture of fishes is a well established technology. Compared to other vertebrates, fish can be easily confined in their culturing water. Among many fish species that can be bred, we selected the loach fish (*Misgurnus anguillicaudatus*) for several reasons.

Fig. 18.4 Phylogenetic tree of animals we eat (except human). Modified from Eigen (1992)



The raising fish in rice paddies is done in many places, and the loach is one of the common species that naturally live in this environment. Loach is a robust species. It is resistant to adverse conditions such as poor water quality and partial drying of the rice paddies. It is able to gulp air into its digestive tube and exhaust it

from its anus after absorbing oxygen in its gut. During the winter dry season, loach survive in the deep muddy layer until spring.

Loach has a high nutritional value. A model diet with 120g of loach fish added to the vegetable and insect diet meets nutritional requirements all within the allowable range or close to the recommended level. Key point of this fulfillment is utilizing the whole loach body including internal organs. Because of small body size of loach, traditional cooking of loach accomplishes this requirement. A Korean recipe, Chueotang, minces the whole fish body into a tasty soup.

Co-culture of loach in rice paddies is extended to tri-culture of rice, *Azolla*, and loach. *Azolla* is an aquatic fern that houses symbiotic cyanobacteria in its body. Cyanobacteria in *Azolla* can quickly fix nitrogen when there is a shortage of nitrogen fertilizer. *Azolla* is an effective green manure, and also suppresses the growth of weeds by covering the water surface and blocking solar light. Even though *Azolla* itself is an edible substance, the co-culture of fish is advantageous, by converting *Azolla* to animal meat, and upgrading biomass to fish meat. The mixing action of fish in the mud and water layer in rice paddies also has other positive effects on rice production. Based on these factors, co-culture in rice paddies enhances productivity under limited available resources (FAO 1988; Watanabe and Liu 1992).

18.6 Compost Waste Safely and Quickly

Historically, recycling of human waste to fertilizer for farming has been widely performed since ancient times (Takahashi 1987). Once lifestyles were modernized and eating raw vegetables became popular, recycling of human feces and urine to soils was considered inappropriate, because of the high risk of propagation of parasites and pathogens. Even though there remain technical problems to be solved, composting is recognized as a very beneficial practical technology for solving problems of solid and organic wastes (garbage) collection and its processing. One way to accomplish a reduction in waste is by composting garbage through biological combustion. A new composting method is based on fermentation under high temperatures at 80°C to 100°C. Typical operation time of a kitchen trash box-sized composting machine is 4.5 hours for processing 1.5 kg of fresh weight of vegetable and other organic wastes. In a small sized composter, the fermentation temperature is elevated and controlled electrically to the appropriate range, because heat loss is dominant at its large ratio of surface area and volume.

In this new composting system, hyperthermophilic aerobic bacteria are utilized to attain a higher operating temperature than in an ordinary composting system. Bacteria in the hyper-thermophilic composter are active and viable under high temperatures in the range of 80°C to 100°C or even higher. Biological combustion releases heat and the temperature increases, when air is force fed through the reaction bed. Since microbial activity decreases at higher temperatures (above 100°C), temperature in the reacting bed is naturally regulated to that temperature range. Large-scale facilities using hyperthermophilic aerobic microbial systems process active sludge from sewage treatment plants and are also being used for processing other waste materials brought from farming sites or food-related industries. The ecology of these composting bacteria is

structured on intensive symbiotic interactions among multiple species that participate in various reaction networks together (Ueda et al. 2002; Oshima et al. 2007).

When hyperthermophilic composting system will be operated in the Martian greenhouse, thermal condition of reacting bed might differ from that of its terrestrial facility. Under reduced total pressure, diffusion of oxygen is enhanced. However, partial pressure of oxygen is set to a half. Reduced gravity and ambient pressure make less loss of heat from the surface. These factors affect on whether composting temperature could reach high without supplementary heating or not at a certain scale of composting facility.

Compost produced by hyper-thermophilic aerobic bacteria was well characterized by Kanazawa et al. (2003) for its application as fertilizer for cultivated plants. Soil fertilized by this compost keeps nutrient ions in forms easily accessible for uptake by plant roots. Nitrogen in organic compounds is converted to ammonium ion through hyper-thermophilic aerobic fermentation. Since typical pH of compost is weak alkaline, ammonia may be lost from composting product. The oxidative process producing nitrate ion does not take place during composting because *Nitrosomonas* or *Nitrobactor* cannot survive or be active under temperatures higher than 80°C. For the same reason, the denitrification process might not be activated. Organic nitrogen, typically amino or heterocyclics group of bio-chemicals, is either converted to ammonium or remains in an undigested form in the compost, thus reducing any losses of nitrogen.

The fate of phosphate in hyper-thermophilic aerobic composting has not yet been studied in detail. Precipitates that appear to be calcium phosphates are present on side walls of the reactor after long operation. However, identification has been made neither on this precipitate, nor any phosphate in the compost.

Potassium is the last of the three major macro-elements in fertilizer. Ring structures formed in clay minerals are known to enclose potassium ions. High affinity of potassium to such mineral structures may function as storage of potassium to provide resistance to wash-out during watering. Heavy metal ions such as copper and zinc ions are known to accumulate in farm soils where sewage sludge is misapplied. Products of the hyper-thermophilic aerobic fermentation have heavy metal ions chelated or fixed by either organic or inorganic compounds. This leads to reductions in the levels of soluble heavy metal ions. A related microbial technology can be applied in the process to clean up soil that is contaminated with heavy metal ions. It should be also recalled that certain trace amount of heavy metal ions are required for plant growth and human health.

Microbes and other organisms such as nematodes in soils are important for agriculture either positively or negatively. Little has been thus far studied about the mechanisms and factors that explain the advantage of hyper-thermophilic aerobic composting, but there are many examples showing their good performance in agriculture, and positive impact on the ecology of soil bacteria and arbuscular mycorrhizal fungi. This soil ecology is known to be effective and essential for supplying nutrients around plant root systems. No negative impacts from thermophilic aerobic composting products on the symbiosis between plant roots and soil bacteria and fungi have thus far been seen. Because of exposure to high temperatures close to 100°C during fermentation, harmful or pathogenic organisms are killed and excluded by this natural

autoclaving. Even though it should be further verified, composting bacteria themselves are presumed to be safe for humans, agricultural plants, and animal species. Hyper-thermophilic bacteria are only active at elevated temperatures, and presumably do no harm at ordinary temperatures. A further advantage of applying the compost in agricultural fields is that the chemical and micro-morphological features of compost produced by the hyper-thermophilic aerobic bacteria effectively control the ecology of bacteria and other organisms, even though the hyper-thermophilic bacteria themselves are inactive (Yamashita, et al. 2005b).

As a safety measure, the processing bed of the composter should be isolated by equipping the inlet and outlet air ports with HEPA filters. This isolation is simpler and less energy-demanding compared with the option of incineration. Hyperthermophilic aerobic composting is superior to ordinary composting in many aspects including processing time and bio-safety. In comparison with physico-chemical waste processing, a biological system consumes less energy, and requires neither high temperature ($>300^{\circ}\text{C}$) nor high pressure ($>3\text{MPa}$ for the wet oxidation process). Emission of odor and harmful volatiles is a big concern in the use of any microbial system for waste processing. Odorous compounds are produced typically in anaerobic condition. The hyper-thermophilic aerobic composting system is capable of trapping many odor components even by just mixing waste materials with its seed bed materials.

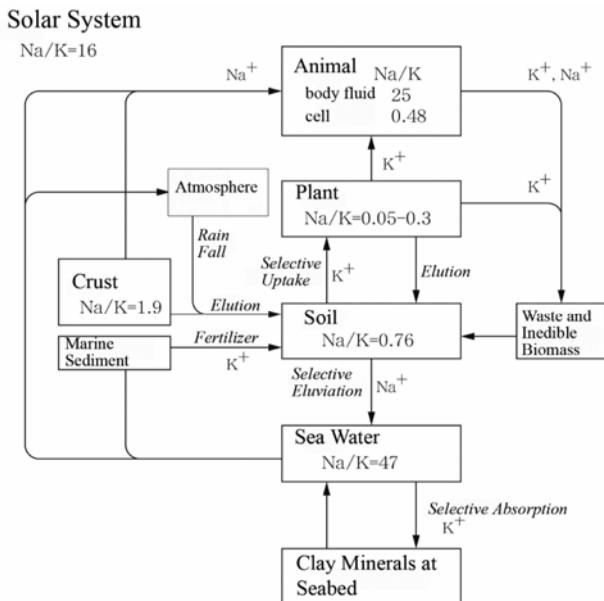
18.7 Sodium Management for Sustainable Agriculture

Human nutritional requirements include sodium salt in addition to other factors. Human excreta contains sodium. Waste and inedible biomass are composted to produce fertilizer for recycling bio-elements. This composting product contains a high concentration of sodium salt relative to potassium. However, sodium is toxic to plants at high concentrations, and ordinary plants are not grown at high salinity. Potassium is an alkali metal like sodium, and is one of the three major components of plant fertilizer.

Circulation of sodium (Na) and potassium (K) in the Earth's biosphere, and the ratio of these two elements in each part of the terrestrial ecosystem are regulated by selective absorption and desorption by clay minerals in soil and seabed, selective uptake and extrusion by plants, and other processes shown in Fig. 18.5. In space agriculture, material circulation through the human body dominates the flow of materials in the ecological system (Yamashita et al. 2005a). Therefore, salination might be a crucial problem for space agriculture.

Sodium should be separated from the compost and fertilizer, or reduced by some other measure, to prevent reduction of plant productivity. Several approaches have been proposed to solve the problem of Na and K processing in space agriculture (Yamashita et al. 1985). One is the physico-chemical process to separate the two elements (Kurokawa et al. 1992). The difference in temperature dependence on solubility of the salts of Na and K could be utilized to separate them by altering temperature to drive the dissolution and precipitation cycle. However, a problem with this kind of process is the removal of some less soluble ionic substances, such as phosphate, which are essential for plant growth.

Fig. 18.5 Circulation of sodium and potassium in terrestrial biosphere. Modified from Yamashita et al. (1985).



Biological processes might be adequate to handle sodium and potassium circulation. One possibility is to harvest potassium taken up in plant growth, and then to compost this organic material. Cultivation of marine algae to harvest potassium from the medium, and increase sodium content in artificial seawater, is one candidate technology for this purpose. An alternative approach is the selection of salinity tolerant halophytes as agricultural plant species (Ushakova et al. 2005).

18.7.1 Marine Algae to Harvest Potassium

The ionic composition of marine algal cells is high in potassium and low in sodium, similar to common land plant cells. The concentration ratio of elements between seawater and various marine algal species indicates that most elements are enormously enriched in algal bodies, compared to seawater. The few exceptions are: sodium (ratio 0.1), boron (0.1-0.3), fluorine (1), and chlorine (1). In the case of potassium, magnesium, and calcium, the ratio varies between less than and greater than 1, depending on the algal species. If algae have a homeostatic capacity to keep intra-cellular sodium and potassium stable at various sodium and potassium incubation medium concentrations, and if the ratio of intracellular sodium and potassium is rich in potassium, then the sodium rich effluent from the composting system could be managed by some marine algae.

Among many marine algal species, *Ulva* is selected as a candidate to harvest potassium from processed compost, and recycle sodium in space agriculture (Yamashita et al. 2009). *Ulva* grows in eutrophic coastal waters. It is tolerant to a wide range of salinity levels, and grows well in estuaries where salinity levels change with the tides. *Ulva* can withstand salinity levels both higher and lower than normal seawater. The ratio of sodium/potassium in *Ulva* was found to be

about 0.58 under a wide range of total salinity levels, and sodium/potassium ratio in the incubation medium spanned almost two orders of magnitude: 47 (sea water) to 0.5. Threshold concentration of potassium for the growth of *Ulva* ranges from 0.052 to 0.062 M. *Ulva* body potassium content was around 0.21 M, based on fresh weight of the specimen. This potassium content is approximately three to four times higher than the growth threshold of *Ulva*.

In space agriculture, a cultivation medium (i.e., artificial seawater) for *Ulva* will be created from human excreta. The sodium/potassium ratio in human excreta is about 2. *Ulva* can harvest potassium in its body at a sodium/potassium ratio of 0.58, therefore, enrichment of potassium can be made in the *Ulva* body from human excreta. The use of marine algae, including *Ulva*, for fertilizing farming areas is well established in many areas in the world (Notoya, 2001). Furthermore, *Ulva* itself is human food. Thus, cultivation of *Ulva* is effective and a good candidate for closing the materials recycling loop in space agriculture. *Ulva* can help solve the problem of the mismatch of sodium and potassium between humans and plants.

Ulva growth and photosynthesis has been studied extensively with regards to its applications in human diet processing, fish and animal feed, and bioremediation of the seashore. Primary photosynthetic production of marine alga reaches 1500–2000 gC m⁻² year⁻¹; equivalent to alfalfa production in soil (Mann 1973). Average dry weight production of *Ulva* is 9.63 g m⁻² day⁻¹ (Maegawa 2001). By postulating 3.2 g of potassium content in 100 g dry weight of *Ulva*, an area of 6.5 m² of farming *Ulva* is required, for one person, to harvest the quantity of potassium excreted by a human (i.e., 2 g per person per day). To extract the amount of cooking salt required by one person, 0.5 liters of artificial seawater must be processed per day.

18.7.2 Sodium Tolerant Plants

One conceivable method to solve the problem of high salinity is to use composted urine through a channel in an agricultural section for halophytes (salt-tolerant plants), which can grow in the salt-affected soil and accumulate sodium in the edible parts of the plant. Candidate halophytes are the ice plant (*Mesembryanthemum crystallinum*) (Adams et al. 1998), the saltwort (*Salicornia herbacea* L.) (Shimizu 2000), and the New Zealand spinach (*Tetragonia tetragonoides*). Among them, the ice plant is the most capable species to remediate salty soil, and has been proposed to modify reclaimed land from the sea to plant ordinary crop species (Shimoda et al. 2003). Sodium salt is excreted to a bladder, which protrudes from the plant body. Sodium salt content reaches a maximum of 30% of dry weight. The ice plant is a gourmet salad vegetable of French cuisine, enjoyed for its salty taste and peculiar texture.

18.8 Phased Development Scenario of Space Agriculture on Mars and Materials Resources

To initiate appropriate agriculture under the harsh Martian conditions, we must adopt a strategy different from that used to establish terrestrial agriculture. We propose a strategy with two stages. During the first stage, water on Mars will be

brought into a pressurized structure, rather than making that recycling loop closed. Regolith, air and water inside the greenhouse will be regulated to approximate their levels on Earth, principally by artificial means, without reliance on the autonomous control in an ecosystem. During the second stage, Martian agriculture will support human life by sufficiently supplying the necessities of subsistence mainly through recycling bio-elements inside the greenhouse. In addition, Martian agriculture will evolve by gradually expanding in its scale by taking resources available on Mars into the recycling loop.

18.8.1 The First Stage of Martian Agriculture

Soon after landing on Mars, habitation and agriculture structures should be set up, and numerous solar panels deployed on the ground surface. The selection of a settlement site is an important early step and must take into account accessibility to on-site resources. Criteria used in the selection of habitation sites on Earth could be applied to the Martian case, with certain modifications. Humans usually dwell in sunny places where both a watering spot and a shelter are within a distance for easy access. This means that detailed topographic maps of Mars are very helpful for selecting the habitation site. Characteristics of the regolith at the selected site is another important factor for conducting agriculture. Because the main goal of Martian exploration is astrobiology to search for extraterrestrial life forms or biotic substances, physical isolation barriers or a minimum distance should be secured between the site of scientific exploration and the dwelling site. Furthermore, careful unmanned exploration should be conducted prior to manned mission and agriculture on Mars (Hashimoto et al. 2006).

The structure of the greenhouse and the materials that sheath them must be tough enough to endure the inner pressure of the greenhouses surrounded by ambient pressure that is 1/100 of terrestrial atmosphere. A multi-layered or -celled structure for the greenhouses is desirable to reduce the mechanical load. In case the greenhouse is constructed on the Martian surface, the filmy sheathing materials themselves should be optically transparent to admit solar radiation necessary for human life and photosynthesis of plants, but opaque to the harmful ultraviolet part of the solar radiation.

Another option is to place the greenhouse sub-surface, where direct exposure to ultraviolet solar radiation is eliminated and cosmic rays can be shielded. Instead of direct introduction of solar radiation, an array of solar light collectors can be set up on the ground surface to guide light into the subsurface greenhouse, after eliminating short wave-length radiation by the color aberration action of an optical lens (Tanatsugu et al. 1985).

18.8.1.1 Treatment of Air and Water

The Martian atmospheric gas outside the greenhouse is pumped into the greenhouse. Appropriate total pressure and partial pressure of oxygen and other gas species in the agriculture area is determined by considering the growth and yield of cultivated plants. A trace amount of nitrogen in the Martian atmosphere is selectively collected and supplemented to the air in the greenhouses. Water can be obtained by melting

frozen regolith mined from subsurface layers. The water obtained from the frozen regolith may contain salts, such as NaCl , Na_2SO_4 , NaHCO_3 , Na_2CO_3 , CaCl_2 , CaSO_4 , or $\text{Ca}(\text{HCO}_3)_2$. Such a saline solution is unsuitable for both drinking and irrigation and needs to be purified by distillation or other means. Some organisms can help with the modification of Martian air and water. For instance, some calcareous algae grow in saline water, and precipitate CaCO_3 from it. *Spirulina*, photosynthetic bacteria and tilapia may flourish in the resulting saline alkaline water, as evidenced at alkaline and/or salty lakes in Africa. *Spirulina* converts CO_2 to O_2 and produces biomass by photosynthesis during its high growth rate. *Spirulina*, tilapia and photosynthetic bacteria are also nutritious foods for humans during the first stage.

18.8.1.2 Formation of Martian Soil from Regolith

The initial mode of agriculture on Mars will be hydroponics. Inedible biomass and human metabolic waste are composted and combined with regolith to form soil to make it habitable for plant roots and soil microorganisms. Unfortunately, information about the regolith is still fragmentary, and somewhat contradictory in certain aspects. We shall explore Mars to map and collect comprehensive data on the regolith, which is required to design space agriculture and select the habitation site. The regolith on Mars is widely covered by hematite. This indicates that ferruginous minerals such as olivine, pyroxene and hornblende have been weathered to release ferrous iron, which is oxidized by mechanisms not yet fully understood. In addition, the presence of CaSO_4 and jarosite indicates that even Ca-, Na-, and K-rich minerals such as feldspars have been weathered. However, the common presence of olivine indicates that the degree of the weathering is generally low. Furthermore, the presence of jarosite would mean the regolith is acidic. On the contrary, the presence of MgCO_3 or CaCO_3 would mean the regolith is somewhat alkaline, as indicated by the Phoenix mission. In any case, the following stepwise efforts are necessary to form Martian soil (Yazawa et al. 2007).

18.8.1.3 Desalinization and Neutralization of Soil

Martian regolith usually contains water-soluble salts. If their concentration is so high as to harm most organisms except for halo-tolerant or halophilous organisms, desalinization (removal of the water-soluble salts) is imperative. Desalinization can be achieved by leaching the salt-affected regolith with non-saline water. The leaching water enriched with the salts can be used to strengthen regolith bricks for use in the construction of space habitation.

The pH of the regolith may vary from place to place. The pH should be very low in places dominated by jarosite and should be high in places dominated by NaHCO_3 or Na_2CO_3 . However, in most places, pH may be somewhat high where CaCO_3 or MgCO_3 is found. Common plants are damaged by soil with too low or too high pH. Accordingly, we must prepare methods of neutralization for both too acidic and too alkaline regolith. The application of calcareous algae rich in CaCO_3 , or fine regolith rich in olivine, is effective for increasing the pH of acidic regolith. For alkaline regolith, the application of carried peat moss with low pH may solve the problem.

18.8.1.4 Improvement of Regolith/Soil Properties

The sediment of fine textured regolith (texture: particle size distribution) has some properties favorable for agriculture, because its cation exchange capability (CEC), mineral nutrient content, and ability to retain organic matter are high. However, fine textured sediment is problematic because of poor aeration and poor drainage, which result in O₂-deficiency for plant roots and microorganisms living in the sediment. Desalinization by leaching is also difficult. A large portion of the pores in the fine textured sediment is liable to be completely filled with water by capillary force. Accordingly, loamy textured soils are considered desirable for common upland crops in terrestrial agriculture. This suggests that a sandy loam or loamy sand texture of Martian regolith is desirable for Martian agriculture under low gravity. However, the aeration and drainage of soil is controlled not only by texture but also by a mode of aggregation of soil particles. For instance, drainage of terrestrial soils with sandy loam or loamy sand texture is often poor. This problem with the physical properties of terrestrial soil is usually overcome by the presence of aggregates of suitable sizes (e.g., > a few mm), which are naturally formed by a concerted action of biotic and abiotic processes and can be artificially formed by applying synthetic polymers or some materials with coarse pores. Pores among the aggregates of suitable sizes are coarse enough for desirable aeration and drainage. Accordingly, the artificial aggregation of Martian regolith should improve its physical properties. A promising candidate is polyvinyl alcohol (PVA) (Dejbhimon and Wada 2005). Another candidate method to improve the physical properties of Martian regolith is to apply peat moss or an equivalent processed plant material produced in agriculture on Mars (Wada 2008). Both are rich in coarse pores for aeration and drainage.

18.8.1.5 Plant Nutrients

Most of the plant nutrients are available on Mars, because they may take forms accessible to plant roots after they are released from minerals by weathering. Nitrogen is not a component of any minerals and accordingly is utterly absent in the regolith. Application of some nitrogen fertilizer (ammonium nitrate, urea) or cultivation of some green manure crop plants capable of nitrogen fixation, such as *Azolla*, is inevitable. However, the nitrogen fixation will not occur if the amount of available phosphorus is far short of the demand of the green manure crop plants.

Phosphorus is contained in the minerals of the Martian regolith, and may be released by weathering in a similar way as other nutrients. However, the released phosphate is considered to be unavailable for plant roots, due to its strong adsorption (phosphate-fixation) on CaSO₄, CaCO₃ and iron oxides. Thus, phosphate fertilizers should be applied by three appropriate methods to avoid this phosphate-fixation. The first is to apply phosphate fertilizers mixed with peat moss (Kawakami et al., 2007). The second is the removal of CaSO₄ by leaching if the main phosphate-adsorbent is this substance. The third is the adjustment of pH at about neutral where phosphate-adsorption with both CaCO₃ and iron oxides is most suppressed if these two substances are dominant phosphate-adsorbents.

Excess amounts of several elements are harmful to common terrestrial crop plants, and this problem may occur on Mars. For instance, if the regolith is enriched with olivine, excess Mg, Ni and Cr may damage common crop plants. The application of CaCO_3 may cure excess Mg, while the application of peat moss, which possesses a strong capability to adsorb Ni and Cr, may effectively counteract the harm caused by excess Ni and Cr.

18.8.1.6 Organic Matter in Soil

Organic matter contributes to many important functions of the soil, two of which may be essential in forming desirable Martian soil. One function is to form and maintain aggregates of suitable sizes in the fine regolith. The other is to store nutrients (bio-elements), especially nitrogen. The nutrients stored in organic matter are released when the organic matter is microbiologically decomposed. Therefore, accumulation of organic matter containing nitrogen in the regolith is imperative to establish space agriculture. This can be achieved by applying compost and processed excreta of animals and people.

Hyper-thermophilic aerobic bacteria can also help conserve Martian environment for astrobiology exploration. Since the organic waste cannot be dumped from the spacecraft flying to Mars to meet the planetary protection requirements, the accumulated organic waste can be quickly composted with hyper-thermophilic aerobic bacteria.

18.8.1.7 Soil Microorganisms

Microorganisms inhabiting the soil are responsible for most of the functions of the soil, such as decomposition of all organic waste, release of nutrients stored in the organic matter, nitrification, denitrification and nitrogen fixation. To endow the regolith with these functions, the Martian regolith should be inoculated with desired microorganisms.

18.8.2 The Second Stage in the Initiation of Martian Agriculture

When the goal of the first stage is almost attained, the second stage of space agriculture should be started. The main objectives in the second stage are to establish a sustainable agriculture that can continuously supply to the Martian emigrants clean air and water, as well as foods, fiber and timber, mainly by recycling substances as described in the earlier section of this chapter and summarized in Fig. 18.1. Reservoirs of the recycling flows of substances are not drawn in Fig. 1, but they are necessary for buffering fluctuation of the flows. Oxygen concentration is most critical for life support, and at low oxygen concentrations oxygen is generated by electrolysis of water, as is presently being done on the International Space Station. During the second stage of Martian habitation, most of the tasks in the first stage should continue to work to keep the conditions inside the greenhouses favorable for recycling materials in the second stage.

18.8.2.1 Composting

Organic waste originates from both the living quarter and the agriculture quarter. It should be composted before application to the farming soil. This is the main method of recycling materials in Martian agriculture, because this is the transfer from consumer to producer by the decomposer in the Martian ecosystem. *Geobacter* is another candidate to process organic substances without oxygen, using the iron oxides commonly available on Mars (Lovely 2006). The hyper-thermophilic aerobic composting discussed earlier becomes the major process in the second stage when quick and safe treatment of a large amount of organic waste is urgently needed.

18.8.2.2 Maintenance of Soil Conditions

The soil formed from the regolith in the first stage should be maintained to be productive and manageable. In other words, the application of a suitable amount of compost and/or organic waste is not only necessary for material recycling, but also necessary for maintaining soil productivity. Furthermore, accumulation of organic matter in the soil helps to ensure high O₂ content in the air and acts as a reservoir for bio-elements.

Careful irrigation is necessary for the following three reasons. The first is to save water. The second is to avoid accumulation of salts in the soil by keeping the evapotranspiration rate lower than the irrigation rate. The third is to avoid the damage of dikes and other agricultural landforms by turbation and the formation of deep cracks in the soil. Such a problem is known in clayey terrestrial soils rich in smectite, which are classified as Vertisols. Vertisols swell when moist, and shrink when dry, leading to cracks and turbation of the clayey soils. If such problems are anticipated in Martian soil, the soil should not be exposed to distinct alternation of wetting and drying by improper irrigation.

18.9 On-Site Energy Resources for Space Agriculture

Energy is required for photosynthesis of plants, environmental control and other usages in space agriculture. Solar radiation is the most important energy source. It is the critical driving energy in our life and activities in the terrestrial biosphere, and its role is unchanged in Mars habitation. Since day length on Mars is almost the same as that of Earth, biological organisms including agricultural plants can be expected to have little difficulty adapting their circadian biological rhythm to the Martian day and night cycle.

18.9.1 Energy for Photosynthetic Production

The choice of an energy source is made after considering the energy quantity required for agriculture. On Earth, area of farm land is 2,300 m² per person. Solar light energy influx on this farming area is in the order of 1 MW. Interestingly, at the Closed Ecology Experiment Facilities (CEEF) in Aomori, two men and two small sized goats consume 1.5 MW of electric energy in a closed ecological life support system based on physico-chemical principles and hydroponic plant

cultivation under mostly artificial light (Nitta et al. 2000). A human being is a 100W animal in terms of energy consumed physiologically. Although a low coverage of plants on Earth's surface and other factors provide the harvest energy index of 5.5×10^4 , the published physiological plant data (Taiz and Zeiger 2002) show that the conversion ratio (light energy to fixed biomass energy) is as high as 5×10^{-2} . The ground surface area required per person in the Martian greenhouse is estimated to be $\sim 2,000 \text{ m}^2$, assuming that the surface coverage of plants multiplied by the ratio of light reaching the surface = 0.2, the ratio of edible parts in biomass = 0.5, the conversion ratio 5×10^{-2} , irradiance of sunlight (190 W m^{-2} as a daily average at the equator zone on Mars) = 590 J/s m^2 , and energy required for habitation = $2 \times 10^3 \text{ W}$ per person. Thus, the minimum requirement of ground surface area for physiological needs (100W human metabolism, and a safety factor of 2) is reduced to 200 m^2 per person. This areal requirement for the pressurized compartments can be further trimmed down by supplementary light irradiation to plants at a level up to twice as much as solar light. Because of the high profile of light energy required, solar light is the right choice for Martian agriculture. For other energy requirements, electricity is the appropriate form of energy for Martian habitation and agriculture. Generation of electric energy by the solar panels is disrupted due to the non-availability of solar radiation at night, and even during daytime, by dust storms. Consequently, energy storage systems, such as batteries or heat baths, are required to supply energy without disruption and supplement lighting for crop control. For the safety of the habitation site on Mars, it might be equipped with a nuclear power unit or a combination of orbiting solar power satellites and ground rectina for receiving microwave energy (Ishikawa et al. 1990) for back-up of the power system.

18.9.2 Energy for Thermal Control in the Greenhouse Structure

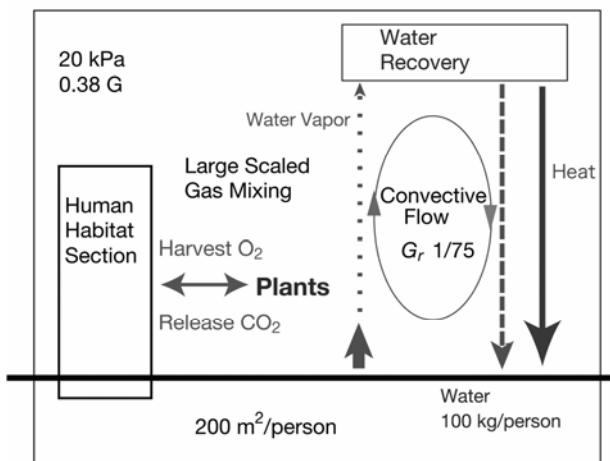
Energy is required to control temperature and drive thermal convection to transport gaseous substances by air currents induced in the greenhouse may be necessary as this section below explains. Terrestrial greenhouses can be either warmed or cooled by solar light and other energy, and inner temperature is maintained at the appropriate set range. Thermal design of the greenhouse starts with examination of its heat balance and budget. Surface temperature on Earth is determined by amount of solar light inflow and equivalent radiative heat outflow proportional to the fourth power of surface temperature. Because solar light intensity is $\sim 1/2$ on Mars compared to that on Earth (Kieffer et al. 1992), surface temperature is low compared to that on Earth. Temperature inside the greenhouse will be warmer than the ambient atmosphere. This is partly because of absorption of solar energy made by light absorbing materials, and trapping of warm air in the greenhouse. Heat dissipation from the greenhouse in the design where the greenhouse is exposed to Martian atmosphere, has two major paths. One is heat radiation and conductive heat transfer to ambient air from the outer surface of the greenhouse. The other is heat loss to the frozen surface regolith of Mars. Heat dissipation from a subsurface greenhouse can take only the latter path, and this can be controlled by a heat insulation layer. The heat transfer coefficient through the former path is

quite low given the very low density of Martian atmosphere. A conductive heat loss from the outer surface of the Martian greenhouse is therefore not extraordinary in spite of the extremely cold outside atmosphere. This makes it possible to maintain a high inner temperature under the condition of 50% heat inflow to the Martian greenhouse as compared with Earth conditions.

18.9.3 Convective Mixing of Gas Components in Greenhouse

Atmospheric convection and localized climates are expected to emerge from heat flows in the Martian greenhouse as shown in Fig. 18.6.

Fig. 18.6 Heat and materials convection in the Martian greenhouse. Modified from Yamashita et al. (2006)



Water vapor evaporated from plants is transferred to the recycling site for collection and recycling. Oxygen produced by photosynthetic reactions is transported to a concentrating apparatus and exchanged with carbon dioxide. All gaseous components need to be distributed uniformly in the greenhouse by the convective mixing. Placement of hot and cold spots in the greenhouse should be well planned in order to induce appropriate convection. The flow induced by thermal convection in the Martian greenhouse is characterized by non-dimensional variables (Kotake and Hijikata 1992). Among them, Grashof number (Gr) is used to compare the Martian greenhouse to its terrestrial counterpart. This Gr number is defined as the ratio between buoyancy force and viscous force in flow, and it characterizes natural convective flow. Since Gr number is proportional to the magnitude of gravity, i.e., $1/3$ of Earth's, as well as the square of gas density, Gr in the Martian greenhouse is $1/75$ of Gr in the terrestrial counterpart provided that other variables (characteristic length, thermal expansion coefficient, dynamic viscosity, and change in temperature) are the same. Low number of Gr means less natural convection, and forced convection might be required in the Martian greenhouse.

18.10 Summary

A conceivable plan is proposed to realize Martian agriculture in greenhouse on the lifeless barren regolith under hostile conditions, based on the knowledge and experiences of terrestrial ecosystems and agriculture. Our scenario for developing space agriculture with phased development largely depends on on-site resource availability. To lessen the export amount from Earth, we propose more than 100% closure of materials recycle in the life support system. Solar light energy, water, carbon dioxide, and other bio-elements introduced from the Mars surface enable an enlargement of the habitation. To test the feasibility of this strategy, many terrestrial experiments should be carried out to examine each part of the plan under simulated Martian conditions, starting with the first step. Precursor mission to Mars should also be conducted to determine the availability of required resources, and characterize the environment in greater detail.

Currently, many terrestrial ecosystems are being quickly destroyed, leading to a decrease in Earth's biodiversity. At the same time, terrestrial agriculture in many places has been losing its productivity, principally due to soil erosion and degradation. This situation with terrestrial agriculture somewhat resembles the challenging conditions for Martian agriculture. This suggests that the success of Martian agriculture may be helpful for reviving damaged terrestrial ecosystems and renewing impaired terrestrial agriculture. This may be the next revolution of our human history, which has developed through the agricultural revolution, the industrial revolution and the information revolution.

References

- Adams, P., Nelson, D.E., Yamada, S., Chmara, W., Jensen, R.G., Bohnert, H.J., Griffiths, H.: Growth and development of *Mesembryanthemum crystallinum* (Aizoaceae). Tansley Review No. 97 New Phytologist 138, 171–190 (1998)
- Akai, H., Kurabayashi, S.: Wild Silk, Science and Technology. Science House, Tokyo (1990)
- Alling, A., Van Thillo, M., Dempster, W., Nelson, M., Silverstone, S., Allen, J.: Lessons Learned from Biosphere 2 and Laboratory Biosphere Closed Systems Experiments for the Mars On Earth Project. Biol. Sci. Space 19, 250–260 (2005)
- Brune, A., Friedrich, M.: Microecology of the termite gut: structure and function on a microscale. Curr. Opinion Microbiol 3, 263–269 (2000)
- Dejphimon, K., Wada, H.: Improvement of the saline soil in Northeast Thailand using polyvinyl alcohol (PVA). In: Proceedings of “Management of Tropical Sandy Soils for Sustainable Agriculture, A holistic approach for sustainable development of problem soils in the tropics, Khon Kaen, Thailand, 27th November-2nd December 2005, pp. 123–128 (2005)
- Eigen, M.: Steps towards life. A perspective on evolution, p. 61. Oxford University Press, Oxford (1992)
- FAO, The “Rice-Azolla-Fish” System. RAPA publication, vol.4, FAO, Regional Office for Asia and Pacific, Bangkok (1988)

- Goto, E., Arai, Y., Omasa, K.: Seed production of *Arabidopsis thaliana* under hypobaric conditions. *Biol. Sci. Space* 17, 267–268 (2003)
- Hashimoto, H., Koike, J., Yamashita, M., Oshima, T.: Space Agriculture Saloon; Proposal for Extension of Planetary Protection Policy to Avoid Biological Contamination of Mars with Manned Exploration Supported by Space Agriculture. *Viva Origino* 34, 86–89 (2006)
- Hinokuchi, T., Akiba, M., Hashimoto, H.: Germination rate of white radish and buckwheat under low pressure in closed environment. *Space Utiliz Res.* 21, 315–318 (2005)
- Hoson, T., Soga, K., Mori, R., Saiki, M., Nakamura, Y., Wakabayashi, K., Seiichiro Kamisaka, S.: Stimulation of Elongation Growth and Cell Wall Loosening in Rice Coleoptiles under Microgravity Conditions in Space. *Plant. Cell. Physiol.* 43, 1067–1071 (2002)
- Ishikawa, Y., Ohkita, T., Amemiya, Y.: Mars habitation 2057: Concept design of a Mars settlement in the year 2057. *J. British Interplanetary Soc.* 43, 505–512 (1990)
- Kanazawa, S., Yamamura, T., Yanagida, H., Kuramoto, H.: New production technique of biohazard-free compost by the hyper-thermal and aerobic fermentation method. *Soil Microorganisms* 58, 105–114 (2003)
- Kanazawa, S., Ishikawa, Y., Tomita-Yokotani, K., Hashimoto, H., Kitaya, Y., Yamashita, M., Nagatomo, M., Oshima, T., Wada, H.: Space Agriculture Task Force. Space agriculture for habitation on Mars with hyper-thermophilic aerobic composting bacteria. *Adv. Space Res.* 41, 696–700 (2008)
- Katayama, N., Ishikawa, Y., Takaoki, M., Yamashita, M., Nakayama, S., Kiguchi, K., Kok, R., Wada, H., Mitsuhashi, J.: Space Agriculture Task Force. Entomophagy: a key to space agriculture. *Adv. Space Res.* 41, 701–705 (2008)
- Katayama, N., Yamashita, M., Wada, H., Mitsuhashi, J.: Space Agriculture Task Force. Entomophagy as Part of A Space Diet for Habitation on Mars. *J. Space Tech. Sci.* 21, 27–38 (2005)
- Kawakami, H., Niijima, Y., Ota, Y., Wada, H., Lan, Z., Wang, Z., Li, S.: Increase in growth and yield of peanut through improvement of the saline alkali soil by applying a domestic peat. *J. Jpn. Peat. Soc.* 4, 22–32 (2007)
- Kieffer, H.H., Jakosky, B.M., Snyder, C.W., Matthews, M.S.: Mars. Univ. Arizona Press, Tucson (1992)
- Kiguchi, K., Shimoda, M.: The sweet potato hornworm, *Agrius convolvuli*, as a new experimental insect: Continuous rearing using artificial diets. *Zool. Sci.* 11, 143–147 (1994)
- Koch, G.W., Sillett, S.C., Jennings, G.M., Davis, S.D.: The limits to tree height. *Nature* 428, 851–854 (2004)
- Kok, R.: The production of insects for human foods. *Can Inst. Food Sci. Technol. J.* 16, 5–18 (1983)
- Kotake, S., Hijikata, K.: Numerical Simulations of Heat Transfer and Fluid Flow on a Personal Computer. Elsevier, New York (1992)
- Kurokawa, H., Sawa, T., Ashida, A., Hamano, N., Oguchi, M., Nitta, K.: Mineral Recovery Systems for Humans in a CELSS. *SAE Tech. Paper Ser.* No. 921237 (1992)
- Levine, L.H., Bisbeea, P.A., Richardsa, J.T., Birmelea, M.N., Priorb, R.L., Perchonokc, M., Dixond, M., Yorioa, N.C., Stuttea, G.W., Wheeler, R.M.: Quality characteristics of the radish grown under reduced atmospheric pressure. *Adv. Space Res.* 41, 754–762 (2008)
- Loveley, D.R.: Bug juice: harvesting electricity with microorganisms. *Nature Reviews Microbiol.* 4, 497–508 (2006)

- Maegawa, Y.: Photosynthesis and productivity. In: Notoya, M. (ed.) Utilization of Ulva and its use for remediation of environment, Seizando-shoten, Tokyo, ch. 2.3 (2001)
- Mann, K.H.: Seaweeds: Their Productivity and Strategy for Growth: The role of large marine algae in coastal productivity is far more important than has been suspected. *Science* 182, 975–981 (1973)
- Mitsuhashi, J.: Use of Insects in Closed Space Environment. *Biol. Sci. Space* 21, 124–128 (2007)
- Mitsuhashi, J.: People who eat insects, Heibon-sha, Tokyo (1997)
- Nagatomo, M., Hashimoto, Y.: Possibility to use Wooden Materials for Human Space Activities. Presented at 10th International Space Conference of Pacific-basin Societies, C5-04, Tokyo (2003)
- Nagatomo, M.: Experimental study on growth of young trees under the pressure of one tenth of earth atmosphere. *J. Space Tech. Sci.* 21(2), 11–26 (2005)
- Nitta, K., Otsubo, K., Ashida, A.: Integration test project of CEEF—A test bed for closed ecological life support systems. *Adv. Space Res.* 26, 335–338 (2000)
- Noda, S., Ohkuma, M., Usami, R., Horikoshi, K., Kudo, T.: Culture-Independent Characterization of a Gene Responsible for Nitrogen Fixation in the Symbiotic Microbial Community in the Gut of the Termite *Neotermes koshunensis*. *Appl. Environ. Microbiol.* 65, 4935–4942 (1999)
- Notoya, M.: Utilization of Ulva and its use for remediation of environment. Seizando-shoten, Tokyo (2001)
- Oshima, T., Moriya, T., Kanazawa, S., Yamashita, M.: Proposal of Hyperthermophilic Aerobic Composting Bacteria and Their Enzymes in Space Agriculture. *Biol. Sci. Space* 21, 121–123 (2007)
- Roberts, S.P., Harrison, J.F., Dudley, R.: Allometry of kinematics and energetics in carpenter bees (*Xylocopa varipuncta*) hovering in variable-density gases. *J. Exp. Biol.* 207, 993–1004 (2004)
- Schaafsma, G.: The Protein Digestibility-Corrected Amino Acid Score. *J. Nutr.* 130, 1865S–1867S (2000)
- Shimizu, K.: Effect of salt treatments on the production and chemical composition of salt wort (*Salicornia herbacea L.*), redgrass and alfalfa. *Jpn. J. Trop. Agr.* 44, 61–67 (2000)
- Shimoda, T., Kodera, A., Nakahara, T., Agarie, S., Nose, A.: NaCl accumulation capacity of ice plant, *Mesembryanthemum crystallinum*. *Jpn. J. Crop. Sci.* 72, 224–225 (2003)
- Taiz, L., Zeiger, E.: Plant Physiology, ch. 4, 3rd edn. Sinauer Associates, Massachusetts (2002)
- Takahashi, E.: Circulation of minerals. In: Scope of Space Biology and Agriculture. ISAS, Tokyo (1987)
- Tanatsugu, N., Yamashita, M., Mori, K.: A Conceptual Design of a Solar-Ray Supply System in the Space Station. *Space Solar Power Review* 5, 221–230 (1985)
- Ueda, K., Saka, H., Ishikawa, Y., Kato, T., Takeshita, Y., Shiratori, H., Ohno, M., Hosono, K., Wada, M., Ishikawa, Y., Beppu, T.: Development of a membrane dialysis bioreactor and its application to a large-scale culture of a symbiotic bacterium, *Symbiobacterium thermophilum*. *Appl. Microbial. Biotechnol.* 60, 300–305 (2002)
- Ushakova, S.A., Kovaleva, N.P., Gribovskaya, I.V., Dolgushov, V.A., Tikhomirova, N.A.: Effect of NaCl concentration on productivity and mineral composition of *Salicornia europaea* as a potential crop for utilization NaCl in LSS. *Adv. Space Res.* 36, 1349–1353 (2005)
- Wada, H.: Roles of peat in realizing agriculture on Mars. *J. Jpn. Peat. Soc.* 7, 21–29 (2008)

- Watanabe, I., Liu, C.C.: Improving nitrogen-fixing systems and integrating them into sustainable rice farming. *Plant and Soil* 141, 57–67 (1992)
- Yamashita, M., Arai, M., Ishii, C., Ishikawa, Y., Izumi, R., Oshima, T., Oshima, H., Omori, K., Katayama, N., Katayama, A., Kanazawa, S., Kariya, T., Kawasaki, Y., Kitaya, Y., Goto, E., Saito, T., Shimizu, T., Shiraishi, A., Takaoki, M., Takahashi, H., Tani, A., Toki, K., Tomita-Yokotani, K., Nakajima, A., Nagatomo, M., Nitta, K., Hashimoto, H., Hirafuji, M., Fujii, Y., Mizutani, H., Mihara, K., Miyagawa, T., Mukai, C., Mori, S., Yano, S., Yamazaki, N., Yokota, H.: Conceptual study of space agriculture. *Space Utiliz. Res.* 21, 323–326 (2005a)
- Yamashita, M., Ishikawa, Y., Oshima, T.: Space Agriculture Saloon, Engineering issues of microbial ecology in space agriculture. *Biol. Sci. Spac.* 19, 25–36 (2005b)
- Yamashita, M., Ishikawa, Y., Nagatomo, M., Oshima, T., Wada, H.: Space Agriculture Task Force. *J. Space Tech. Sci.* 21(2), 1–10 (2005c)
- Yamashita, M., Ishikawa, Y., Kitaya, Y., Goto, E., Arai, M., Hashimoto, H., Tomita-Yokotani, K., Hirafuji, M., Omori, K., Shiraishi, A., Tani, A., Toki, K., Yokota, H., Fujita, O.: An Overview of Challenges in Modeling Heat and Mass Transfer for Living on Mars. *Ann. NY Acad. Science* 1077, 232–243 (2006)
- Yamashita, M., Katayama, N., Hashimoto, H., Tomita-Yokotani, K.: Space Agriculture Task Force Space Agriculture for Habitation on Mars—Perspective from Japan and Asia. *J. Jpn. Soc. Microgravity Appl.* 24, 340–347 (2007)
- Yamashita, M., Ohya, H., Nitta, K., Yatazawa, M.: Sodium and potassium recycle in closed ecosystem-space agriculture. *J. Jpn. Soc. Aeron. Space Sci.* 33, 288–296 (1985)
- Yamashita, M., Tomita-Yokotani, K., Hashimoto, H., Sawaki, N., Notoya, M.: Sodium and Potassium Uptake of *Ulva* – Application of Marine Macro-algae for Space Agriculture. *Adv. Space Res.* 43, 1220–1223 (2009)
- Yazawa, Y., Saito, M., Takeda, H.: Variation between Terrestrial and Lunar or Martian Desert - Surface Materials on Mars and its Possible Use for Agriculture. *Biol. Sci. Space* 21, 129–134 (2007)
- Yokota, H., Ishikawa, Y., Yamashita, M., Oshima, T.: Space Agriculture Task Force. Space agriculture on Mars using hyper-thermophilic aerobic bacteria. *Habitation* 10, 191 (2006)

Chapter 19

Utilization of Regolith for Manufacturing Construction Material on Mars

Yoji Ishikawa

Mars Habitation Institute, Yokohama, Japan

19.1 Habitat on Mars

19.1.1 Requirement for Habitat

Immediately upon arrival to the Martian surface, Martian travelers or settlers require some type of habitat on the Martian surface. First of all habitat is necessary for respiration with air pressure maintained inside, for shielding against cosmic radiation, and in case of emergency, for reserve as a safe heaven.

Air-tightness is the first factor to be considered to separate living quarter air from Martian atmosphere. Since any material or membrane cannot be perfect, leakage is unavoidable, and then it becomes important to provide reliable air revitalization and Martian air processing systems.

Volume of living area in habitat is another important factor. After long interplanetary journey while living in tiny space vehicle that cannot be comfortably commodious, the Martian housing should be planned as being roomier in the psychological point. At the initial stage of our past space travel, astronaut is allowed to occupy merely a space of a few or several cubic meters per person. Then, roughly $10\text{ m}^3/\text{person}$ is given to Space Shuttle astronaut, and $100\text{ m}^3/\text{person}$ to Space Station astronaut. In an architectural view, our knowledge based on the experiences in confined space such as submarine or underground office suggests that people is able to perform a simple work in the space volume of 10 m^3 and their tolerable limit of size is as small as 5 m^3 .

Radiation is a serious harm to be protected from, and micrometeorite might be thought to be another hazard source for human. On Moon where actually no atmosphere exists, micrometeorites reach the surface and can be a significant hazard. On Mars, however, the atmosphere has a column mass of 3 g/cm^2 and will effectively shield against micrometeorites. Anything incoming meteorite large enough to reach the surface of Mars intact is too big to be stopped by a habitat. Fortunately they are rare.

Cost is also important. In-Situ Resource Utilization (ISRU) must be a key to reduce transportation cost.

19.1.2 Type of Habitat

Habitats are classified into three types, i.e., hard shell, soft membrane, and ISRU types (Ishikawa 1990). Details are given in Table 19.1.

Hard shell type habitat can be, for example, described as a metal cylindrical shape structure, and is in most cases common with habitat in interplanetary space vehicle. This most basic type habitat can be utilized at a beginning of a Mars habitation stage. It is robust and safe against harmful radiation. It becomes a safe-heaven especially when covered with Martian regolith. However, disadvantages of this type are its small size and the high transportation cost.

Table 19.1 Habitat Types

Type of Habitat	Characteristics
Hard Shell Type	basic type in common with habitat of space vehicle robust and safe against incoming disturbance such as sand storm, or radiation used as safe-heaven
Soft Membrane Type	better cost performance needs protection against incoming disturbance larger, comfortable space provided
ISRU Type	necessary to lower cost in long term regolith as covering material, for structural material,or for chemical processing

Soft membrane type habitats create a larger space for living area. Even from the first mission of manned flight, this type of habitat should be introduced onto Martian surface because of psychological reason. Membrane structures assure the largest space with a minimum cost. Even if the view of the Martian surface eases people psychologically, a larger living space is important. The way to reduce radiation penetration is required.

Due to the distance between Earth and Mars, In-Situ Resource Utilization (ISRU) is definitely required. Although any resource processing is not considered or realized at the earlier stage of the exploration, people will be determined to proceed towards in-situ manufacturing. For a habitat, there are several ways to apply ISRU. The easiest way is to use regolith itself as it is. Regolith will cover habitat, either hard shell or soft membrane type, to prevent radiation. Counterweight for crane is another application using regolith its other use when constructing and moving heavy habitat parts. The second stage of ISRU is a manufacturing structural material from regolith. It includes brick or sintered material. Chemical processing of regolith will be realized as

the third stage. These materials do not necessarily assure air-tightness, however, at least can be used for protection from radiation and also as structural materials.

19.2 ISRU Construction Materials on Mars

Several kinds of construction materials through ISRU for use in space have been proposed, i.e., concrete, glass, glass fiber, glass composite, and aluminum or iron metal. The simpler method is cast basalt method (Binder al. 1990) which makes glassy material as a product by melting and cooling regolith. As an ISRU construction material on Mars, duricrete was proposed (Boyd 1989). On Earth, a similar process is known as a cold press method. The principle of this method is to bind clay mineral particles by pressing under the presence of water using two necessary ingredients, clay mineral and water.

19.3 Available Resources on Mars

19.3.1 Water

It has been clear that water exists on Mars; in atmosphere, subsurface, regolith, and polar caps. Sole polar region might hold water of Water Equivalent to a Global layer (WEG) 11 m deep (Plaut et al. 2007), and Phoenix Mars Lander has confirmed the presence of water even near the landing surface.

19.3.2 Regolith

The detailed data of Viking data revealed that the Martian sediment holds >1-3 % of water content and 1150 -1600 kg/m³ of density (Gooding 1992). Martian primary minerals are believed to be igneous rocks such as quartz, pigeonite, augite, ortho-pyroxene, olivine, plagioclase, and titanomagnetite (Baird 1976, Clark 1982, Wright 1988, Ming 1993, Ishikawa 1996). The igneous rocks have been chemically weathered under the presence of water in the Early Noachian era (4.6 to 4.1 Ga) and then the secondary clay minerals mainly of phyllosilicate group, smectite, have been created (Chevrier 2007). Thus, the clay mineral which can be considered to be raw material for Martian brick is found abundantly on Martian surface.

19.4 Manufacturing Martian Brick

As one of methods to manufacture Martian brick, a cold press method has been experimentally demonstrated to examine applicability to Mars (Ishikawa 1992). The objectives of this demonstration have been:

- (1) To verify that construction materials can be actually manufactured from Martian regolith simulants. For this purpose, these simulants must represent the proper characteristics of Martian regolith.

(2) To investigate how useful and efficient the examined methods are for their implementation on Mars. To this end, the efficiencies of these methods are analyzed.

Martian regolith is not yet clearly defined as mentioned above, however it is believed to consist of a basaltic type material and a Smectite type clay mineral. In this experiment, a bentonite, which primarily consists of Monmorillonite, is used as the Martian regolith simulant.

A cold press method is basically the method to produce a construction material by pressing a mixture of sand, clay mineral, and water without heating. This method only requires a mold to hold the regolith mixture and a machine to apply a constant load. As to the load applied to a sand mixture, 9.8 MPa is applied whenever it is possible. In some cases, the loads smaller than 9.8 MPa are applied with monitoring the condition of the pressed regolith mixture to avoid the break downs of molds. The regolith mixture, after being compressed, is removed from the mold and stored in a plastic bag until its compressive strength is measured (Brook 1985). The experiments have been carried out for different mixture ratios of sand, clay, and water within the regolith mixture. The results are shown in Table 19.2.

The results of the compressive strength indicate that the regolith mixture which includes 20 % water has a higher strength than that of the 30 % water content.

To investigate the effect of dry curing process, some specimens are left dry after pressed. Results are shown in Table 19.3, where in all the cases the water content before pressing is kept at 20%. It is clear that the specimen, when cured in a dry environment, demonstrates a higher strength than non-dried samples. Also, it should be noted that the richer in bentonite the specimen is, the higher strength is. The maximum strength is 7.39 MPa.

This strength does not fulfill the Earth's general standard of concrete for road use. However under Martian gravity it may be used as a primary structural material.

The Martian regolith simulant is shown in Fig. 19.1. Here bentonite is mixed with crushed basalt and fine particles of iron oxide which makes the simulant reddish. Then water is added to the mixture before processing. Figure 19.2 shows the simulant packed in pressing mold of 10 cm x 10 cm x 10 cm. The simulant is pressed by pressing machine with up to 30 MPa, and then the red brick is made (Fig. 19.3).

Table 19.2 Cold Press Method – Pressing Pressures and Strength Results (The figures are average of several runs, modified from Ishikawa et al. 1992)

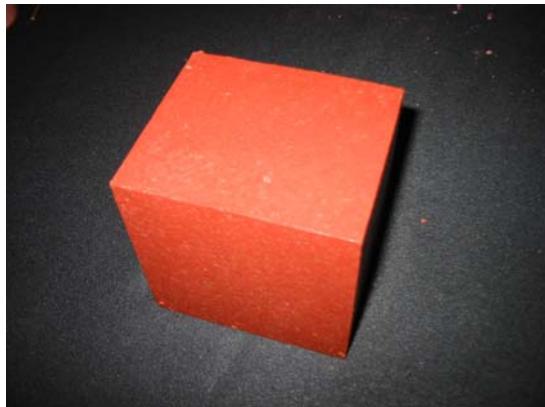
Regolith Mixture before Mixing			Strength Results		
Bentonite Content (%)	Sand (Silica) Content (%)	Water Content (weight % of Bentonite)	Density (g/cm ³)	Water Content (%)	Uniaxial Compressive Strength (MPa)
100	0	20	1.775	28	1.92
		30	1.757	38.4	0.36
50	50	20	1.888	12.7	1.45
		30	2.032	17.6	1.12
70	30	20	1.946	19.8	2.04
		30	1.945	25.5	0.74

Table 19.3 Cold Press Method Experiment to Examine the Effect of Dry Curing (The figures are average of several runs, modified from Ishikawa et al. 1992)

Regolith Mixture before Pressing			Curing Condition	Strength Results		
Bentonite Content (%)	Sand (Silica) Content (%)	Water Content (weight % of bentonite)		Density (g/cm ³)	Water Content (%)	Uniaxial Compressive Strength (MPa)
50	50	20	Dry	1.953	13.4	5.67
			Non-Dry	1.896	13.4	1.45
70	30	20	Dry	1.963	19.6	7.39
			Non-Dry	1.992	19.6	2.00
30	70	20	Dry	1.784	7.5	1.33
			Non-Dry	1.810	7.5	0.87

Fig. 19.1 Martian Regolith Simulant**Fig. 19.2** Martian Regolith Simulant packed in Mold

Fig. 19.3 Martian Brick made by Cold Press Method from Martian Regolith Simulant



19.5 Application

Brick has been used for construction on Earth from an ancient age. House walls, fences, and bridges are typical structure. Martin bricks manufactured with cold press method can be mainly used for structural material. Air tightness cannot be expected to the bricks, though cold pressed bricks have dense and impermeable properties. The interesting way to use brick on Mars is a structural material for vault. Vault structure can be seen as bridges or arch shape gate on Earth, however, on Mars will be applied to the shielding cover structure of inflatable habitat. The brick-made vault can stand by itself with the compressive tension between bricks, it will give no load to the inflatable habitat which is placed inside.

The red brick used on Earth is said to have compressive strength of approximately more than 10 MPa whereas concrete more than 20 MPa. Though the Martin brick seems to be fragile at the Earth standard, however, under Martian gravity the load of the building will become smaller by roughly one third. Consequently, the strength of the Martian brick will become at equivalent level to the Earth concrete.

The cold press method uses the adhesive forces of the soil particles. The reason why the soil particles stick together at the cold press method is as follows. The major players of this method are Smectite type clay mineral and water.蒙脱土矿物是这种方法的主要成分。蒙脱土矿物是一种主要由层状硅酸盐-铝硅酸盐层组成的层状矿物，具有较大的阳离子交换容量。蒙脱土也具有吸收大量水分并膨胀的能力。当与适量的水混合并施加适当的机械压力时，蒙脱土的粘性性质会成为颗粒之间的结合力，使压成的块状物坚固耐用。因此，很明显，火星砖的制造方法与普通地球砖（例如， Adobe）不同。 Adobe是由粘土烧制或烧制而成的砖，其产品具有多孔性且具有不同的物理化学特性。

蒙脱土广泛应用于土木工程领域作为不透水浆液屏障材料。它还有效用于限制和隔离放射性物质。

material and has been a common material as a packing and sealing material for its confinement. The advantage of using such clay mineral is its long life. Concrete might deteriorate in hundreds years, however, clay mineral can maintain its impermeable property even in hundreds of thousands years because it is naturally occurring mineral.

One of the other uses of Martian brick is to prevent radiation exposure. The shielding ability depends on the total mass of the material that the radiation passes through. Rapp (2006) examined the effect of regolith as the shield media and discussed that the regolith has relatively little effectiveness in shielding against high energy components of space radiation on Martian surface, even though the Mars atmosphere reduces the radiation to some extends. It is shown in Rapp (2006) that the BFO dose equivalent of annual galactic cosmic radiation, after penetrating 7 g/cm² Martian atmosphere, is reduced from 40 cSv to 30 cSv with 20 g/cm² Martian regolith thickness, and to 25 cSv with 50 g/cm². If regolith itself is supposed to have bulk density of 2 g/cm³, 50 g/cm² becomes equivalent to around 25 cm thickness. Although the thicker shield is expected to reduce radiation even more, it is unlikely to build thick layer of regolith shield around habitat. Martian brick, on the other hand, can be piled up to wall or vault shield thick enough to reduce the radiation level inside the habitat. It is also obvious that the required thickness of the brick shield is shorter compared to the regolith because of the heavier weight of Martian brick. Thus, it is likely that all the Martian habitats will be bricked in with walls or vaults structure.

19.6 Conclusion

Humans on Mars require a habitat, and In-Situ Resource Utilization (ISRU) will be required to lower cost for construction of the habitat. Regolith, which is abundant on the Martian surface, can be utilized either as it is or as a raw material to manufacture concrete, glass, metal, and brick. As one of such manufacturing processes using regolith, a cold press method is a simple and effective method to make brick by just pressing a mixture of regolith and water without heating. The manufacturing process is experimentally examined and the feasibility on Mars is assured. It is concluded that the Martian brick can be used as construction material and shielding block to protect a Martian habitat from radiation. The Martian brick will be one of the plausible materials for constructing habitat on Mars in a future.

References

- Baird, A.K., Toulmin, P., Clark, B.C., Rose, H.J., Keil, K., Christian, R.P., Gooding, J.L.: Mineralogic and petrologic implications of Viking geochemical results from Mars; Interim Report. Science 194, 1288–1293 (1976)
- Binder, A.B., Culp, M.A., Toups, L.D.: Lunar Derived Construction Materials; Cast Basalt. Engineering, Construction, and Operations in Space II. In: Wetzel, J.P. (ed.) Proceeding of Space 1990, Albuquerque, New Mexico, April 22–26, 1990, pp. 117–122. American Society of Civil Engineers Aerospace Division (1990)

- Boyd, R.C., Thompson, P.S., Clark, B.C.: Duricrete and Composites Construction on Mars.
In: Stoker, C. (ed.) *The Case for Mars III*, Boulder, Colorado, July 18-22, 1987.
American Astronautical Society (1989)
- Brook, N.: The Equivalent Core Diameter Method of Size and Shape Correction in Point Load Testing. *Int. J. Mech. Sci. & Geomech. Abstr.* 21, 61–70 (1985)
- Clark, B.C.: Chemical composition of Martian fines. *J. Geochemical Research* 87(B12), 10059–10067 (1982)
- Gooding, J.L.: Soil mineralogy and chemistry on Mars: Possible clues from salts and clays in SNC meteorites. *Icarus* 99, 28–41 (1992)
- Ishikawa, Y., Kobayashi, K., Saito, T.: Martian Soil Analysis, Its Implication for Life on Mars. In: Chela-Flores, J., Raulin, F. (eds.) *Chemical Evolution: Physics of the Origin and Evolution on Life*, pp. 389–395. Kluwer Academic Publishers, Dordrecht (1996)
- Ishikawa, Y., Ohkita, T., Amemiya, Y.: Mars Habitation 2057: Concept design of a Mars settlement in the year. *J. British Interplanetary Soc.* 43, 505–512 (1990)
- Ishikawa, Y., Sasaki, T., Higasayama, T.: Simple and Efficient Methods to Produce Construction Materials for Lunar and Mars Bases. *Engineering, Construction and Operations in Space III*. In: Miller, R.J. (ed.) *Proceeding of Space 1992*, Denver, Colorado, May 31-June 4, 1992, pp. 1335–1346. American Society of Civil Engineers Aerospace Division (1992)
- Rapp, D.: Radiation Effects and Shielding Requirements in Human Missions to the Moon and Mars. *Mars* 2, 46–71 (2006)
- Wright, I.P., Grady, M.M., Pillinger, C.T.: Carbon, Oxygen, and Nitrogen Isotopic Compositions of Possible Martian Weathering Products in EETA 79001. *Geochimica et Cosmochimica Acta* 52, 917–924 (1988)

Chapter 20

Perspectives on the Utilization of Martian Iron and Carbon Dioxide

J. Marvin Herndon

Transdyne Corporation, San Diego, California, USA

20.1 Introduction

The idea of colonizing Mars has long intrigued both scientists and laymen as reflected in fiction and in non-fiction writings, in books, in magazine articles, in scientific journal articles, in movies, and in popular websites. The author's purpose in this chapter is not to review in detail the extensive, extant literature, but rather, from the perspective of a multi-disciplinary scientist with commercial mining experience, to share some thoughts pertaining to the utilization of two of the major Martian natural resources, iron and carbon dioxide, each of which has been discussed in the literature (Hepp et al. 1991; Landis 2009).

Mars is the only planet in the Solar System that is potentially feasible for self-sufficient human habitation, but the environment is climatically harsh (Jakosky and Phillips 2001; Leovy 2001) and, in several important ways directly related to atmosphere, ground, and underground, quite unlike Earth (Solomon et al. 2005). In any colonization activity, survival is paramount. Although humans have a long and reasonably well documented history of colonizing new areas of Earth, those experiences will be of little benefit for Mars colonizers. Generally, in Terrestrial colonization the first priorities are to secure shelter and sources of potable water and food. The same priorities hold for Martian colonization with the additional priority of securing a source of breathable oxygen. At least for Mars, though, there is as yet no evidence of hostile denizens that people must defend themselves against.

Colonized areas on Earth, even the most remote, are generally not too different from the places where the colonists originated. A notable exception is Antarctica, where virtually everything needed for survival, except air and water, must be continuously supplied and re-supplied. The great distance between Earth and Mars, varying between 57 and 400 million km, precludes major supply and re-supply. Unlike the Antarctic example, because of the great distances involved, the long

transit times, and the extended human exposure to cosmic radiation in transit, a Martian colony must be largely self-sufficient.

In any colonization effort anywhere, the best approach is to make the maximum use of materials at hand. The absence of trees, or in fact any vegetation, on Mars means that initial permanent shelters must be composed of rock, including rock composites, or must be constructed underground. The low tensile strength of rock without means for reinforcement places serious construction limitations on structural and building height above ground as well as depth below the ground. Iron is the most desirable material for structural reinforcement, as well as for machinery, and other uses, such as for counter weights. Fortunately, iron on Mars is known to be readily available over wide regions of the planet's surface (Christiansen et al. 2001). In fact, iron is potentially more abundant and more readily accessible on Mars than on the surface of Earth. The nature of the technology to be employed for mining, smelting, or otherwise processing Martian iron, as well as the milling or other forms of producing useful items will depend on the chemical state of iron resources and on the relative availability of energy for processing. Ultimately, though, the technology for utilizing Martian resources will depend as well and, in fact mainly, on the time-frame for initial colonization, the monetary resources made available, and the degree and extent of beforehand planning, research, and development.

In the course of past human development, the progression from Stone Age to Bronze Age to Iron Age occurred through hard-won step-by-step metallurgical understanding and related technological development, and occurred as a consequence of readily available raw materials, especially the raw materials necessary for the production of heat. At first wood from virgin forests was used, but that was replaced by the use of charcoal produced from wood by slow pyrolysis, heating in the absence of oxygen to drive off water and volatile hydrocarbons. Then coal replaced charcoal for firing iron furnaces, and then coal was ultimately replaced by coke, which is produced from coal by distilling off hydrocarbons. A very serious impediment to the utilization of Martian iron resources is the absence on Mars of plant matter, and, presumably, the complete absence of any reduced, elemental carbon that might resemble, or be able to be used in place of, coal.

Nevertheless, as serious efforts are made to prepare for establishing habitable, self-sufficient colonies on Mars, preparations should be made for a new technological age, the Martian Iron Age. But at present, preparations can at best be described as sketchy or highly idealized. Far too many fundamental factors are unknown and cannot be assumed with any reasonable certainty, such as available energy and water. Moreover, there are serious logistical problems. For example, very heavy equipment is required to move sand and rock on Mars, which would be extremely costly and time consuming to bring from Earth. Further, heavy-equipment breakdowns, especially common in windy, arid places, are difficult problems on Earth, but would be of astronomical difficulty on Mars because of the remoteness.

In this chapter, thoughts as to the utilization of Martian resources first address relatively near-term colonization, generally taking into account in-hand technology as much as possible. Against that framework, some thoughts are presented to suggest yet-undeveloped potential future technology.

20.2 Sources of Martian Iron

Meteorites land on Earth as well as on Mars. Some meteorites consist of nickel-iron metal combined with silicate minerals; others consist exclusively of nickel-iron metal. In ancient times, some of the iron meteorites found on Earth were directly hammered into crude hand tools (Coghlan 1941); a similar activity might take place on Mars, although with minuscule impact for colonization because of the scarcity and paucity of such deposits, and because the resulting implements would at best be quite crude, as meteoritic nickel-iron is hard but can be worked with the application of heat work from a forge like blacksmiths once commonly used.

As iron meteorites crash through the atmosphere, material spalls from the surface, producing small spheres, sometimes referred to as “cosmic spherules” (Park and Reid 1965). These may be expected to comprise a small fraction of Martian sands, and can be separated magnetically either directly from Martian sand, or, more practically, from density concentrates of Martian sands obtained by using devices similar to the “dry washers” employed by precious metal miners in arid regions of Earth to separate heavy minerals from sand without using water (Rancourt 1994).

As iron meteorites impact the Martian surface at high velocities, all or parts of the meteorite may melt and/or evaporate. These evaporates and melts cool quickly, condensing and solidifying, ending up either as small shards or spherules, which can be separated in the same way as “cosmic spherules” may be separated from Martian sand.

Small masses of iron, typically in the size range of 1-10 mm, such as “cosmic spherules” or shards, at high temperatures undergo rapid chemical reaction with the ambient atmosphere. Depending on the prevailing oxygen fugacity at the time chemical equilibrium is frozen-in, the iron may remain metallic or it may become an iron-oxide.

The nature of metallurgical processing of recovered Martian iron resources depends on the available energy, and on whether the recovered iron is combined with oxygen as the oxide (Sects. 20.3 and 20.4) or occurs as the native metal (Sects. 20.7 and 20.8).

Generally, with isolated rare exceptions, ore-grade iron on Earth occurs combined with oxygen as black magnetite (Fe_3O_4) or as usually red-hued hematite (Fe_2O_3). Surface ore-grade iron on Mars is the reddish-colored iron-oxide, which gives the planet its red color and the nickname *Red Planet*. Indeed, hematite spherules have been observed on Mars (Squyres et al. 2004), which if present in great numbers, would be a convenient source of iron-ore for processing into iron metal.

Utilization of iron for structural purposes, whether on Earth or Mars, necessarily involves reducing the ore to the metal by chemical means and then processing and fabricating into useful forms.

20.3 Reduction of Martian Iron Oxide by Carbon Monoxide

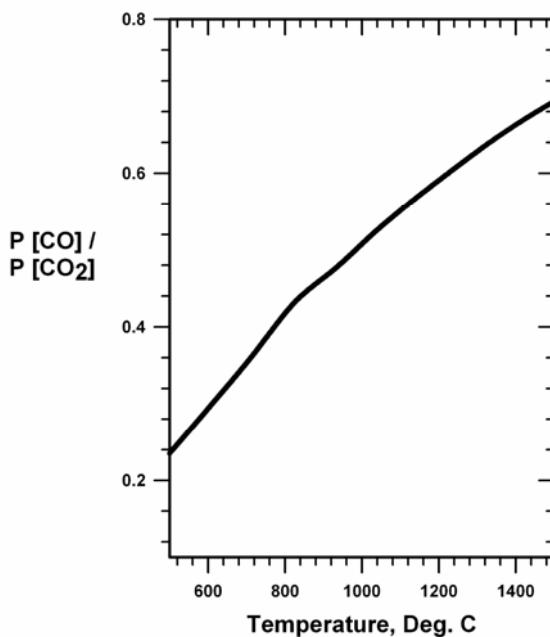
Industrially on Earth, iron-oxide is reduced to the metal mainly by heating with carbon, produced from wood or from coal. Limestone is added to facilitate producing

slag to scavenge impurities and, while floating on the surface, shields the molten iron from reaction with atmospheric oxygen (Brack 2008). In the blast furnace, oxygen (O_2) is forced into the mix and reacts with the carbon (C) to form carbon monoxide (CO) and to produce heat. Significantly, it is the carbon monoxide that reduces the iron-oxide (Fe_2O_3) to iron metal (Fe) by reaction Eq. (20.1), the thermodynamic boundary conditions of which are shown in Fig. 20.1:



Mars appears to have similar surface ore-grade iron oxide compounds as Earth and in relatively greater abundance. However, Mars is essentially devoid of elemental carbon. Consequently, iron metal production on Mars will require greater energy expenditure. Although Mars is virtually devoid of elemental carbon, the main constituent of the Martian atmosphere is carbon dioxide (CO_2) which can be utilized for processing iron.

Fig. 20.1 The pressure independent equilibrium ratio of the pressure of carbon monoxide to the pressure of carbon dioxide as a function of temperature for the reaction $3 \text{ CO} + \text{Fe}_2\text{O}_3 = 2 \text{ Fe} + 3 \text{ CO}_2$. Prompt removal of carbon dioxide drives the reaction to the right



The atmosphere of Mars is 95.7% carbon dioxide (CO_2), which is available everywhere at the surface and which can readily be compressed. The Martian carbon dioxide can be, simultaneously, a source of carbon monoxide (CO) and a source of oxygen (O_2) through the thermal decomposition reaction:



Reaction Eq. (20.2) requires a temperature of about 1100°C. Note that the iron-reduction reaction Eq. (20.1) can take place at about 700°C, but a higher temperature, such as 900°C, would be desirable. As Mars is truly remote and accessible only at great expenditure of time, energy and perhaps human lives, it would seem that the simplicity of the above approach would be a great advantage. The technological shortcoming of the approach, however, is that a simple, rugged, and fool-proof method of separating carbon monoxide and oxygen would need to be developed to work efficiently and successfully in that distant place. And, of course, there is a serious energy requirement to attain operating temperatures. In addition, extreme degradation of equipment occurs at those operating temperatures, necessitating frequent replacement.

Pure iron melts at 1535°C. Thus the reduced iron produced by reduction with carbon monoxide will not alone lead to the separation of iron from impurities. Consequently, the iron obtained through reaction Eq. (20.1) would necessarily need additional processing, such as described in Sects. 20.6 and 20.7.

20.4 Reduction of Martian Iron Oxide by Hydrogen

The carbon monoxide reduction methodology, described in Sect. 20.3, for using atmospheric carbon dioxide (CO_2) as feed for a system to reduce iron-oxide and produce oxygen (O_2), while having some advantages and some difficulties, is not the only available reduction reaction. Iron oxide can as well be reduced by hydrogen (H_2) through reaction Eq. (20.3), the thermodynamic boundary conditions of which are shown in Fig. 20.2:



The hydrogen necessary for reaction Eq. (20.3) can be produced either from the electrolysis of water by reaction Eq. (20.4):



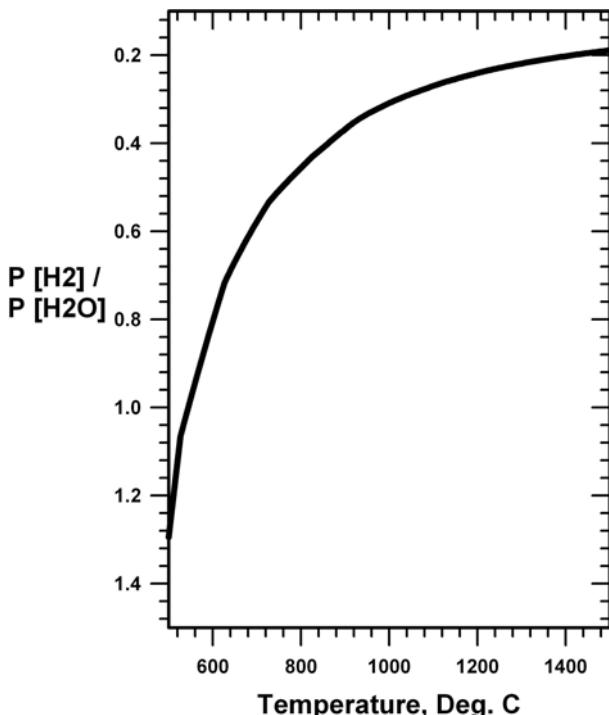
or by the reaction of carbon monoxide and water gas (H_2O), i.e.



The source of carbon monoxide for reaction Eq. (20.5) would come from reaction Eq. (20.2). In any remotely located technical activity, simplicity is truly the key to eventual success.

It is known from experience that reaction Eq. (20.3) requires for efficiency a minimum temperature of about 700°C, although 900°C is probably more practical from the standpoint of reaction rate and reaction completion. Recall that pure iron melts at 1535°C. Thus the reduced iron produced by reduction with hydrogen, like the reduced iron produced by reduction with carbon monoxide, by reaction Eq. (20.1), will not alone lead to the separation of iron from impurities. Consequently, the iron obtained either through reaction Eq. (20.3) or reaction Eq. (20.1) would necessarily need additional processing, such as described in Sects. 20.5, 20.6 and 20.7.

Fig. 20.2 The pressure independent equilibrium ratio of the pressure of hydrogen to the pressure of water gas as a function of temperature for the reaction $3 \text{ H}_2 + \text{Fe}_2\text{O}_3 = 2 \text{ Fe} + 3 \text{ H}_2\text{O}$. Prompt removal of gaseous water drives the reaction to the right



20.5 Smelting Martian Iron Metal

Most of the iron metal produced on Earth is reduced from the oxide and smelted in the same step. Recall that Terrestrial iron-oxide is reduced to the metal mainly by heating with elemental carbon. Limestone is added, which decomposes to yield calcium oxide (CaO), which then combines with silica (SiO_2) to form slag (CaSiO_3). The slag serves to scavenge impurities and to shield the surface of the molten iron from reaction with atmospheric oxygen. Heat produced by the reaction of carbon with oxygen helps to bring the temperature to above the melting point of iron. Iron, being denser than the slag, settles by the action of gravity. The advantage of this process is that the end-point iron is molten, into which additional alloying elements can easily be added to improve physical properties, and the iron produced can be directly processed into structural units, such as I-beams.

The virtual absence of elemental carbon on Mars precludes Earth-like iron smelting, although a variant might be undertaken. Martian iron produced by either reaction Eq. (20.1) or reaction Eq. (20.3) could in principle be smelted, using an electric furnace, provided slag-forming minerals are available to be mined and provided that sufficient energy sources are available to power the operation, which cannot depend upon energy produced by the oxidation of carbon. Aside from high energy demands, the high temperatures involved can pose serious technological problems, especially during the initial periods of Mars colonization.

20.6 Sintering Martian Iron Metal

On Mars there would be great advantages to be gained from being able to avoid temperatures near the melting point of iron (1535°C). Sintering is one mechanism of treating the iron produced from reactions Eqs. (20.1) and (20.3) to produce specific shapes and simultaneously impart improved physical properties. Sintering involves the application of heat and pressure to iron powder produced by either reaction Eq. (20.1) or reaction Eq. (20.3), or more ideally from the processes described in Sects. 20.7 and 20.8. The temperatures involved would be below the melting point of iron, possibly as low as 900°C. A protective or reducing atmosphere during processing is necessary to keep the iron from oxidizing.

If the iron-oxide being reduced via reaction Eq. (20.1) or reaction Eq. (20.3) is quite pure, sintering could be accomplished in the same reduction chamber, if piston-fitted for subsequent compression. If the reduced iron contains too many impurities, such as silica, after reduction the charge could be disrupted by tumbling, and jiggled to separate the iron from the gangue before sintering.

20.7 Production of Martian Iron Metal from Iron Carbonyl

As has been known since the 1920s, with subsequent and improvements (Unger et al. 2007), if iron metal is heated in the presence of carbon monoxide (CO) at a pressure above about 50 bars and at a temperatures in the range 100-200°C, the highly volatile organo-metallic compound, iron pentacarbonyl, $\text{Fe}(\text{CO})_5$ is formed by the reaction:



At 1 bar pressure, iron pentacarbonyl is yellowish-brown, readily volatile, liquid which becomes a gas at temperatures above 103°C. At present, iron carbonyl has commercial applications as a catalyst for organic chemical reactions and as a battlefield obscurant.

A classic laboratory demonstration of the thermal decomposition of iron pentacarbonyl involves passing vapors of the compound, carried along by an inert carrier gas such as nitrogen, through a glass tube to which, when heat is applied (ca. 170°C), causes iron to be deposited on the hot inner walls of the glass tube. The metallic iron deposition is notable because it happens quickly at a low temperature and because it produces a mirror-like reflective surface. Thermal decomposition of iron pentacarbonyl is the means by which iron-product fabrication might be achieved on Mars (Jenkin 1990).

The starting point for producing iron pentacarbonyl is iron metal, not iron oxide. The iron metal produced by either reaction Eq. (20.1) or reaction Eq. (20.3) would be appropriate for producing iron pentacarbonyl, provided oxidizers, such as H_2O are first purged.

Because iron pentacarbonyl is liquid, with a relatively high vapor pressure, and because it readily decomposes on surfaces with temperatures as low as 170°C, it would seem to be the ideal medium for fabricating iron structures on Mars. Because of the relatively low temperatures involved, material handling and fabrication techniques can be extensively developed on Earth.

Iron pentacarbonyl should not be exposed to air or sunlight and can be stored under water. Interestingly, the $\text{Fe}(\text{CO})_5$ molecule is unstable with respect to sheer stresses; the author once destroyed an expensive piston vacuum pump by admitting iron pentacarbonyl vapor into the vacuum cylinder. A note of caution: Iron pentacarbonyl is toxic, causing lung irritation, toxic pneumonitis, and pulmonary edema.

20.8 Production of Iron and Nickel Carbonyl from Martian Meteorites

Martian meteoritic metal, which contains nickel, in the form of “cosmic spherules” or tiny shards, provided that its abundance is sufficiently great for mining, might be used as feed stock for making both iron pentacarbonyl, according to reaction Eq. (20.6), and nickel carbonyl, $\text{Ni}(\text{CO})_4$, by the following reaction:



At ambient pressure, nickel carbonyl is a gas at temperatures above 40°C. The author is unaware of any experimental, simultaneous production of iron pentacarbonyl and nickel carbonyl from meteoritic metal, or what potential difficulties or benefits might be encountered. As nickel carbonyl is considerably more volatile than iron carbonyl, separation of the two, in principle at least, should pose no difficulty. Another note of caution: Nickel carbonyl is considerably more toxic than iron pentacarbonyl.

20.9 Effect of Uncertainty

What is the best approach for utilizing Martian iron and carbon dioxide resources? The answer to that question depends critically upon information yet unknown. Consider first these critical unknowns:

- (1) Is there a sufficient quantity of appropriately-sized meteoritic nickel-iron metal available?
- (2) Can meteoritic nickel-iron metal efficiently be converted into iron carbonyl and nickel carbonyl?

The answer to (2) can be obtained on Earth from laboratory studies, but the answer to (1) is not known. If the answers to (1) and (2) were both “yes”, then, to this author, the preferred approach for processing Martian iron would be to use meteoritic nickel-iron metal in reactions Eqs. (20.6) and (20.7) simultaneously. After separating the iron carbonyl and nickel carbonyl, these could be thermally decomposed either to plate surfaces or to produce metal powders for sintering.

If, on the other hand, meteoritic nickel-iron metal resources were insufficient, or if some insurmountable problem were found that would render impossible the production of carbonyl compounds from meteoritic nickel-iron metal, then it seems to the author that the best approach for utilizing Martian iron and carbon dioxide resources is through reactions Eqs. (20.1), (20.2), and (20.6). In other words, from atmospheric carbon dioxide, through reaction Eq. (20.2), produce

oxygen for breathing and carbon monoxide for iron-making. Reduce iron oxide with carbon monoxide through reaction Eq. (20.1) to produce iron metal. Use that iron metal with carbon monoxide to produce iron pentacarbonyl through reaction Eq. (20.6). Then, make use of the thermal decomposition of iron pentacarbonyl vapor to produce iron metal objects or to produce iron powder for sintering. One major technological shortcoming of the approach, however, is that a simple, rugged, and foolproof method of separating carbon monoxide and oxygen would need to be developed to work efficiently and successfully in the great remoteness and inaccessibility of the Martian Surface. The other major shortcoming relates to energy availability.

One of the greatest uncertainties in planning for Martian iron production is the uncertainty surrounding the availability of appropriate energy sources. Nuclear energy sources, based upon the heat of natural radioactive decay, might be sufficient for powering scientific instruments, but not for processing iron, even on a relatively modest scale. Utilizing solar energy for that purpose is even more impractical than it would be on Earth, due to the lower intensity of incident sunlight and because of the technological difficulties posed by Martian sandstorms. Similarly, utilizing wind energy for iron production is impractical due to the low atmospheric pressure on Mars. It is not known whether the Martian equivalent of geothermal energy might be available, but from Terrestrial experience it would seem at best to be insufficient.

To this author, one known way of obtaining the power required for driving a Martian Iron Age would be if fully functional and fueled nuclear fission reactors are brought from Earth. But, because as yet there is no evidence of uranium ore deposits on Mars, self-sufficiency would be an impossible; Mars colonies would require continuing and expensive re-supply from Earth. There are, however, other solutions, described in Chaps. 7, 9 and 10.

20.10 Perspectives

The Internet hosts a variety of websites and blogs devoted to Mars and, especially, to dreams of Martian colonization. And, while it is human nature to “dream the impossible dream”, it is of benefit to pause and reflect a bit on what might be gained and the price to be paid. Robotic exploration of Mars is not only doable, but it is an on-going, reasonably successful, endeavor. So, one might ask, why and under what circumstances would planning for the establishment of a permanent, self-sufficient colony on Mars be justified?

Some might argue that priorities at home should come first. What justification can be made to plan for Martian colonization while there is homelessness, hunger, and disease here on Earth?

Recall that Mars is the only planet in the Solar System that is potentially feasible for self-sufficient human habitation. If Earth were to face impending doom, from plague, nuclear holocaust, asteroid impact, or as the result of some other catastrophic natural or man-made disaster, Mars may be the only location where members of the human species might be able to survive. The prospect of impending humanity-doom may be remote, a very low-probability event, but an event

with non-zero probability. From the standpoint of human survival, one-way transport to Mars may be the only possibility. In other words, a Martian colony would of necessity need to be entirely self-sufficient, totally without the possibility for re-supply. Planning therefore should be focused upon non-re-supply, which has profound implications with respect to the utilization of Martian iron and carbon dioxide resources.

Producing carbonyl from meteoritic nickel-iron metal appears, from energy requirements, to be the only potentially feasible mechanism for making iron and nickel metal on Mars without re-supply from Earth. Although the amount of available meteoritic nickel-iron metal on the Martian surface is unknown, iron meteorites are known to exist there (Landis 2009). Laboratory experiments to produce iron carbonyl and nickel carbonyl from meteoritic nickel-iron metal are not expensive to make. In fact, only one scientist with modest equipment is required. So, from the standpoint of Mars' colonization, this is where emphasis should be placed for the production on Mars of iron and nickel metal. But there is a conflict in that concept, a paradox. In general, it would seem that any self-sufficient colonization on Mars, based upon iron reinforcements and iron structural members, will require substantial quantities of high tensile strength, high-carbon iron-alloys not only for construction, but for farming, mining, transportation, and manufacturing. The structural strength of steel arises from the exsolution of iron carbide (Fe_3C) lamellae, called pearlite. But to add carbon to iron to produce high tensile strength steel is a high temperature, high energy-consuming process. The high energy requirements of Martian steel production are inescapable.

From the long-term prospect of establishing a fully self-sufficient, non-re-supply, colony, the author suggests considering bamboo as an alternative to iron metal as structural reinforcement for construction materials (Ghavami 1995). Of course, that would mean "engineering" a species of bamboo that could grow well first in Martian greenhouses, and, perhaps later, outside in the Martian environment. Such a prospect is really not as extreme as it may seem at first. If humanity is to be able to exist in wholly self-sufficient colonies without the possibility for re-supply from Earth, Mars-adapted plants would be an absolute necessity. In addition to providing food and structural reinforcement material, plants, as is well known, utilize carbon dioxide and give back in return breathable oxygen. While the need for establishing self-sufficient Martian colonies to save humanity may never arise, preparation, especially development of Mars-adapted botanicals, may have the fringe benefit of helping humans to create new plant species to utilize excess carbon dioxide on Earth and to make up for reduced oxygen production that has resulted from denuding our forests and reducing the phytoplankton content of our oceans.

References

- Brack, H.G.: Handbook for Ironmongers. A Glossary of Ferrous Metallurgy Terms: A Voyage through the Labyrinth of Steel- and Toolmaking Strategies and Techniques 2000 BC to 1950. Pennywheel Press, Hulls Cove (2008)
- Christiansen, P.R., Morris, R.V., Lane, M.D., Bandfield, J.L., Malin, M.C.: Global mapping of Martian hematite mineral deposits: remnants of water-driven processes on early Mars. *J. Geophys. Res.* 107(E10), 23873–23885 (2001)

- Coghlan, H.H.: Prehistoric iron prior to the dispersion of the Hittite Empire. *Man* 41, 74–80 (1941)
- Ghavami, K.: Ultimate load behavior of bamboo-reinforced lightweight concrete beams. *Concrete & Cement Composites* 17, 281–288 (1995)
- Hepp, A.F., Landis, G.A., Kubiak, C.P.: Chemical Approaches to Carbon Dioxide Utilization for Manned Mars Missions. NASA Technical Memorandum TM-104405 (1991)
- Jakosky, B.M., Phillips, R.J.: Mars' volatile and climate history. *Nature* 412, 237–244 (2001)
- Jenkin, W.C.: Process for coating a metal substrate by chemical vapor deposition using a metal carbonyl, US Patent No. 4,938,999 (1990)
- Landis, G.A.: Meteoritic steel as a construction resource on Mars. *Acta Astronautica* 64(2–3), 183–187 (2009)
- Leovy, C.: Weather and climate on Mars. *Nature* 412, 245–249 (2001)
- Park, F.R., Reid, A.M.: A comparative study of some metallic spherules. *Ann. NY Acad. Sci.* 119, 250–281 (1965)
- Rancourt, V.: Method and apparatus for separating heavy particles from particulate matter, US Patent no. 5,407,079 (1994)
- Solomon, S.C., Aharonson, O., Aurnou, J.M., Banerdt, W.B., Carr, M.H., Dombard, A.J., Frey, H.V., Golombek, M.P., Hauck II, A.A., Head III, J.W., Jakosky, B.M., Johnson, C.L., McGovern, P.J., Neumann, G.A., Phillips, R.J., Smith, D.E., Zuber, M.T.: New perspectives on ancient Mars. *Science* 307, 1214–1220 (2005)
- Squyres, S.W., et al.: The Opportunity Rover's Athena Science Investigation at Meridiani Planum. *Science* 305, 1698–1703 (2004)
- Unger, G., Schönenfelder, H., Watzenberger, O., Sterzel, H.-J., Kühling, K.: Preparation of Iron Pentacarbonyl, US Patent No. 7,276,220 B2 (2007)

Chapter 21

Mars Aqueous Processing System

Mark Berggren, Robert Zubrin, Cherie Wilson, Heather Rose, and Stacy Carrera

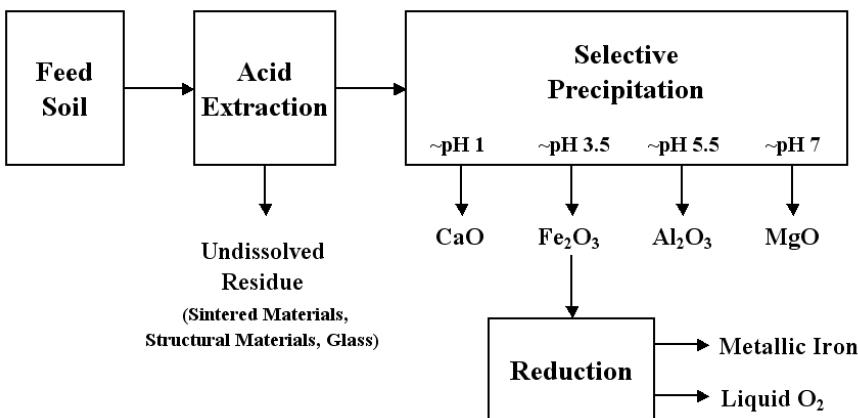
Pioneer Astronautics, Lakewood, Colorado, USA

21.1 Introduction

The Mars Aqueous Processing System (MAPS) is a novel technology for producing iron, high-grade metal oxides, and oxygen from Mars soils. These materials are suitable for use as-is and are also excellent feeds for further refining to produce steel, ceramics, glass, and light metals. The closed-loop process employs sulfuric acid extraction of soil followed by selective recovery of constituents by crystallization or precipitation initiated by pH adjustment using magnesium hydroxide base. Process reagents can be prepared on Mars from magnesium sulfate salts and water. Oxygen is recovered as a byproduct of iron oxide reduction. Alumina, magnesia, and calcia products (at greater than 94 percent purity) are useful refractory materials for manufacture of insulation, furnace liners, casting molds, and structural components. MAPS soil extraction residues are elevated in silica content and are suitable for preparation of thermally-fused materials and glass. MAPS residues can also be used to produce water-based structural materials such as salt- or magnesia-bonded concrete. The MAPS technology would be implemented after an initial human presence is established and fundamental *in-situ* resource utilization technologies have been demonstrated. MAPS will enable fabrication of complex structural components for habitats, machinery, spare parts, and other materials to support human self-sufficiency on Mars.

The MAPS flow sheet was developed with a focus on low-severity processing so that many of the tanks, piping, valves, and other hardware can be fabricated from lightweight, corrosion-resistant plastics or polymers. Only small amounts of process concentrates are subjected to high-temperature treatments, thereby reducing Earth-supplied equipment mass and minimizing power input for furnaces. The unit operations are designed to use as much common equipment as possible. Figure 21.1 shows a simplified schematic of the MAPS process.

Thermodynamic evaluations, laboratory experiments, and process modeling were carried out in support of a preliminary engineering design for the Mars Aqueous Processing System. The preliminary engineering design includes mass and energy balances, estimated hardware and reagent masses, reagent make-up requirements, and

**Fig 21.1** Mars Aqueous Processing System schematic

consumables mass. Each unit operation was validated through experimentation at the specified process conditions using NASA JSC Mars-1 soil simulant. Multi-cycle, closed-loop extraction and recovery experiments were conducted in a scale-model column system. Metallic iron as well as high-grade calcia, alumina, and magnesia were produced from Mars soil simulant. Other products including glass and concrete-like materials were also prepared from Mars soil simulant extraction residues.

Table 21.1 summarizes the MAPS baseline design rates and product purities. Data are based on laboratory results applied to a typical Mars soil. The baseline design assumes 100 kilogram feed soil batches with a 12.5-hour process cycle time (70,080 kilograms annual soil feed).

Table 21.1 MAPS Production Summary (For 100 kg Soil Feed Batches).

Product	Yield [kg per 100 kg feed soil]	Production Rate [kg/year]	Purity [%]
Metallic Iron Product	12.1	8,470	85.2 % Fe ^a
CaO Product	0.9	620	98.0 % CaO
Al ₂ O ₃ Product	4.9	3,430	90.1 % Al ₂ O ₃
MgO Product	3.5	2,460	91.9 % MgO
O ₂	4.4	3,070	100.0 % O ₂
Total	25.8	18,050	

^a When utilized directly as a sintered product; impurities can be removed using molten blast-furnace methods.

The electrical energy input per batch is about 75 kilowatt-hours at a peak electrical power demand of 16 kilowatts. The thermal energy input per batch is about 150 kilowatt-hours at a peak thermal power demand of 39 kilowatts. Table 21.2 summarizes the key design and operating specifications.

The breakeven time for MAPS at the design throughput is about 50 days (based on the time required to generate a combined product mass equal to the initial

Table 21.2 MAPS Design Specifications.

Parameter	Value
Hardware Mass [kg]	1,300
Start Up Reagent Mass [kg]	660
6-Month Reagent Make Up Mass [kg]	260
6-Month Hardware Consumables Mass [kg]	260
Daily Consumables Mass [kg]	1.4
Peak Electrical Power Input [kilowatts]	16
Peak Thermal Power Input [kilowatts]	39
Electrical Energy [kw-hr/kg total products]	3
Thermal Energy [kw-hr/kg total products]	6

delivered hardware, reagent, and consumables mass). MAPS process leverage is about 35 (based on the ratio of combined product mass to consumables mass for seals and wear components).

This chapter focuses on production of iron, metal oxides, and oxygen on Mars. However, a subset of the MAPS flow sheet applies to lunar soils if reagents are brought from Earth and are fully recovered and recycled. The types of products and their quality were found to be very similar to those obtained from Mars soil simulant. A later section of this chapter summarizes lunar applications of the Mars Aqueous Processing System.

21.2 Process Description

The Mars Aqueous Processing System employs partial dissolution of several elements from soil using sulfuric acid. After solid-liquid separation, the acid solution is chilled to crystallize dissolved constituents. The crystals are recovered and then dissolved in weak acid. The dissolved crystal solution is then subjected to sequential pH adjustment using magnesium hydroxide to precipitate individual concentrates. Concentrates are recovered from the solution and are thermally treated to prepare high-grade oxides while recovering sulfur to regenerate sulfuric acid.

The sulfuric acid and magnesium hydroxide required for initial MAPS operations can be prepared from sulfates and water available on Mars. For example, hydrated sulfates including kieserite, $MgSO_4 \cdot H_2O$ were detected by the OMEGA spectrometer on board the Mars Express orbiter (Bibring, 2006). Mars Odyssey observations of apparent water even in equatorial regions can be explained by the stability of various sulfate salt hydrates at Mars pressure and temperature conditions (Bish and Scanlan 2006).

Despite the apparent availability of reagent precursors on Mars, all start-up reagents plus a six-month make up supply are supplied from Earth for the conservative design outlined in this chapter. The small amounts of make up needed for ongoing operations are derived from constituents including magnesium and sulfur contained in many Mars soils.

For advanced implementation, all MAPS reagents can be prepared on Mars. A method for extracting magnesium sulfate to prepare sulfuric acid and magnesium

hydroxide is discussed in a later section. Potential refining methods to separate sulfates from chlorides and other salts are discussed in the Process Considerations section.

21.2.1 Soil Composition

Process development was carried out using the NASA JSC Mars-1 simulant. Results were then applied to a Mars soil composition representing the average of Viking and Pathfinder analyses. Table 21.3 summarizes the simulant and soil analyses (Allen et al. 1997).

Table 21.3 Mars Simulant and Soil Analyses

Constituent [Weight %]	Normalized NASA JSC Mars-1 Simulant	Average, Normalized Viking/Pathfinder Soil
Fe ₂ O ₃	15.62	19.46
MnO	0.28	---
MgO	3.40	7.00
CaO	6.19	6.34
Na ₂ O	2.40	2.32
K ₂ O	0.61	0.22
Al ₂ O ₃	23.31	8.03
SiO ₂	43.50	47.91
TiO ₂	3.79	0.86
SO ₃	---	7.22
Cl	---	0.63
P ₂ O ₅	0.89	---

Where no values are shown, analyses were not performed. Separate analyses by x-ray fluorescence showed less than 0.15 percent of S as SO₃ and less than 0.02 percent Cl in the Mars-1 soil simulant.

The Mars soil simulant was calcined at 600°C to remove moisture and organic matter prior to conducting laboratory experiments. The size fraction of soil less than 0.85 millimeter was used for all process development experiments.

21.2.2 Unit Operations

Each unit operation described below was validated in the laboratory using NASA JSC Mars-1 soil simulant. Solutions were carried forward through each unit operation and then recycled to represent near-steady-state conditions for the closed-circuit process. Modeling was carried out in parallel with the bench-scale experiments to generate material and energy balances and to guide process development. Iron, manganese, magnesium, calcium, sodium, potassium, aluminum, and silicon were tracked individually in the process model. Many other elements in the feed

and products were analyzed to determine their relative extraction and distribution in the process. The closed-loop MAPS flow sheet is shown in Figure 21.2.

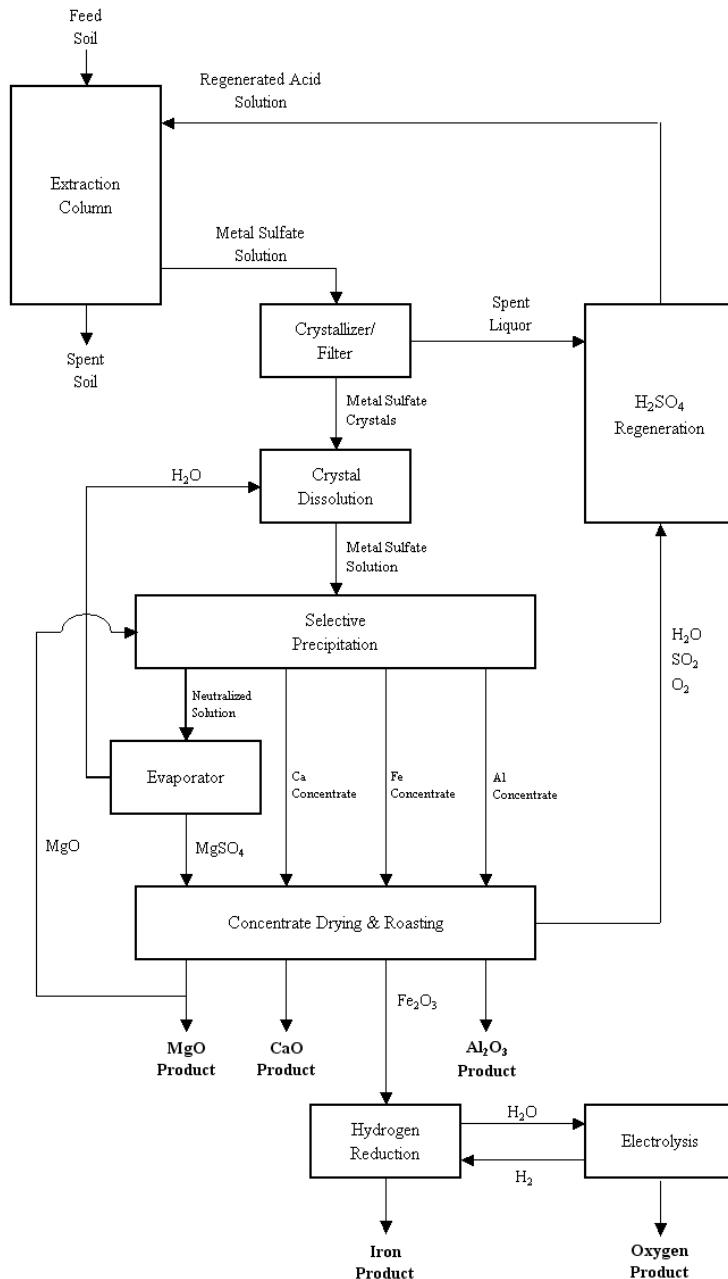


Fig 21.2 MAPS process flow sheet

The MAPS operating sequence is summarized in Table 21.4.

Table 21.4 MAPS Operating Sequence

Process Step	Time ^a [minutes]
1 st Stage Soil Extraction	140
1 st Stage Crystallization	(50)
2 nd Stage Soil Extraction	140
2 nd Stage Crystallization	(50)
Crystal Dissolution and CaSO ₄ Recovery	40
Iron Recovery	70
Aluminum Recovery	70
Crystal/Precipitate Drying	30
MgSO ₄ Solution Evaporation	60
Soil Residue Washing/Drying	50
Crystal Roasting and Acid Regeneration	120
Product Recovery and Mg(OH) ₂ Preparation	30
Hydrogen Reduction	(240)
Electrolysis, Oxygen Liquefaction	(750)
Total Cycle Time	750

^a Times in parentheses indicate activities conducted simultaneous with other process steps

21.2.3 Magnesium Sulfate Extraction and Recovery

The process design presented in this chapter is based on delivering start up reagents from Earth. However, a method that can be implemented on Mars for preparing sulfuric acid and magnesium hydroxide required for process start up was developed and demonstrated. This method significantly reduces the mass of supplies delivered from Earth and is suitable for advanced process implementation. Magnesium sulfate is first extracted from selected rich deposits on Mars using hot water. Nearly pure magnesium sulfate is then crystallized from solution and thermally decomposed to generate sulfuric acid precursor gases and magnesium oxide, which is hydrated to produce magnesium hydroxide base. With operation of subsequent MAPS steps in a closed-circuit mode, only small additional amounts of magnesium sulfate are needed to make up for process losses.

The extraction and recovery process is based on magnesium sulfate solubility in water, which increases from about 18 weight percent at 0°C to about 36 percent at 60°C (Dean 1992). Therefore, nearly saturated solutions produced by water extraction at elevated temperature can be cooled to crystallize about one-half of the dissolved magnesium sulfate. The recovered crystals are dried and thermally decomposed to produce MgO and acid precursor gases. The spent solution is reheated and used for further hot water extractions.

This hot water extraction and crystallization technique was verified using a Mars soil simulant infused with magnesium sulfate. Upon cooling a filtered solution

obtained after water extraction at about 60°C, magnesium sulfate crystals were recovered in the form of $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$. The hydrated magnesium sulfate released nearly pure moisture up to about 300°C. After heating to about 1,200°C, nearly complete decomposition took place to generate SO_2 and O_2 while forming nearly pure MgO .

21.2.4 Acid Production and Regeneration

A critical step in the Mars Aqueous Processing System involves the production of sulfuric acid needed for soil extraction. The overall acid formation reaction is shown below:

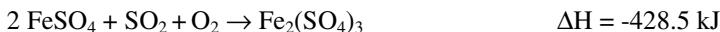


Oxidation of SO_2 with O_2 to form SO_3 prior to reaction with water is key for production of sulfuric acid. SO_3 formation proceeds very slowly in the absence of a catalyst. Without catalyst, solutions containing primarily dissolved SO_2 in water will be formed.

In commercial practice where high concentrations of nearly pure sulfuric acid are needed for industrial purposes or to reduce transportation costs, oxidation of SO_2 to SO_3 is conducted over expensive and easily poisoned catalysts at high temperatures. This technology is not well suited to MAPS.

Instead, an acid production process referred to as “autoxidation” is performed at low temperatures (10 to 75°C) by catalyzing the oxidation of SO_2 with O_2 in the presence of dissolved metal sulfates to generate sulfuric acid (Rampacek and Clemmer 1963).

Effective autoxidation catalysts are iron and manganese sulfates. Manganese becomes active at lower concentrations (about 1 gram per liter). However, it is prone to poisoning by the presence of small amounts of impurities such as copper. On the other hand, iron sulfate (above about 10 grams per liter) is more robust in terms of its tolerance to other dissolved species. The autoxidation process is based on the following chemistry:



The first reaction occurs quickly if the gases are provided as small bubbles with a low $\text{SO}_2:\text{O}_2$ ratio. The second reaction is favored by a high $\text{SO}_2:\text{O}_2$ ratio or larger gas bubbles. The autoxidation process is particularly suitable to MAPS since the main catalyst (iron sulfate) is present in suitable concentrations in acid regeneration feed solutions.

A laboratory gas dispersion test rig was assembled and tested to determine per-pass conversion of SO_2 , excess oxygen requirements, temperature effects, and sensitivity to catalyst concentration. Because the process is conducted at moderate temperatures (<50°C), conventional plastic equipment was used.

The literature discussed the need for an excess of oxygen to aid SO_2 oxidation to SO_3 , which is attained in MAPS by recirculating oxygen through the acid regeneration system. Test results confirmed that fine gas dispersion is important.

The use of manganese plus iron was reported in the literature to be completely resistant to poisoning while providing a higher rate of acid production. Fortunately, both iron and manganese are extracted with sulfuric acid from lunar and Mars soil simulants. The engineering design assumes that the iron and manganese sulfates required at start up are dissolved in the sulfuric acid delivered from Earth. Laboratory investigations showed that sufficient quantities of iron and manganese sulfates remain in solution during the closed-loop processing described below to catalyze sulfuric acid production.

Virtually complete oxidation of SO_2 to SO_3 followed by conversion to up to 16 weight percent H_2SO_4 was achieved in the laboratory at temperatures of less than 50°C at about 1 atmosphere absolute pressure. The literature citation indicates that the autoxidation process has been used to produce concentrations of up to 40 percent H_2SO_4 . MAPS modeling assumed further optimization and direct production of acid at the required 28 percent concentration. Alternatively, a low-temperature concentration step can be used to boost the strength to the 28 percent range if necessary. An examination of the partial pressure of water over solutions of sulfuric acid (Gmitro and Vermeulen 1964) showed that distillation enhanced by inert gas sparging at 50 to 70°C and about 0.5 atmosphere absolute pressure can remove nearly pure water to concentrate H_2SO_4 in solution. The distillation method was confirmed during laboratory experiments.

21.2.5 Acid Extraction of Soil

Soil is subjected to an acid extraction step to partially dissolve several elements. Conditions were selected to target high iron extraction.

For the purposes of process evaluation and modeling of the extraction step, constituents were expressed as their respective oxides, rather than as silicates, aluminates, or other likely mineral forms. Thermodynamic data were mostly derived from HSC Chemistry v5.11® software.

Shown below are the reactions of ferric oxide and other key constituents with sulfuric acid along with their reaction enthalpies:

$\text{Fe}_2\text{O}_3 + 3\text{H}_2\text{SO}_4 = \text{Fe}_2(\text{SO}_4)_3 + 3\text{H}_2\text{O}$	$\Delta H = -175.5 \text{ kJ}$
$\text{MnO} + \text{H}_2\text{SO}_4 = \text{MnSO}_4 + \text{H}_2\text{O}$	$\Delta H = -151.9 \text{ kJ}$
$\text{MgO} + \text{H}_2\text{SO}_4 = \text{MgSO}_4 + \text{H}_2\text{O}$	$\Delta H = -159.0 \text{ kJ}$
$\text{CaO} + \text{H}_2\text{SO}_4 = \text{CaSO}_4 + \text{H}_2\text{O}$	$\Delta H = -274.5 \text{ kJ}$
$\text{Na}_2\text{O} + \text{H}_2\text{SO}_4 = \text{Na}_2\text{SO}_4 + \text{H}_2\text{O}$	$\Delta H = -441.8 \text{ kJ}$
$\text{K}_2\text{O} + \text{H}_2\text{SO}_4 = \text{K}_2\text{SO}_4 + \text{H}_2\text{O}$	$\Delta H = -547.9 \text{ kJ}$
$\text{Al}_2\text{O}_3 + 3\text{H}_2\text{SO}_4 = \text{Al}_2(\text{SO}_4)_3 + 3\text{H}_2\text{O}$	$\Delta H = -180.6 \text{ kJ}$

After laboratory evaluation of operating parameters, a two-stage extraction at 70°C and about 1 atm pressure was selected. Each stage is carried out for two hours at a liquid:solid mass ratio of 2:1. A solid-liquid separation is conducted between stages. Sulfuric acid concentrations of 28 and 25 weight percent are used for the first and second stages, respectively. Nearly three-fourths of the iron is extracted from the soil after two stages of extraction. About 95 percent of the iron is

extracted from the finest soil size fraction, resulting in a silica-rich byproduct suitable for manufacture of glass and ceramics. Table 21.5 summarizes the overall laboratory Mars soil extraction results.

Table 21.5 Extraction of Constituents from Mars Soil Simulant

Extraction Stage	Soil Wt Loss [%]	Fe ₂ O ₃ Extn [%]	MnO Extn [%]	MgO Extn [%]	CaO Extn [%]	Na ₂ O Extn [%]	K ₂ O Extn [%]	Al ₂ O ₃ Extn [%]	TiO ₂ Extn [%]	P ₂ O ₅ Extn [%]
First	24.3	53.0	62.8	7.3	10.6	0.0	23.5	51.2	43.8	71.8
Second	32.9	73.8	79.1	50.8	14.1	12.6	36.5	61.9	67.3	85.0

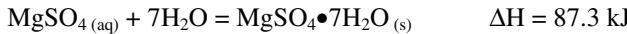
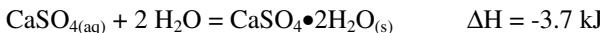
Silica extraction is very low under the conditions targeted for MAPS (about 0.1 percent). However, silica was tracked because it represents a significant portion of the soil mass. In addition, high-silica extraction residues are of use for materials production. For interpretation of laboratory results, silica extractions were set to zero. This allowed calculation of sample weight changes from analytical data, which reduced errors related to laboratory experiment sample losses and residue weighing uncertainties.

A water wash is applied after the final solid-liquid separation to displace and recover dissolved constituents from the fluid between particles. The wash liquor is advanced with the primary filtrate to the crystallization step. The extraction residue solids contain about 20 percent moisture and can be dried for use as building materials or glass. Moisture is recovered and recycled.

21.2.6 Crystallization

The extraction liquor contains excess acid used to provide a driving force for dissolution of the soil constituents. Therefore, a concentration of about 11 percent acid remains in solution after extraction. The crystallization step prior to selective precipitation avoids the need to neutralize this excess acid and eliminates a significant amount of acid regeneration that would otherwise be required.

The hot, filtered, acidic extraction solutions are cooled (from 70°C to 0°C) to force saturation for recovery of iron, aluminum, calcium, and magnesium as mixed, hydrated sulfate crystals. The reactions shown below were used for crystallization modeling of key constituents:



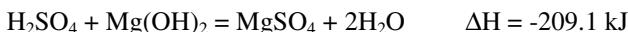
The degree of hydration can vary considerably depending on the temperature and other conditions during crystallization. For modeling purposes, only the forms shown above were considered.

In most cases, about half of the dissolved elements of interest are recovered from solutions fed to crystallization. The bulk crystals are advanced to selective precipitation. The spent liquor is fed to acid regeneration and is then recycled to the next extraction step. The dissolved elements remaining in the spent liquor after crystallization are of sufficiently low concentration to allow dissolution of soil constituents at elevated temperature. In addition, the iron and manganese sulfates contained in the spent liquor are used to catalyze acid production.

21.2.7 Selective Precipitation

The objective of selective precipitation is to isolate and separately recover iron and other constituents from solution by addition of MgO to gradually raise pH by neutralization of residual acid. The technique requires iron to be fully oxidized to ferric (+3) form to achieve sharp separation from aluminum and other compounds.

Bulk crystals are dissolved in water to form an acidic sulfate salt solution at about pH 1 (resulting from residual acidity in the bulk crystals). Most of the calcium sulfate present in bulk crystals remains undissolved, allowing recovery of nearly pure calcium sulfate by filtration immediately after the crystal dissolution step. The remaining solution contains the primary constituents iron, aluminum, and magnesium along with other minor dissolved elements. The solution is sparged with oxygen to ensure complete conversion of any ferrous sulfate to ferric sulfate. Next, MgO base is added as a finely dispersed slurry in water to increase solution pH. Neutralization proceeds according to the following reaction:



The iron and aluminum concentrates are recovered sequentially as pH increases according to the following precipitation reactions:



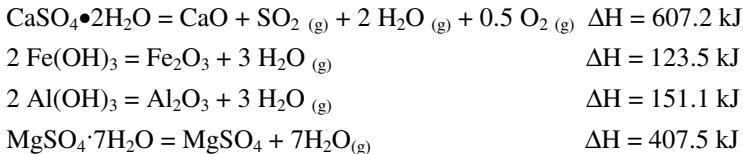
Ferric hydroxide precipitates from solution first at about pH 3 and is recovered by filtration. After more MgO base is added, aluminum begins to precipitate at about pH 4. After filtration, further neutralization can be employed to precipitate manganese and other elements above pH 7. Manganese recovered by this method produces a low-grade concentrate that requires further refining. After final filtration, the neutralized magnesium sulfate solution is subjected to partial evaporation and is then cooled to force saturation and removal of hydrated magnesium sulfate crystals.

The precipitates and sulfate crystals are advanced to drying and roasting as described below to prepare oxide concentrates and to recover water and gases for acid regeneration.

21.2.8 Drying and Roasting

The individual precipitates and sulfate crystal products are dried and then roasted at elevated temperatures (900°C for iron and aluminum precipitates, 1,200°C for

magnesium sulfate, and about 1,500°C for calcium sulfate) to generate oxides while releasing various amounts of H₂O, SO₂, O₂, and residual H₂SO₄. Nearly pure water is recovered at lower temperatures and is recycled to the process for washing after solid-liquid separations, preparation of MgO neutralization slurry, and make up in recycled acid extraction solutions. A mixture of water, H₂SO₄, SO₂, and O₂ is released from concentrates at higher temperatures. The following reactions represent the roasting step for the key MAPS constituents:



The relatively high product purities, which aid resistance to sintering or fusing during roasting, lead to generally free-flowing powders. Released gases are used to regenerate sulfuric acid solution. The iron oxide concentrate is sent to hydrogen reduction for manufacture of metallic iron and oxygen byproduct. The other oxides are used as refractory materials or are refined further. The compositions of calcium, iron, aluminum, and magnesium concentrates prepared from Mars soil simulant were determined by x-ray fluorescence and are shown in Table 21.6 on a normalized, sulfur-free basis following roasting. Extraction and refining results were applied to Mars soil compositions for modeling and engineering design.

Table 21.6 Analysis of Products from Mars Soil Simulant Extraction

Constituent	Calcium Concentrate [Wt %]	Iron Concentrate [Wt %]	Aluminum Concentrate [Wt %]	Magnesium Concentrate [Wt %]
Fe ₂ O ₃	0.09	83.7	1.25	<0.04
MnO	<0.03	0.01	0.27	0.68
MgO	<0.14	0.06	2.31	95.3
CaO	98.9	0.01	<0.01	2.05
Na ₂ O	<0.14	<0.05	<0.82	0.65
K ₂ O	<0.03	0.01	0.03	0.80
Al ₂ O ₃	0.09	1.17	94.6	<0.11
SiO ₂	0.37	0.11	<0.03	<0.11
TiO ₂	<0.03	10.79	0.42	<0.04
P ₂ O ₅	<0.14	4.08	0.18	<0.19

Concentrate purity ranged from as high as 98.9 percent for CaO to as low as 84 percent for Fe₂O₃. The primary impurity in the iron oxide concentrate was titanium, which exists in the Mars simulant in an amount of over four times that of the average actual Mars soil. If desired, the iron oxide impurities can be removed using fluxing agents and slagging procedures similar to those employed in blast furnaces. Soil extraction residue, oxygen, or other metal oxides can be used to aid impurity separations.

21.2.9 Iron Oxide Reduction

Iron oxide is reduced to generate metallic iron and water vapor. The high surface area of the fine particles formed by selective precipitation are particularly amenable to hydrogen reduction. The MAPS design specifies that iron oxide reduction be performed in the same furnace used for roasting. A reduction temperature of 750°C is used according to the following reaction:



Note that the enthalpy for the reduction reaction is shown at standard state. At 750°C, the reaction is actually endothermic and requires heat input. The exhaust gas is cooled to condense water. The water is electrolyzed for production of oxygen and for recycle of hydrogen to the reduction process.

21.2.10 Metal Forming

After reduction, the iron powder can be heated further to either produce sintered products or molten metal. Examples are discussed below.

21.3 Products

The following paragraphs describe some of the materials produced from MAPS concentrates.

21.3.1 Metallic Iron

An integrated reduction-metal forming process was demonstrated during MAPS development. For this procedure, an iron oxide concentrate is loaded into a ceramic crucible, reduced by hydrogen at about 750°C, and then heated to form a sintered product (at 1,500 to 1,550°C). Figure 21.3 shows the progression of treatment of a MAPS iron oxide concentrate produced from soil simulant. The iron oxide concentrate is shown on the left. A sintered metallic iron button produced by reduction and subsequent thermal treatment at 1,550°C is shown in the center. An example product machined from the sintered button is shown on the right.



Fig 21.3 Iron oxide concentrate, metallic button, and machined product.

The MAPS iron oxide concentrate is also a suitable feed for molten treatment similar to blast furnace operations, in which impurities are transported to a slag phase. Purifying additives including oxygen and fluxing agents can be derived from other MAPS products or residues.

21.3.2 Calcium, Aluminum, Magnesium and Iron Oxides

Preparation of refractory powders was demonstrated by thermal treatment of the calcium, aluminum, and magnesium concentrates generated by crystallization and selective precipitation. These products exhibit very high sintering and melting temperatures and are therefore useful for manufacture of insulation, furnace liners, casting molds, and durable ceramic components.

The pure oxides have melting temperatures between about 2,050°C (alumina) and 2,800°C (magnesia). The MAPS oxide products exhibited only slight agglomeration and were easily broken into powder form after thermal treatment at 1550°C. The calcia, alumina, and magnesia products can be combined with high-silica soil residue to form a number of ceramics including mullite, which exhibits good strength and resistance to thermal stresses. Other silica-free ceramics of potential use on Mars include magnesia-alumina spinel, calcia-alumina formulations, and iron oxide-alumina compositions. The oxide products can be used as additives to bulk soil or extraction residues to alter melting temperature or to improve strength or thermal expansion properties. The aluminum and magnesium oxide concentrates would also be excellent feeds for light metal production.

21.3.3 Concrete and Salt-bonded Products

Production of durable pellets and bricks was demonstrated using Mars soil simulant and extraction residues bound with MAPS-derived anhydrous magnesium sulfate and magnesium oxide binders. Dry soil or extraction residue is first mixed with binder powders. Water is then blended with the soil and binder mixture to form a thick concrete-like material. The freshly mixed materials are transferred to molds and allowed to cure for at least 24 hours before removal.

Brick preparation methods attempted without binder failed, indicating no natural cement-like properties of the soil simulant or extraction residue. The magnesium sulfate binder works by hydrating to a plaster-of-Paris type of material that attaches to the soil or extraction residue particles and exhibits considerable strength as long as further moisture additions are avoided. The magnesia binder works as a substitute for calcium aluminosilicates such as cement. Figure 21.4 shows examples of nine-centimeter-long bricks prepared from extraction residues. The brick on the left was bound with $MgSO_4$; the brick on the right was bound with MgO .

Soil or residue was mixed with binder at about a 1:1 ratio for most experiments, which were not optimized with respect to binder content, moisture addition, particle size distribution, cure time, or cure temperature. Water was added at a ratio of about 60 to 70 percent of total dry solids mass. Compressive strengths of up to 57 kg/cm^2 were obtained. This is about one-third of terrestrial concrete compressive strength

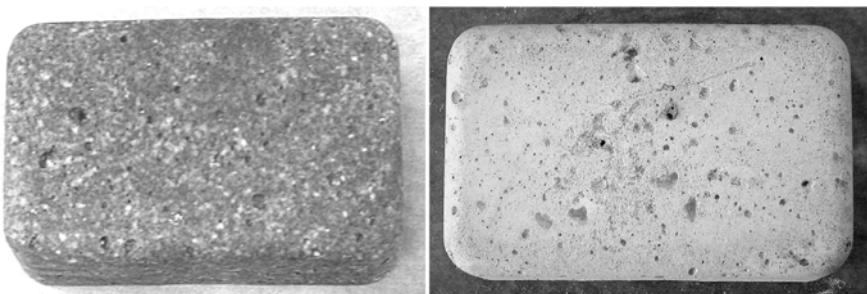


Fig 21.4 Bricks prepared from extraction residues.

requirements (NRMCA 2003) and is potentially suitable for use in the reduced gravity of Mars. In general, structural materials prepared using extraction residues were superior to those prepared from raw soil simulant. Materials prepared using magnesium sulfate binder were generally stronger than those prepared using magnesium oxide binder. However, extraction residue prepared using the magnesium oxide binder exhibited better long-term stability. Bricks bound with magnesium oxide did not exhibit the small cracks that developed over a period of several weeks in some bricks bound with magnesium sulfate.

21.3.4 Sintered Products

The Mars soils and extraction residues are candidate feeds for manufacture of sintered components. Non-optimized procedures showed that the sintering temperature can be altered by the retention or further addition of metal sulfate salts such as would exist in the bulk extraction solutions. Alternatively, the MAPS aluminum, calcium, or magnesium oxide products can be used as enhancers to alter the sintering temperatures and mechanical properties. Combined with optimized particle size distribution, heating techniques, and cooling techniques, MAPS-based modifiers are capable of generating a wider range of product characteristics than could be achieved using unaltered soils.

21.3.5 Alkali Metal Sulfate Salts

The MAPS engineering design conservatively assumed that any extracted alkali metal oxides would exit the system as sulfates contained in excess MgO product, representing a process sulfur loss. This sulfur loss is made up from magnesium sulfate extracted from feed soil. Much of the water-soluble alkali metal sulfates can be removed by water washing of the roasted magnesium oxide product. Alternatively, a portion of the alkali metal sulfate salts that are extracted into the sulfuric acid solution and recovered with the magnesium sulfate crystals can be recovered during high-temperature roasting of magnesium sulfate. The vapor pressure of alkali metal sulfate salts allows for their removal in the gas phase. Recovery of condensed Na_2SO_4 and K_2SO_4 in relatively pure form could be useful for preparation of H_2SO_4 and strong bases such as NaOH/KOH upon installation of an additional process module.

21.3.6 Silica

Silica has application for manufacture of photovoltaics, semiconductors, glass, and ceramics. The finest fraction of the Mars soil simulant extraction residue contains nearly 80 percent silica. This material could serve as feed to additional sulfuric acid extraction with a focus on removing the remaining impurities (primarily calcia, alumina, and iron oxide). Alternatively, the high-silica residue could be used as feed to other silica extraction and purification methods such as strong base, HF, pyrolysis, or fluorine gas processes. The high-quality silica feed would minimize equipment size and refining complexity for these processes.

At about 80 percent purity, MAPS silica could make a good ceramics ingredient for production of a wide range of useful materials similar to porcelain ($\text{Al}_3\text{O}_3 \cdot (\text{K},\text{Na})_2\text{O} \cdot 2\text{SiO}_2$), mullite ($3\text{Al}_2\text{O}_3 \cdot 2\text{SiO}_2$), and cordierite ($2\text{MgO} \cdot 2\text{Al}_2\text{O}_3 \cdot 5\text{SiO}_2$). Refractory reinforcement and insulating fibers with maximum use temperatures in the range of 1,300°C can be made from silicates that also include alumina, calcia, and magnesia (Waugh 1993).

21.3.7 Glass

Glass has applicability for transparent or translucent covers. Commercial glass (soda-lime-silica glass) contains about 70 percent SiO_2 and is made by adding Na_2CO_3 (soda) and CaCO_3 to sand. The silica required for conventional glass formulations can be obtained using the high-silica fine fraction of soil extraction residue.

Glass preparation was demonstrated with the high-silica soil extraction residue using only its residual potassium, sodium, calcium and aluminum. This material contains less potassium and sodium than would be used for manufacture of conventional glass. A Mars simulant extraction residue (about 79 percent SiO_2) was heated to 1,550°C under a low flow of helium to determine its melting properties and clarity after fusion.

The high-silica material formed a green-color glass. Because no effort was made to anneal the product, it broke into small pieces upon cooling. However, the product exhibited good light transmission properties and appears to be suitable for manufacture of colored glass by application of proper curing times and temperatures.

The same high-silica Mars simulant residue was used in another glass formulation with the addition of potassium carbonate. A mixture of high silica residue and K_2CO_3 was prepared and placed in a stainless steel mold. The mixture was heated to 1,100°C and was then annealed overnight at 485°C. Although the formulation was not optimized, the resulting glass was translucent and structurally stable.

Yet another glass manufacturing approach was identified in citations related to alumina-based glass. Alumina-based glass can be used to produce strong and optically desirable products (Weber et al. 2004). In some cases, modifiers such as calcium and/or rare-earth-element compounds are required to reduce melting temperature and to improve thermal properties (Rosenflanz et al. 2004). Alumina glass composites can also be fabricated in molds where complex shapes with good dimensional control are possible (Zhu et al. 2005). The basic ingredients for this glass making approach are available in the form of the MAPS alumina and calcia (or magnesia) products.

21.3.8 Manganese and Minor Elements

Lower-grade products including manganese are precipitated from solution at higher pH values. With further refining to separate magnesium contamination, a small amount of ferro-manganese or manganese concentrate could be prepared. Manganese increases iron strength in a manner similar to carbon addition (McGannon 1971). Selective precipitation during initial MAPS development focused on constituents capable of being collected at high grades (80 percent purity or greater) in a single stage. However, methods including chelating ion exchange, solvent extraction, sulfide precipitation, and electrochemistry can be applied in various combinations to target the recovery of manganese and other minor Mars soil constituents.

21.4 Hardware Description

MAPS hardware was designed to accommodate multiple process steps in order to minimize delivered equipment mass. Table 21.7 contains the equipment list and estimated component mass. The design basis and features of key process equipment are described in the following sections.

Table 21.7 Hardware Summary.

Component	Mass [kg]
Extraction Column	100
Crystallizer	81
Precipitator	108
Evaporator	108
Acid Regenerator	115
Mg(OH) ₂ Slurry Tank	54
Condenser	62
Vacuum Drying Oven	39
Roasting-Reduction Furnace	165
Electrolyzer	69
Oxygen Liquefaction Module	187
Oxygen Storage System	110
Pneumatic Gas System	15
Instrumentation/Controls/Ancillary	20
Equipment Frames/Supports	62
Total Hardware Mass	1295

21.4.1 Extraction Column

A scale model column extraction unit was designed and demonstrated during MAPS process development. The system involves only one soil feeding and one residue discharge step per batch cycle. The soil remains in the extraction column

for two stages of extraction, solid-liquid separations, residue washing, and drying prior to discharge.

The extraction column is integrated with a separate reservoir connected by tubing on the bottom for fluid transfer and on the top for gas transfer. An overhead pump, which is isolated from direct contact with hot and corrosive process liquors, is used to pneumatically transfer liquid from the reservoir to the column, to provide agitation by gas bubbling, and to pull extraction liquor through the extraction column solids support back to the reservoir. An integrated polypropylene filter felt is installed at the bottom of the column to facilitate solid-liquid separation. The laboratory column was constructed mostly from PVC (polyvinyl chloride). Polyvinylidene fluoride is specified for the full-scale unit.

The scale model extraction column, reservoir, and pneumatic pump system was developed and tested at a scale of 200 to 300 grams of feed soil per batch and scaled to the 100 kg soil per batch used for design engineering. Figure 21.5 shows a schematic of the extraction column and a photograph of the laboratory scale model.

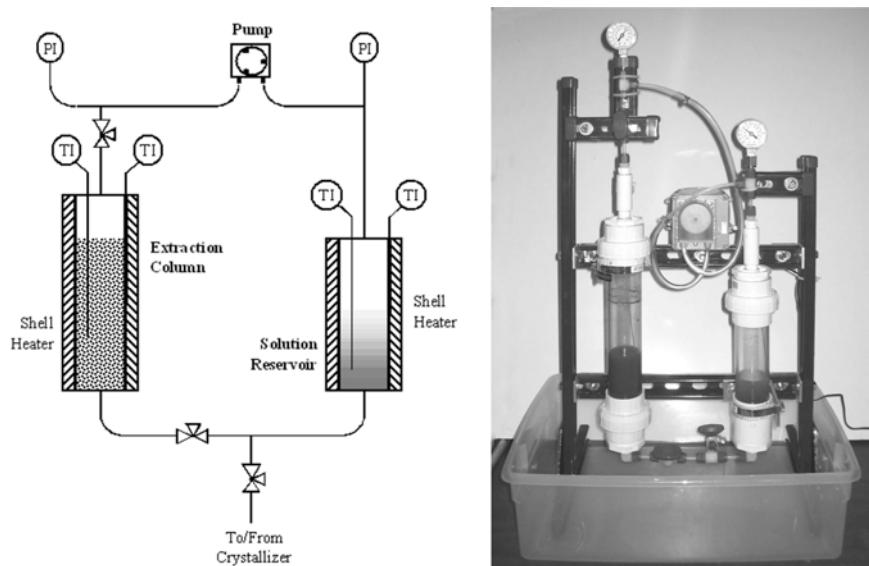


Fig 21.5 Extraction column schematic and photograph of scale model.

Solution, preheated as necessary in the adjoining reservoir, is transferred upward into the bed of soil for extraction. The pump is periodically operated to provide agitation via bubbling of inert gas through the slurry. The initial supply of inert gas is provided with reagents transported from Earth. Gas make up requirements would be met by recovering nitrogen or argon components of the Mars atmosphere, likely in parallel with other *in-situ* resource utilization for fuels or oxygen production.

Near the conclusion of an extraction cycle, the pump is used to further agitate the bed by an upward flow of gas. The flow is stopped, and particles are allowed to settle. This serves to sort particles, with the coarsest material settling first to act as a sand filter to assist solid-liquid separation.

The 51-centimeter-diameter extraction column was specified with a total volume of 315 liters and a height:diameter ratio of 3:1. PVDF was specified for most wetted parts. The mass of fittings, valves, heaters, and the pump were estimated. An additional 10 percent contingency was added to this and all other hardware mass estimates.

21.4.2 Crystallizer

The extraction solution is transferred to a column system similar to that described above. This system crystallizes and filters the bulk components from the extraction solution. The column is chilled for crystallization to about 0°C. An integrated filter felt, similar to that used for the extraction column, is installed in the lower flange of the crystallizer for solid-liquid separation of hydrated crystals from the spent solution. Gas is bubbled through the solution during crystallization. After solid-liquid separation, the spent crystallization liquor is transferred to the acid regeneration column. The crystals are transferred from the crystallizer column to the selective precipitation column. Although crystallization and selective precipitation could be carried out in the same system, separate systems are used to allow simultaneous operations in order to reduce batch cycle time.

The 45-centimeter-diameter crystallizer was specified with a total volume of 210 liters and a height:diameter ratio of 3:1.

21.4.3 Selective Precipitation Column

Bulk crystals are dissolved in weak acid solution prior to selective precipitation of calcium, iron, and aluminum concentrates. Recycled water from the condenser is used to dissolve the bulk crystals in the small amount of residual acid carried forward with the crystals. Un-dissolved calcium sulfate is recovered first using filtration methods similar to those described above. Iron and aluminum precipitates are filtered and recovered separately following pH adjustments. Solution remaining after selective precipitation is pumped to the evaporator described below for recovery of magnesium sulfate. Like the columns discussed previously, pneumatic agitation is used.

The 53-centimeter-diameter precipitator was specified with a total volume of 360 liters and a height:diameter ratio of 3:1.

21.4.4 Evaporator

The Evaporator is used as a transfer and holding tank during precipitation operations in addition to its primary function of producing magnesium sulfate crystals by evaporation of spent liquor. The design and geometry is similar to that of the Selective Precipitation Column. The Evaporator is designed to operate with vacuum up to the PVDF maximum operating temperature of 138°C.

21.4.5 Acid Regeneration Column

Acid regeneration utilizes a taller column that provides for contact and reaction of spent crystallization liquor with sulfur dioxide, sulfuric acid fumes generated by crystal roasting, and recirculating oxygen. Operating experience shows that the majority of recirculating gas consists of oxygen because sulfuric acid and sulfur dioxide readily absorb into the solution.

The 42-centimeter-diameter extraction column was specified with a total volume of 359 liters and a height:diameter ratio of 6:1. PVDF was specified for most wetted parts.

21.4.6 Mg(OH)₂ Slurry Preparation Tank

A tank is provided for preparation and holding of the magnesium hydroxide slurry used during selective precipitation and acid neutralization. Design of the tank is similar to those described above. The required portion of magnesium oxide, generated by roasting the magnesium sulfate resulting from acid neutralization, is slurried with water recycled from the condenser to produce the magnesia slurry. As with the other operations, pneumatic agitation and slurry transfer methods are assumed for this system.

The 37-centimeter-diameter extraction column was specified with a total volume of 121 liters and a height:diameter ratio of 3:1.

21.4.7 Water Condenser

A condenser including a tank similar in construction to process extraction columns is used to recover water from the evaporation of precipitates or crystals and drying of washed soil residue. Chilling power is required to remove the heat of condensation and to hold the contents at about 0°C.

The 52-centimeter-diameter condenser was specified with a total volume of 338 liters and a height:diameter ratio of 3:1.

21.4.8 Vacuum Drying Oven

A vacuum drying oven is used for removal of free and bound moisture from precipitates and sulfate crystals. Operating temperatures of up to 300°C and pressures as low as 0.1 atm are specified. The conditions are intended to remove nearly pure moisture by adjusting temperature and pressure during drying. Concentrates are dried separately in the same oven. The oven consists of a horizontal cylindrical chamber of 57 liters (38 centimeters inside diameter) with a single, sealable hatch for insertion of ceramic trays containing moist feeds. An insulated aluminum shell with an alloy liner is specified for the oven. Tubing ports to control pressure and to exhaust evaporated water are included in the oven design.

Upon completion of the drying step, trays with dried precipitates or crystals are conveyed out of the drying oven and into the roasting furnace.

21.4.9 Vacuum Roasting and Reduction Furnace

A sealed furnace with a ceramic liner is used for removal of residual H₂O and H₂SO₄ and for decomposition of metal hydroxides or sulfates to produce acid precursor gases (SO₂, O₂, and H₂O). Most sulfates are decomposed in the 1,200°C range. Up to 1,500°C may be needed for complete decomposition of calcium sulfate, although reduced pressures enable operation at lower temperatures. The calcium, aluminum, and magnesium oxide concentrates are removed from the furnace after roasting.

The roasted iron precipitate (iron oxide) remains in its ceramic tray and is reduced with hydrogen at a temperature of about 750°C. The resulting water is condensed, and hydrogen is recycled after electrolysis.

The furnace volume is 203 liters based on a cube shape with inside dimensions of 59 centimeters per side.

21.4.10 Electrolyzer

The electrolysis unit is used to split water into hydrogen for recycle to the iron oxide reduction operation and to generate byproduct oxygen. Some oxygen is recycled within the MAPS process for oxidation of any ferrous iron to ferric iron prior to selective precipitation and to maintain oxygen-rich conditions during acid production.

The electrolyzer specifications are adapted from previous *in-situ* resource utilization work (Rosenberg et al. 1992). The electrolyzer is assumed to operate continuously at the average gross oxygen production rate. The electrolyzer mass and power requirements were proportioned from information in the literature citation, and an additional allowance for mass was provided due to the smaller scale of MAPS operation.

21.4.11 Oxygen Liquefaction Unit

The oxygen liquefaction unit operates continuously in conjunction with the electrolyzer. The liquefaction unit condenses the net oxygen product from the electrolyzer and recovers any oxygen vented during storage. Factors for the mass and power using radiative cooling to minimize chilling requirements were taken from the same source cited above with additional allowance provided for the smaller scale of MAPS operations.

21.4.12 Oxygen Storage

The oxygen storage unit operates in conjunction with the electrolyzer and liquefaction units. Similar factors as those used for electrolysis and liquefaction were applied to oxygen storage.

21.4.13 Ancillary Hardware

An allowance was provided for the mass and power for ancillary hardware and operations including instrumentation, automation, controls, heat exchangers, and gas collection and handling.

21.4.14 Equipment Frames and Supports

An allowance of five percent of the equipment mass was provided for the mass of equipment frames and supports (to secure process modules and interconnections).

21.5 Engineering Design

The MAPS engineering design was based on equipment sizes and specifications established for each major unit operation assuming 100 kg soil batches fed every 12.5 hours (average 192 kg/day soil feed). Power was calculated from reaction enthalpies and from projected heating or chilling requirements assuming thermal recovery between each process step of about 80 percent. The hardware mass values summarized above were projected from equipment volume, wall thicknesses, and materials of construction. Mass of ancillary equipment was factored from the mass of primary equipment. The plant footprint is estimated to be about 9 square meters with a height of 3 meters. Overall engineering design data were extracted from detailed design calculations and are summarized in Tables 21.1 and 21.2.

Virtually all initial reagent needs and make up for reagent losses can be provided by indigenous resources. Therefore, only the hardware consumables for Mars materials production need be brought from Earth.

The process model included complete mass balances and make up requirements as a function of reagent losses to the various process streams. The main process losses include moisture and acid retained in washed and dried soil extraction residue, sulfur retained in roasted concentrates and alkali metal salts, and hydrogen retained in metal or lost during iron oxide reduction. Electrical power and energy requirements were calculated on the basis of the pressures and mass flows of liquids and gases in each unit operation. Thermal power and energy requirements were calculated on the basis of reaction enthalpies, process temperatures, and assumptions for heat exchange and heat transfer. The power source for the Mars Aqueous Processing System was not specified in the engineering design, but the system was developed with an emphasis on minimal energy consumption and low peak power demand.

Leverage was calculated by dividing daily production mass by daily hardware consumables mass. A process leverage of 35 was determined from the results of the detailed material balances and projected consumables requirements. A break-even operating time of about 50 days was calculated by dividing the delivered mass of initial hardware plus six-months of reagents and consumables by the daily production rate of materials plus oxygen. The overall results show MAPS to be attractive for iron, metal oxide, and oxygen production on both Mars and the Moon. The relatively high leverage and short breakeven operating time allow for expansion modules to produce a broader product suite, including advanced ceramics, glass, steel, and light metals.

21.6 Lunar vs Mars Application

Sulfuric acid, water, and hydrogen must be nearly completely recovered and recycled for MAPS applied to lunar processing. Process losses are partially made up

by recovery of volatile hydrogen from regolith and by sulfur recovery from lunar soil sulfides. Additional losses are made up from Earth supplies. Low alkali content reduces sulfate losses on the Moon compared to Mars. Engineering analysis showed operating leverage and breakeven times similar to those for Mars, assuming that iron extraction of the glassy lunar soil can be improved from 40 percent (established from lunar soil simulant testing) to the 75 percent range.

A key difference in Mars and Lunar application is the oxidation state of iron. On Mars, iron exists almost exclusively in ferric form. On the Moon, iron exists in ferrous and metallic states. Dissolved iron must be oxidized to ferric form prior to selective precipitation in order to maintain a sharp separation from aluminum and other products. Therefore, a small additional recirculating load on the electrolyzer is required to oxidize iron for precipitation prior to reduction of ferric oxide to its metallic state.

21.7 Process Considerations

Some issues related to hardware design and others related to the chemistry of the closed-loop MAPS are summarized below.

Solid-liquid separations steps were found to be superior in the column compared to conventional laboratory vacuum filtration equipment. This is in part due to the greater pressure differential that can be achieved in the column. Limited laboratory work suggested that temperature control and seeding can further boost filtration performance of precipitates (especially the ferric hydroxide precipitate).

The Martian surface is a rich source of salts for processing by MAPS. Sulfate salts, especially magnesium sulfate, are of interest for preparation of sulfuric acid and magnesium oxide base. However, other ions such as chlorides, perchlorates, and bromides exist with sulfate salts in some Mars soils. Although the presence of non-sulfate anions may not be deleterious to the extraction of constituents using MAPS, the presence of these other components may interfere with some of the crystallization and selective precipitation techniques employed by MAPS. Methods to isolate sulfates from other anions beyond selective mining are therefore of interest. Candidate methods include anion exchange processing, electrodialysis, nanofiltration membrane processing, and fractional crystallization. Of these methods, fractional crystallization is preferred due to reduced hardware and specialty resin requirements along with greater compatibility with other MAPS process steps. Such methods are employed for recovery of various products from Dead Sea brines (Aral et al. 2004, 2004a).

Mars soils can contain up to about 0.5 percent phosphorus (Greenwood and Blake 2006) while simulant contains about 0.35 percent. Roughly 85 percent of the feed phosphorus was extracted from the Mars soil simulants. Nearly all of the extracted phosphorus reports to the iron concentrate during selective precipitation. Phosphorus is a constituent in high-strength, low-alloy steels. Its presence provides corrosion resistance and markedly increases strength, but contributes to brittleness at concentrations greater than 0.1 percent (McGannon 1971). However, the embrittling effect is less pronounced in steels with carbon contents of less than 0.15 percent, which would be representative of the MAPS iron product. Although

the phosphorus concentration in the Mars product would be considered high for conventional applications, its effect on structural properties has not been determined. Phosphorus is conventionally removed from an iron melt via oxygen injection. Phosphorus has other potentially beneficial applications if it can be isolated. Some applications include desiccants, fertilizers, dopant for semiconductors, and organophosphorus compounds used for plasticizers, flame retardants, pesticides, extraction agents, and water treatment.

MAPS is a closed-loop process. All constituents extracted from the feed soils will build to a steady state concentration that is dependent on the rate of removal in various products or by process losses. It is possible that a constituent will eventually build to a concentration harmful to one or more of the unit operations. One means of dealing with this potential problem is to periodically evaporate the neutralized selective precipitation solution. Water would be recovered in nearly pure form. The resulting residue could be subjected to the standard roasting steps for recovery of the contained sulfur and oxygen for acid regeneration. Built up impurities could then be discharged from the system as impurities in the recovered oxides.

21.8 Conclusions

The Mars Aqueous Processing System was demonstrated to be effective for production of high-grade iron, metal oxides, and oxygen on Mars and the Moon. The engineering designs indicated favorable leverage and breakeven operating times. The process is amenable to addition of modules to broaden the product suite to include steel, advanced ceramics, glass, silicon, and light metals.

Acknowledgements

The work described in this chapter was conducted under funding provided to Pioneer Astronautics by the National Aeronautics and Space Administration (NASA) through the Small Business Innovation Research (SBIR) program. Mark Berggren was the Principal Investigator for Pioneer Astronautics. Kris Romig was the Contracting Officer's Technical Representative for NASA's Johnson Space Center. Pioneer Astronautics employees including James Kilgore, Anthony Muscatello, and Melanie West provided significant contributions to the development of process concepts, hardware design and fabrication, and process demonstrations.

References

- Allen, C.C., Morris, R.V., Jager, K.M., Golden, D.C., Lindstrom, D.J., Lindstrom, M.M., Lockwood, J.P.: Martian regolith simulant JSC Mars-1. In: 29th Lunar & Planetary Science Conference Proceedings, Lunar and Planetary Institute, Houston (1998) Abstract 1690
- Aral, H., Hill, B.D., Sparrow, G.J.: Value adding to salts recovered from saline waters in disposal basins in the Murray-Darling Basin, Proof of concept study. In: CSIRO Minerals Report DMR-2378C; Appendix 2, Production of Salts from Brines and Bitterns. Canberra City, Australia (2004)

- Aral, H., Hill, B.D., Sparrow, G.J.: Value adding to salts recovered from saline waters in disposal basins in the Murray-Darling Basin, Proof of concept study. In: CSIRO Minerals Report DMR-2378C; Appendix 3, Salts from saline waters and value added products from the salts, Canberra City, Australia (2004a)
- Bibring, J.P.: Mineralogic Diversity of Mars and Discovery of Sulfate Deposits from the OMEGA/Mars Express Investigation. In: Workshop on Martian Sulfates as Recorders of Atmospheric-Fluid-Rock Interactions, Lunar and Planetary Institute, October 22–24 (2006)
- Bish, D.L., Scanlan, M.K.: The Hydration and Dehydration of Hydrous Mixed-Cation Sulfates. In: 37th Lunar and Planetary Science Conference Proceedings, Lunar and Planetary Institute, League City, Texas (2006) Abstract 1011
- Dean, J.A.: Lange's Handbook of Chemistry, 14th edn. McGraw-Hill, New York (1992)
- Gmitro, J.I., Vermeulen, T.: Vapor-liquid equilibria for aqueous sulfuric acid. AIChE Journal 10(5), 740–746 (1964)
- Greenwood, J.P., Blake, R.E.: Evidence for an acidic ocean on Mars from phosphorus geochemistry of Martian soils and rocks. Geology 34(11), 953–956 (2006)
- McGannon, H.E.: The Making, Shaping, and Treating of Steel, 9th edn. United States Steel, Herick & Held, Pittsburgh (1971)
- NRMCA, Testing Compressive Strength of Concrete. In: National Ready Mixed Concrete Association bulletin CIP 35/3090/6.0/DCTR, Silver Spring, Maryland, National Ready Mixed Concrete Association (2003)
- Rampacek, C., Clemmer, J.B.: Sulfuric acid from sulfur dioxide by autoxidation in mechanical cells. In: Report of Investigations 6236, Bureau of Mines, U.S. Department of the Interior, Washington DC (1963)
- Rosenberg, S.D., Beegle Jr., R.L., Guter, G.A., Miller, F.E.: The Onsite Manufacture of Propellant Oxygen from Lunar Resources. In: McKay, M.F., McKay, D.S., Duke, M.B. (eds.) Space Resources - Materials, NASA SP-509, vol. 3, NASA Lyndon B. Johnson Space Center, Houston (1992)
- Rosenflanz, A., Frey, M., Endres, B., Anderson, T., Richards, E., Schardt, C.: Bulk glasses and ultrahard nanoceramics based on alumina and rare-earth oxides. Nature 430, 761 (2004)
- Waugh, R.A.: Refractory Fibers. In: Kirk-Othmer Encyclopedia of Chemical Technology, 4th edn., vol. 21, John Wiley & Sons, Inc., New York (1993)
- Weber, R., Tangeman, J.A., Nordine, P.C., Scheunemann, R.N., Hiera, K.J., Ray, C.S.: Rare-earth-aluminum oxide glasses for optical applications. Journal of Non-Crystalline Solids 345, 359–365 (2004)
- Zhu, Q., de With, G., Dortmans, L.J.M.G., Feenstra, F.: Near net-shape fabrication of alumina glass components. Journal of the European Ceramic Society 25(5), 633–638 (2005)

Chapter 22

Fundamentals of Human Physical Growth on Mars

Enrique Canessa

Abdus Salam International Centre for Theoretical Physics, Trieste, Italy

22.1 Introduction

Space colonization makes sense if we humans will be able to renew life and humanity beyond the Earth. We can foresee the possibility of human growth on the Martian surface since this planet has an environment that could be suitable to host humans and some other organic organisms (Bennett et al. 2003, Tokano 2005). The relative similarities from Earth - including the availability of ground ice on Mars, a Martian rotational period of only 37 minutes longer than 24 hours, and the presence of Mars seasons much like those on Earth (though about twice longer because the Martian orbital period is 1.88 Earth years), can make Mars arguably an ideal hospitable planet. There are also important differences to consider in any future Mars colonization. A low average surface temperature of -53°C due to the Mars eccentric orbit, the reduced atmospheric pressure of 0.005 bar, and a surface gravity of a little more than one third that of the Earth, namely $g_{\text{Mars}}=3.7 \text{ m/s}^2$ (or 38% of the mean gravitational acceleration on Earth).

Gravity can be seen as an important local resource available to humans and organic life on Mars. The effects of a Mars gravity, though lower, could be enough to prevent health problems associated with weightlessness. In this regard experiments testing the effects of a Mars-like gravity on a population of mice are under development in the Mars Gravity Biosatellite project (website: www.marsgravity.org). We need to exploit and take advantage of the existing force of gravity to facilitate Mars exploration and its subsequent colonization throughout the complete life cycle of humans, animals, insects, plants and other organisms. For example, gravitational pull is well known to influence plant roots to grow downward on Earth - a phenomenon known as *Gravitropism* for plant movement triggered by gravity stimuli.

In this chapter we show how physics laws also seem to influence the physical growth of humans under a given gravity field g by considering the human body as a dynamic system of variable mass m (and volume). The theory is based on

Newton's second law and it was first introduced in Canessa (2007) for modeling body mass index (BMI) on Earth (i.e., ratio of human "weight" and height squared: $B(t) \equiv w/h^2$). We reproduce and revise here this theory, and extend it using similar notation to the physical growth on Mars. To some extent, one can also study theoretically an equivalent human growth on Mars by applying universal physics laws as done previously on planet Earth.

We deduce analytical relations for the dependence on age (i.e., time t) of human height $h(t)$, "weight" $w(t) = m(t)g$ and combinations of them such as the weight-for-height and BMI. From the derivative dh/dt , then the growth velocity v follows, which depends on common parameters that could relate to environment, genetics and natural processes (known on Earth) like metabolism, energy supplied by food, etc. For comparison, physics estimates of human growth velocity on Earth and their "predicted" counterparts on Mars are given by considering an static smaller gravity field of 0.38 g (Martin et al. 2009). The limiting case of physical growth in an essentially weightless environment ($g \rightarrow 0$) is also discussed in connection to the Laird-Gompertz formula valid on Earth (Regner 1980).

22.2 Theory

For open systems like those in which there is influx of mass, one can apply the conservation of linear momentum to study their growth dynamics. Examples in which momentum is gained from, or lost to, the surroundings include rockets (Halliday and Resnick 1978, Sousa 2002) and the falling of a snow ball (Diaz et al. 2003), respectively. The human body on Earth is a system of fluctuating mass (and volume), under the influences of the acceleration of gravity g and food consumption. It is not unreasonable to consider it also as a dynamic physics system and then make an attempt to derive a relationship based on physics laws and observed physical growth.

We consider the human body as a system of variable mass $m(t)$, where mass is the amount of matter present in the body measured using balance. In everyday usage, weight and mass are used interchangeably. The reported "weight" in the smoothed growth charts for humans (see e.g. website: www.cdc.org) corresponds to measured matter or mass in kilogram and not to the weight " $w(t) = m(t)g$ " measured in Newton. In the following such "weight" data is referred as measured mass for the sake of correctness. In scientific usage mass is an intrinsic property of matter and weight is a force that results from the action of gravity on matter. Changes of mass and momentum will be assumed to be continuous during the growing process. The surface of the earth is taken as the zero level of gravitational potential energy.

To derive the equation of motion of the body system whose mass is not constant, we then use Newton's second law to define the external force F on the human body as derived explicitly by Halliday and Resnick (1978), namely

$$m \frac{dv}{dt} = \mathbf{F}_{ext} + \mathbf{v}_{rel} \frac{dm}{dt} , \quad (22.1)$$

where $dm/dt > 0$ is the rate at which mass is gain (i.e., equivalent to particles entering an open system) and $\mathbf{v}_{rel}(t)$ is the relative velocity of the gained mass with respect to the body moving with a growth velocity of magnitude $v(t) \equiv |\mathbf{v}(t)|$. The above term $\phi(t) \equiv v_{rel} (dm/dt)$ corresponds to the momentum flux being transferred into the body by the added mass. This is interpreted as the force exerted on the body by the mass that joins it (Sousa 2002) and \mathbf{F}_{ext} is the external force of gravity acting on the variable mass system. When the volume is allowed to vary, a term accounting for the additional momentum that the system acquires as a result of its changing volume (distict from the momentum flux $\phi(t) \equiv v_{rel} (dm/dt)$) needs to be added to Eq. (22.1) (Tiersen 1969).

In applying the Newton law for the vertical growth of humans, it corresponds to the analysis of an ensemble of N particles rather than to the analysis of a single particle of mass m_i and velocity \mathbf{v}_i at time t . In this case one can only consider averaged values for all the physical quantities -such as the external force, due to the complexity of the body system. Hence, if $n(h_i)$ is the number of objects at various heights h_i , of a total of n objects distributed among ℓ heights, then the mean average value of any given function f for those objects is in general

$$\hat{f} \equiv \langle f_i \rangle = \frac{1}{n} \sum_{i=1}^{\ell} n(h_i) f_i = \sum_{i=1}^{\ell} \frac{n(h_i)}{n} f_i = \sum_{i=1}^{\ell} \rho_i f_i , \quad (22.2)$$

where ρ_i is interpreted as the probability that a fraction of the objects $n(h_i)/n$ has the height h_i . So as the sum of all fractions of the whole system is unity ($f=1$), the probability distribution function is also normalized such that $\sum_{i=1}^{\ell} \rho_i = 1$. It can be seen that ρ_i directly affect the outcome of averaged results since it weights the values of the given function f at each height i . The above mean value equation is used throughout the theory of statistical physics.

Let us consider next the number of n objects distributed among ℓ heights as being N sets or groups of particles of added mass m_i each at time t . Then the mean value for the external force of Eq. (22.1) in one longitudinal direction becomes

$$\hat{F}_{ext} \equiv \langle F_{i,ext} \rangle = \sum_i^N \rho_i \left\{ m_i \frac{dv_i}{dt} - v_{i,rel} \frac{dm_i}{dt} \right\} . \quad (22.3)$$

Statistically stationary states in which all the probabilities ρ are time independent fo all i will be only considered.

To an observer at rest on the planet, the accumulating particles in the body at a given time appear to move at a velocity, say $v_a(t)$, proportional to the growth velocity $v(t)$ of the center of mass. For simplicity let us consider an homogenous system so that such a proportionality is set equal to one, i.e. $v_a(t) \approx v(t)$, which corresponds to the maximum velocity that the added mass alone can achieve during the grow. This is so because the human growth in stature is longitudinal until reaching a maximum height $h(\infty)$ with respect to earth's surface (and a maximum mass $m(\infty)$ under healthy conditions) at adolescence. On the other hand, to an observer on the center of mass of the body, the added mass appears to have a vertical motion with a velocity of magnitude $v(t) + v_a(t) \approx 2v(t)$ in the opposite direction of the growth with respect to the center of mass displacements. Therefore, for the present purposes, the relative velocity $v_{rel}(t)$ of the gained mass appears to move at twice the velocity of growth as seen from within the body system, namely $|v_{rel}| = 2|v|$, or for each set of particles of variable mass

$$v_{i,rel} = 2v_i \quad (i=1, \dots, N) . \quad (22.4)$$

Hence Eq. (22.3) for the mean value of the net external force on the body as a whole is the superposition of contributions

$$\hat{F}_{ext}(t) = \langle F_{i,ext}(t) \rangle = \sum_{i=1}^N \rho_i m_i(t) \frac{dv_i(t)}{dt} - 2 \sum_{i=1}^N \rho_i v_i(t) \frac{dm_i(t)}{dt} . \quad (22.5)$$

with ρ the normalized probability distribution function of N objects or set of particles of added mass m_i each at time t .

The force of gravity F_g due to the planet mass acting thoughout the human body must be also taken into account. Similarly to suspended chains systems (Sousa 2002), this force is not only proportional to the total weight of the body at an given height $h(t)$, but also to the total force due to the amount of mass at rest lying on the surface of the earth. This extra term can also be understood by the fact that the volume (besides the mass) of the human body also varies. It accounts for an additional momentum that the system acquires as a result of changing volume (in one-dimension vertical) with years. In this case, the force of gravity acting on the N objects of added mass m_i is then given by

$$\hat{F}_g(t) = \sum_{i=1}^N \left\{ \rho_i \frac{h_i(t)}{h_i(\infty)} m_i(t) \right\} g - w(\infty) = \langle \lambda_i(t) h_i(t) \rangle g - m(\infty) g , \quad (22.6)$$

where $\lambda_i(t) \equiv m_i(t)/h_i(\infty)$ represents mass per unit height. The weight at rest term $w(\infty) = m(\infty)g = \langle m_i(\infty) \rangle g = \sum_{i=1}^N \rho_i m_i(\infty) g$ is deduced according to the boundary condition $\hat{F}_g \rightarrow 0$ as $t \rightarrow \infty$.

22.3 Biometric Relations

In order to derive analytical relations for the dependence of stature, measured mass (weight), growth velocity (and their mix as the body mass index) on age, we proceed by considering Newton's law of motion in relation to the two distinct mean forces acting on the system, *i.e.* the net external force of Eq. (22.5) and the gravity force of Eq.(22.6). This means to set $\hat{F}_g = \hat{F}_{ext}$, or more explicitly

$$\sum_{i=1}^N \rho_i \left\{ m_i(t) \frac{d^2 h_i(t)}{dt^2} - 2 \left(\frac{dh_i(t)}{dt} \right) \left(\frac{dm_i(t)}{dt} \right) - \frac{h_i(t)}{h_i(\infty)} m_i(t) g \right\} = -m(\infty) g , \quad (22.7)$$

with $v_i(t) = dh_i(t)/dt$.

In order to derive possible solutions of this master equation an additional hypothesis is needed since this equation contains two unknown functions: $m_i(t)$ and $h_i(t)$ for each set of particles. Let us consider

$$h_i(\infty) = \frac{1}{2} g \tau_i^2 , \quad (22.8)$$

with τ_i time lags for i sets of particles. The mean value of τ_i is deduced from the limit $t \rightarrow \infty$, which leads to $h(\infty) = \langle h_i(\infty) \rangle = \sum_{i=1}^N \rho_i h_i(\infty) = g \langle \tau_i^2 \rangle / 2$. In this way, the mean time lag $\langle \tau_i \rangle$ is interpreted as the time at which the growth velocity becomes zero and the maximum body height is achieved with respect to the earth's surface $h(0) = 0$ (recall that $x(t) = x(0) + v(t) \times t + (1/2)a(t) \times t^2$ from Newton's mechanics for constant mass systems). On the other hand, t_i are the time lags at which the single height contributions $h_i(t)$ of the N objects or set of particles become zero.

Using the well-known trigonometric relation: $\text{sech}^2 + \tanh^2 = 1$ and the above condition for $h_i(\infty)$, it is then straightforward to check that our second order differential equation of Eq. (22.7) has the general solutions

$$h(t) = \langle h_i(t) \rangle = \sum_{i=1}^N \rho_i h_i(\infty) \tanh \left(\frac{t - t_i}{\tau_i} \right) , \quad (22.9)$$

and

$$m(t) = \langle m_i(t) \rangle = \sum_{i=1}^N \rho_i m_i(\infty) \tanh \left(\frac{t - t_i}{\tau_i} \right) . \quad (22.10)$$

and from Eq. (22.9), we also have

$$v(t) = \frac{dh(t)}{dt} = \left\langle \frac{dh_i(t)}{dt} \right\rangle = \sum_{i=1}^N \rho_i \frac{h_i(\infty)}{\tau_i} \operatorname{sech}^2 \left(\frac{t-t_i}{\tau_i} \right) , \quad (22.11)$$

which for $t \rightarrow \infty$, it gives $v(\infty) \rightarrow 0$ as expected (after completion of the physical growth).

It follows immediately that physics theory predicts that the height and mass functions reach maximum values (plateau) on age due to the presence of the hyperbolic tangent functions. The mass per unit height $\lambda_i(\infty) = m_i(\infty)/h_i(\infty) \neq 0$ corresponds to the slope of $m_i(t)$ versus $h_i(t)$. For systems of variable mass, the center of mass velocity $v(t)$ is not the same as that for the N sets of particles in the system v_i . In turn, the moving particles may not have the same velocities for each state i .

22.4 "Biological" Parameters

The amount of growth parameters in the present physics based biometric formulae may be seen to be large. However, these are not only comparable with those in the alternative asymptotic mathematical models tested for the modeling of human anthropometric data (see, e.g., Flegal 1999, Perevozskaya and Kuznetsova 2002), but their number can still be reduced - at least for small N sets of particles.

It is easy to verify by a plot that for $y \geq 0$, the monotonicity of the hyperbolic tangent implies the relation (Malisoff and Sontag 2000)

$$\tanh y \leq 1 \leq \frac{\tanh(xy)}{\tanh x} . \quad (22.12)$$

Therefore, using this inequality it is possible to group together model parameters to set lower bounds to all functions in this work. In fact, by setting $y = \tau/\tau_i \rightarrow 1/\varepsilon_i \geq 0$ (independent of t) and $x \rightarrow (t-t_i)/\tau$, such that $\tau_i \equiv \varepsilon_i \tau$, one obtains

$$\begin{aligned} \tanh(xy) &= \tanh \left(\left[\frac{t-t_i}{\tau} \right] \left[\frac{\tau}{\tau_i} \right] \right) \\ &\geq \tanh \left(\frac{\tau}{\tau_i} \right) \tanh \left(\frac{t-t_i}{\tau} \right) = \tanh y \tanh x . \end{aligned} \quad (22.13)$$

Hence Eqs. (22.9) and (22.10) can be approximated as

$$h(t) \geq \sum_{i=1}^N A_i \tanh \left(\frac{t-t_i}{\tau} \right) . \quad (22.14)$$

and

$$m(t) \geq \sum_{i=1}^N B_i \tanh\left(\frac{t-t_i}{\tau}\right), \quad (22.15)$$

where, for all i , it is defined

$$A_i \equiv \rho_i h_i(\infty) \tanh\left(\frac{1}{\varepsilon_i}\right) = \rho_i h_i(\infty) \tanh\left(\frac{\tau}{\tau_i}\right), \quad B_i \equiv \lambda_i(\infty) A_i, \quad (22.16)$$

with λ is mass per unit height as before, which makes B_i to be independent of gravity g . Since the τ_i parameters are related to h_i via the Newtown equation in Eq. (22.8), then this leads to the positive dimensionless quantities $\varepsilon_i = (1/\tau)\sqrt{2h_i(\infty)/g}$. Furthermore, the momentum flux transferred to the body by the added m_i becomes $\phi_i(t) \equiv v_i (dm_i/dt) = (B_i/A_i) v_i(t)^2$, similarly to nonrigid systems (Sousa 2002). Taking the derivative of Eq. (22.14) it follows that

$$v(t) \geq \frac{1}{\tau} \sum_{i=1}^N A_i \operatorname{sech}^2\left(\frac{t-t_i}{\tau}\right). \quad (22.17)$$

which is consistent with Eq. (22.11) in the lower bound since $(1/\tau_i) \operatorname{sech}^2(xy) \geq (1/\tau) \tanh y \operatorname{sech}^2 x$.

In this way, the unknown growth parameters $h_i(\infty)$, $m_i(\infty)$ and ρ_i can reduce to A_i and B_i , which will be referred to as "*biological*" parameters. Together with the time lag τ and the peak positions of each tanh function, given by t_i , these parameters allow to describe non-linear percentile curves.

At the origin, the boundary conditions $h(0)=0$ and $m(0)=0$ in Eqs. (22.14) and (22.15) lead to the relations between the biological parameters

$$A_1 \tanh\left(\frac{t_1}{\tau}\right) \leq \sum_{i=2}^N A_i \tanh\left(\frac{t_i}{\tau}\right); \quad B_1 \tanh\left(\frac{t_1}{\tau}\right) \leq \sum_{i=2}^N B_i \tanh\left(\frac{t_i}{\tau}\right), \quad (22.18)$$

where $t_1 < 0$ and $t_2, \dots, t_N > 0$ which leads to reduce further the number of parameter values. In fact, as $t \rightarrow \infty$, $\tanh t_{i>N} \rightarrow 1$ and $B_{i>N} \rightarrow 0$ for finite N sets of particles (of added matter $m_{i \leq N}$).

22.5 Estimates of Human Growth on Earth and Mars

The raw (and theoretically smoothed) statistical data of different aspects of human physical growth from birth to adolescence have been collected for decades as a function of gender and nationality. There are smoothed growth charts for boys and

girls such as those maintained by the USA National Center of Health Statistics (NCHS) (website: www.cdc.org). Typically, the charts consist of a set of non-linear percentile curves displaying the dependence on age of height $h(t)$, "weight" $w(t)$ and combinations of them. According to these growth charts, the 85th percentile of BMI for children is considered the overweight threshold, and the 95th percentile is the obesity threshold on Earth.

As an example we plot in Fig. 22.1 theoretical curves (full lines) and selected smoothed empirical percentiles of growth velocity as a function of chronological age in months from the boys data reported on the medical literature. The curves are obtained using Eq. (22.17) assuming $N = 6$ sets of particles with biological parameters A_i/B_i of Eq. (22.16) as previously used by Canessa (2007). These parameter values can easily be derived using any fit data plotting software. The equivalent growth on Mars as "predicted" by the Newton laws of motion are also shown by full lines assuming common boundary conditions and a reduced Mars gravity of 38% g .

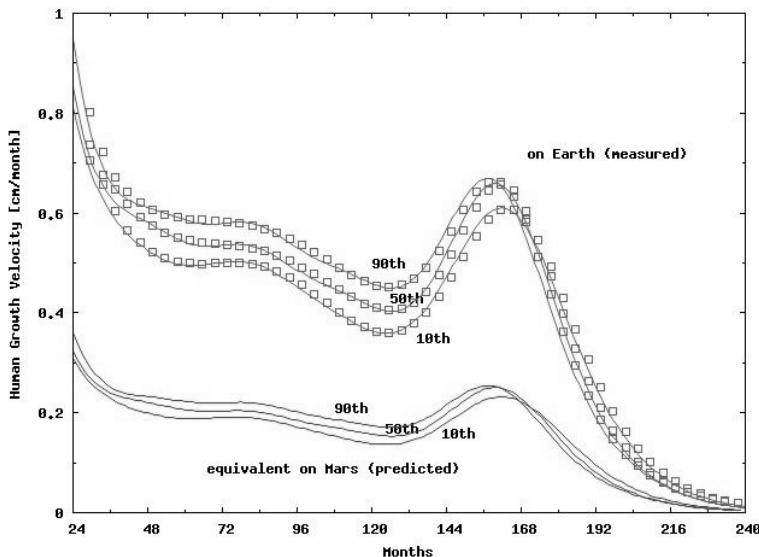


Fig. 22.1 Growth velocity curves for boys on Earth at selected percentiles. Real data (dots) collected for decades using smoothed charts from NCHS and theoretical results (full lines) from Eq.(22.17). The equivalent growth on Mars as "predicted" by the same formulae assuming common biological parameters and the Mars gravity $g=3.7 \text{ m/s}^2$ is also shown by full lines.

Within measurements error, the variation of estimates of growth charts on Earth is close to all smoothed empirical data in Fig. 22.1 usually measured to the nearest 0.1 cm (Xu et al. 2002). The growth velocities on Earth oscillate down as a

function of time to vanish in the adolescence. However, $v(t)$ on Mars may become almost constant for several years prior to puberty (about 14 years old in the figure). The system acceleration dv/dt changes sign over the years due to the tanh term in the derivative of Eq. (22.17) whereas both the real and the mass-for-height model curves increase monotonically. Increasing N in the sums of Eqs. (22.15) and (22.14) allows a better minima fit at higher percentiles of the BMI-for-age curves on Earth.

The main consequence of an equivalent human growth and development on Mars (or, in general, under a reduced gravity environment), is the substantial 60% reduction on the magnitude of the growth velocity as compared to terrestrial measurements assuming analogous biological parameters. The human height on Mars is also "predicted" to be shortened via Eq. (22.14). The human mass described by Eq. (22.15) behaves similarly on Earth and Mars since from Eq. (22.16) B_i is independent of $h(\infty)$ and, thereafter, independent of the gravitational field (the "weight" $w(t) = m(t)g$ changes instead). Humans on Mars will then grow much shorter in stature having a growth velocity much slower than on Earth and almost constant for several years.

According to the non-linear results in the plot (and those by Canessa 2007), we can argue that available human growth data is correlated in relation to physics laws over many years and at different percentiles. The underlying physics of Newtonian mechanics is the base of the derivation of all equations. The fundamental function, $\tanh(x)$ appearing throughout the resulting relations for the body height in Eq. (22.9), mass in Eq. (22.10) and BMI thereafter is not arbitrary. On the other hand, the peaked growth velocity of Eq. (22.11) depends on the different hyperbolic function $\text{sech}(x)$.

The primary interest in human physical growth and its relationship to physics laws is to enhance the fit quality to observed data and to understand the underlying phenomena. The claim here is not modified by any class of ρ_i distributions - relating the biological parameters in Eq. (22.16). In principle, one may associate these parameters to changes in kinetic energy of the center of mass of the system (Sousa 2002).

22.6 The Limit of Null Gravity

The relationship between natural laws and human physical growth under a gravity field may be also suitable to understand the biometrics of other living species growing vertically such as Primates (Leigh 2001) and perhaps leaving trees and plants too. Furthermore, assuming a 1/3 reduced g as on Mars all the above equations should ideally be valid. At this point, it is also worth to ask if the same physics laws discussed already can also be applied in the "absence" of gravity, i.e., $g \rightarrow 0$. This includes the case of horizontally growing terrestrial organisms, i.e., aquatic organisms growing in an essentially weightless environment (Ferrero 1996).

Let us consider the general Eq. (22.5) for \hat{F}_{ext} again and set the force of gravity equal to $\hat{F}_g = 0$. Alternatively, from Eq. (22.6) this implies $h_i(t) \rightarrow h_i(\infty) = \text{const}$. which means to compensate out the weight $w(t)$ with other forces. Therefore

$$\sum_{i=1}^N \rho_i m_i(t) \frac{dv_i(t)}{dt} - 2 \sum_{i=1}^N \rho_i v_i(t) \frac{dm_i(t)}{dt} = 0 \quad . \quad (22.19)$$

One solution satisfies

$$\frac{dv_i(t)}{v_i(t)} = 2 \frac{dm_i(t)}{m_i(t)} \quad , \quad (22.20)$$

which, by integration, implies the relation

$$v(t) = \langle v_i(t) \rangle \equiv \langle \frac{dx_i(t)}{dt} \rangle = \gamma \langle m_i^2(t) \rangle \quad , \quad (22.21)$$

with γ a given constant and for the x -axis of a horizontal structure.

On the other hand, let us consider (and expand in Taylor series) the well-known Laird-Gompertz empirical formula for fish growth on Earth (Regner 1980, Ferrero 1996). That is

$$\frac{x(t)}{x_0} = e^{a(1-e^{-\alpha t})} \approx 1 + a(1 - e^{-\alpha t}) \approx 1 + a \left(1 - \frac{1}{1 + \alpha t} \right) \quad , \quad (22.22)$$

where $x(t)$ is the lenght at age t (at zero gravity) and (a, α) are specific rates of the exponential growth. It follows that the growth velocity $v(t) \equiv dx(t)/dt$ satisfies

$$v(t) = \frac{ax_0\alpha}{(1+\alpha t)^2} = \frac{\alpha}{ax_0} [(1+a)x_0 - x(t)]^2 \quad . \quad (22.23)$$

Using the Chebyshev's sum inequality $\sum_{k=1}^n a_k b_k \geq (1/n) \sum_{k=1}^n a_k \sum_{k=1}^n b_k$, such that $a_k = b_k = \rho_i m_i \equiv \rho m_i$, then comparison of Eqs. (22.21) and (22.23) results in $\langle m_i(t) \rangle^2 \leq N\rho \sum_{i=1}^N \rho m_i(t)^2 = N\rho \langle m_i^2(t) \rangle$ and

$$m(t) \leq \left(\frac{N\rho\alpha}{ax_0\gamma} \right)^{1/2} |(1+a)x_0 - x(t)| \quad . \quad (22.24)$$

which for values $x(t) > (1+a)x_0$, it implies an approximated linear relation between mass and length for growth under $g \rightarrow 0$. To this end, collected data for Dolphins (Ferrero 1996) suggest that their postnatal dentine thinkness (used as an index of age in these animals) increases linearly with the total body length. It is reasonable to justify then the linear proportion between the aquatic species mass

and length of the type in Eq. (22.24), as a consequence of the application of physical laws (c.f., Eq. (22.5)), in conjunction with the Laird-Gompertz growth function (c.f., Eq. (22.22)). Equations (22.22) and (22.24) also allows to study the BMI for aquatic organisms and Eq. (22.23) and relate their growth velocity curves in terms of the square of the mass (c.f., Eq. (22.21)) - at least within a period of time in the physical growth phenomena. These results could be particularly interesting for future studies on aquatic animals growth on Mars.

22.7 Concluding Remarks

Everybody living on the surface of the Earth experiences some ground reaction constraints. The force of gravity on our bodies of variable mass (and volume) is not only proportional to the weight at a given height, but also to the total force due to the amount of mass at rest lying on the surface of the planet (similarly to suspended chains systems). Based on an original reference to Newton's law of motion (Canessa 2007), we extended this theory to the problem of human physical growth on the reduced gravitational field of Mars. In this context we deduced biometric relations for the temporal dependence of human height, mass, growth velocity and combinations of these (such as BMI). We also analysed growth in a completely weightless environment.

Within the Newtown-based approach we have not assumed a unique state of change for all metric aspects of the body. The observed allometric human growth (e.g., length of the head of a newborn being 25% of total body length and 15% on adults) (Lampl et al. 1992, Cameron 2002), can be represented by ensemble averages. These are here composed of N different sections (i -sum of particles in Eq. (22.2) behaving differently each). The complexities created by the adolescent growth spurt on Earth, and "predicted" on the red planet under similar conditions, has been dealt with by the magnitude of the so-called "biological parameters" A_i, B_i .

Our understanding of human growth as depicted in Fig. 22.1, depends on the frequency of measurements which are usually done at intervals of few months to reduce errors. The resultant curves are smooth and continuous reflecting that growth is a continuous process from conception to adulthood. They also demonstrate the well known changes in growth velocity at clearly discernable times such as mid-childhood and adolescence. Assuming analogous biological conditions, the main consequence of an equivalent human growth and development under the partial gravity of Mars, is a substantial reduction on the growth velocity magnitude as compared to measurements on Earth. Humans on the surface of Mars will most likely grow much shorter in height and will show a slowing down growth velocity behaviour.

Given the likelihood of human colonization in at least few Martian locations having an appropriate ecosystem, a weaker gravity field could be in fact an advantage. To be born and raised smaller on Mars, as well as to weight less, could be positive in terms of lower energy consumption and food needs. Any "*flora et fauna*" on Mars is foreseen to be similarly influenced.

References

- Bennett, J., Shostak, S., Jakosky, B.: Life in the Universe. Addison Wesley, San Francisco (2003)
- Cameron, N. (ed.): Human Growth and Development. Academic Press, London (2002)
- Canessa, E.: Modeling of Body Mass Index by Newton's Second Law. *J. Theor. Biology* 248, 646–656 (2007)
- Diaz, R.A., Gonzalez, D.L., Marin, F., Martinez, R.: Comparative kinetics of the snowball respect to other dynamical objects (2003) (Preprint) arXiv:physics/0310010 v2
- Ferrero, R.: Growth Patterns of the Pacific White-sided Dolphin in the Central North Pacific Ocean. AFSC Quarterly Report October-November-December (1996), <http://www.afsc.noaa.gov/quarterly>
- Flegal, K.M.: Curve smoothing and transformations in the development of growth curves. *Am. J. Clin. Nutr.* 70, 163S–165S (1999)
- Halliday, D., Resnick, R.: Physics -I,II, ch. 9. Wiley, New York (1978)
- Lampl, M., Veldhuis, J.D., Johnson, M.L.: Saltation and Stasis: A Model of Human Growth. *Science* 258, 801–803 (1992)
- Leigh, S.R.: Evolution of Human Growth. *Evol. Anthr.* 10, 223–236 (2001)
- Malisoff, M., Sontag, E.D.: Universal formulas for feedback stabilization with respect to Minkowski balls. *Systems and Control Letters* 40, 247–260 (2000)
- Marty, J.C., Balmino, G., Duron, J., Rosenblatt, P., Le Maistre, S., Rivoldini, A., Dehant, V., Van Hoolst, T.: Martian Gravity Field Model and its Time Variations from MGS and Odyssey Data. *Planetary and Space Sci.* 57, 350–363 (2009)
- Perevozskaya, I., Kuznetsova, O.M.: Modeling Longitudinal Growth Data and Growth Percentiles with Polynomial Gompertz Model in SAS Software. *Stat and Data Analysis Paper* 277-25 (2002)
- Regner, S.: On Semigraphic Estimation of Parameters of Gompertz Function and its Application on Fish Growth. *Acta Adriat.* 21, 227–236 (1980)
- Sousa, C.A.: Nonrigid systems: mechanical and thermodynamic aspects. *Eur. J. Phys.* 23, 433–440 (2002)
- Tiersten, M.S.: Force, Momentum Change, and Motion. *Amer. J. Phys.* 37, 82–87 (1969)
- Tokano, T. (ed.): Water on Mars and Life. *Adv. Astrobiol. Biophys.* Springer, Heidelberg (2005)
- Xu, X., Wang, W., Guo, Z., Karlberg, J.: Longitudinal Growth During Infancy and Childhood in Children from Shanghai: Predictors and Consequences of the Age at Onset of the Childhood Phase of Growth. *Pediatric Res.* 51, 377–385 (2002)

Chapter 23

Artificial Environments on Mars

Alexander A. Bolonkin

C&R, New York, USA

23.1 Introduction

This chapter has two independent sections, important to Mars resource utilization. Section 23.2 refers to inflatable domes for Mars as well as for Mars satellites. Section 23.3 is devoted to the proposed AB method of agriculture on Mars without added water (based on a closed-loop water cycle). Both sections are based on a common idea, i.e. the isolation of a limited area by a cheap transparent inflatable dome and creating more suitable conditions (in other words, micro-terraforming) inside this dome for humans and agriculture.

23.2 Inflatable Dome for Mars and Mars Satellites

23.2.1 Brief Description of Problem

This section is based on the previous paper (Bolonkin 2007d). The large-scale development of permanent human life on Mars requires two conditions: (i) all-sufficient space settlement and (ii) artificial life conditions close to those prevailing currently on the Earth. The first condition demands production of all main components needed for human life: food, oxidizer, and energy within the outer space and Solar System body colony. The second requisite condition is a large surface settlement having useful plants, flowers, water pool, walking and sport areas, etc. All these conditions may be realized within large 'greenhouses' (Bolonkin 2006a,b,c; Bolonkin and Cathcart 2006a,b,c,d) that will produce food, oxidizer and all needed conditions of "the good life". Settlement will not be attractive to potential settlers unless free of privation.

In other words, future human lifestyles on Mars must be more comfortable than commonly projected now for humans to explore and properly exploit this planet. Human life there will be more comfortable if employing the author's macro-project proposal - staying in Martian atmosphere without a special spacesuit (mass of current spacesuit reaches 180 kg, cost \$22 millions) (Bolonkin 2006b,c, 2008; Bolonkin and Cathcart 2006c).

The current conditions in Mars are far from comfortable. For example, Mars does not have a terrestrial-like atmosphere, there is space radiation, etc. Especially during wintertime, Mars could provide only a meager and uncomfortable life-style for humans, showing low temperatures and strong winds.

23.2.2 Brief Information about Mars

Mars' average distance from the Sun is roughly 230 million km (1.5 AU) and its orbital period is 687 (Earth) days. The solar day (or sol) on Mars is only slightly longer than an Earth day: 24 hours, 39 minutes, and 35.244 seconds. A Martian year is equal to 1.8809 Earth years, or 1 year, 320 days, and 18.2 hours. Mars's axial tilt is 25.19 degrees, which is similar to the axial tilt of the Earth (other details are given in Chap. 2).

Mars is a terrestrial planet with a thin atmosphere, having surface features reminiscent both of the impact craters of the Moon and the volcanoes, valleys, deserts and polar ice caps of Earth. It is the site of Olympus Mons, the highest known mountain in the Solar System, and of Valles Marineris, the largest canyon. In addition to its geographical features, Mars' rotational period and seasonal cycles are likewise similar to those of Earth. Still, of all the planets in the Solar System other than Earth, Mars is the most likely to harbor liquid water, and perhaps life.

Although Mars has no intrinsic magnetic field, observations show that parts of the planet's crust have been magnetized and that alternating polarity reversals of its dipole field have occurred. Current models of the planet's interior imply a core region about 1,480 kilometres in radius, consisting primarily of iron with about 14–17% sulfur. Mars is also scarred by a number of impact craters: a total of 43 000 craters with a diameter of 5 km or greater have been found.

The red-orange appearance of the Martian surface is caused by iron(III) oxide, more commonly known as hematite, or rust. Based on orbital observations and the examination of the Martian meteorite collection, the surface of Mars appears to be composed primarily of basalt. Some evidence suggests that a portion of the Martian surface is more silica-rich than typical basalt, and may be similar to andesitic rocks on Earth; however, these observations may also be explained by silica glass. Much of the surface is deeply covered by a fine iron(III) oxide dust that has the consistency of talcum powder.

In June, 2008, the Phoenix Lander returned data showing Martian soil to be slightly alkaline and containing vital nutrients such as magnesium, sodium, potassium and chloride, all of which are necessary for living things to grow. Scientists compared the soil near Mars' north pole to that of backyard gardens on Earth, saying it could be suitable for plants such as asparagus. However, in August, 2008, the Phoenix Lander conducted simple chemistry experiments, mixing water from Earth with Martian soil in an attempt to test its pH, and discovered traces of the salt perchlorate. Its presence, if confirmed, would appear to make the soil more exotic than previously believed. Further testing is necessary to eliminate any possibility of the perchlorate readings being influenced by terrestrial sources which may have migrated from the spacecraft, either into samples or into the instrumentation.

Water, in the state of ice, was found by the Phoenix Mars Lander on July 31, 2008. Geological evidence gathered by various missions suggests that Mars previously had large-scale water coverage, while observations also indicate that small geyser-like water flows have occurred during the past decade. Observations by NASA's *Mars Global Surveyor* show evidence that parts of the southern polar ice cap have been receding. Liquid water cannot exist on the surface of Mars with its present low atmospheric pressure, except at the lowest elevations for short periods but water ice is in no short supply, with two polar ice caps made largely of ice. In March 2007, NASA announced that the volume of water ice in the south polar ice cap, if melted, would be sufficient to cover the entire planetary surface to a depth of 11 metres. Additionally, an ice permafrost mantle stretches down from the pole to latitudes of about 60°.

The polar caps at both poles consist primarily of water ice. However, there is dry ice present on their surfaces. Frozen carbon dioxide (dry ice) accumulates as a thin layer about one metre thick on the north cap in the northern winter only, while the south cap has a permanent dry ice cover about eight metres thick. The northern polar cap has a diameter of about 1000 kilometres during the northern Mars summer, and contains about 1.6 million cubic kilometres of ice, which if spread evenly on the cap would be 2 kilometres thick. (This compares to a volume of 2.85 million cubic kilometres for the Greenland ice sheet.) The southern polar cap has a diameter of 350 km and a thickness of 3 km. The total volume of ice in the south polar cap plus the adjacent layered deposits has also been estimated at 1.6 million cubic kilometres. Both polar caps show spiral troughs, which are believed to form as a result of differential solar heating, coupled with the sublimation of ice and condensation of water vapor. Both polar caps shrink and regrow following the temperature fluctuation of the Martian seasons.

Mars atmosphere. Mars lost its magnetosphere 4 billion years ago, so the solar wind interacts directly with the Martian ionosphere, keeping the atmosphere thinner than it would otherwise be by stripping away atoms from the outer layer. Both Mars Global Surveyor and Mars Express have detected these ionised atmospheric particles trailing off into space behind Mars. The atmosphere of Mars is now relatively thin. Atmospheric pressure on the surface varies from around 30 Pa (0.03 kPa) on Olympus Mons to over 1155 Pa (1.155 kPa) in the depths of Hellas Planitia, with a mean surface level pressure of 600 Pa (0.6 kPa). This is less than 1% of the surface pressure on Earth (101.3 kPa). Mars's mean surface pressure equals the pressure found 35 km above the Earth's surface. The scale height of the atmosphere, about 11 km, is higher than Earth's (6 km) due to the lower gravity. Mars' gravity is only about 38% of the surface gravity on Earth.

The atmosphere on Mars consists of 95.72% Carbon dioxide, 2.7% Nitrogen, 1.6% Argon, 0.2% Oxygen, 0.07% Carbon monoxide, 0.03% Water vapor, 0.01% Nitric oxide, 2.5 ppm Neon, 300 ppb Krypton, 130 ppb Formaldehyde, 80 ppb Xenon, 30 ppb Ozone, 10 ppb Methane. The atmosphere is quite dusty, containing particulates about 1.5 µm in diameter which give the Martian sky a tawny color when seen from the surface.

During a pole's winter, it lies in continuous darkness, chilling the surface and causing 25–30% of the atmosphere to condense out into thick slabs of CO₂ ice (dry ice). When the poles are again exposed to sunlight, the frozen CO₂ sublimes, creating enormous winds that sweep off the poles as fast as 400 km/h. These seasonal actions transport large amounts of dust and water vapor, giving rise to Earth-like frost and large cirrus clouds. Clouds of water-ice were photographed by the *Opportunity* rover in 2004.

Of all the planets, Mars's seasons are the most Earth-like, due to the similar tilts of the two planets' rotational axes. However, the lengths of the Martian seasons are about twice those of Earth's, as Mars' greater distance from the Sun leads to the Martian year being about two Earth years in length. Martian surface temperatures vary from lows of about –140 °C (–220 °F) during the polar winters to highs of up to 20 °C (68 °F) in summers. The wide range in temperatures is due to the thin atmosphere which cannot store much solar heat, the low atmospheric pressure, and the low thermal inertia of Martian soil. The summer temperatures in the south can be up to 30 °C (54 °F) warmer than the equivalent summer temperatures in the north.

Mars also has the largest dust storms in our Solar System. These can vary from a storm over a small area, to gigantic storms that cover the entire planet. They tend to occur when Mars is closest to the Sun, and have been shown to increase the global temperature.

Moons of Mars. Mars has two moons, Phobos and Deimos, which are small and irregularly shaped. They orbit very close to the planet and are thought to be captured asteroids. Both satellites were discovered in 1877 by Asaph Hall. From the surface of Mars, the motions of Phobos and Deimos appear very different from that of our own moon. Phobos rises in the west, sets in the east, and rises again in just 11 hours. Deimos, being only just outside synchronous orbit—where the orbital period would match the planet's period of rotation — rises as expected in the east but very slowly. Despite the 30 hour orbit of Deimos, it takes 2.7 days to set in the west as it slowly falls behind the rotation of Mars, then just as long again to rise. Because Phobos' orbit is below synchronous altitude, the tidal forces from the planet Mars are gradually lowering its orbit. In about 50 million years it will either crash into Mars' surface or break up into a ring structure around the planet.

23.2.3 ‘Evergreen’ Inflated Domes

Possibly the first true architectural attempt at constructing effective artificial life-support systems on the climatically harsh Mars will be the building of greenhouses. Greenhouses are maintained nearly automatically by heating, cooling, irrigation, nutrition and plant disease management equipment. Humans share commonalities in their responses to natural environmental stresses that are stimulated by night cold and day heat, for example. Darkness everywhere inflicts the same personal visual discomfort and disorientation. However, with special clothing and shelters, humans can adapt successfully to Mars. Incontrovertibly, living on the Red Planet is difficult, even when tempered by strong conventional protective buildings.

This section describes and theorizes the macro-engineering concept of inexpensive-to-construct-and-operate “Evergreen” inflated surface domes. Innovations are wanted, and needed, to make possible within such structures a recreation in space of our unique and continuously exploration on Mars.

23.2.4 Description and Innovations

Dome. Our design for a Mars people-housing “Evergreen” dome is presented in Fig. 23.1, which includes the thin inflated film dome concept. The innovations are listed here: (1) the construction is air-inflatable; (2) each dome is fabricated with very thin, transparent film (thickness is 0.2 to 0.4 mm) without rigid supports; (all in tension, not compression) (3) the enclosing film is a two-layered structural element with air between the layers to provide insulation; (4) the construction form is that of a hemisphere and part of the film has control transparency and a thin aluminum layer about 1 μm or less that functions as the gigantic collector of solar incident solar radiation. Surplus heat collected may be used to generate electricity or furnish mechanical energy; and (5) the dome is equipped with sunlight controlling louvers with one side thinly coated with reflective polished aluminum of about 1 μm thickness. Real-time control of the sunlight’s entrance into the dome and nighttime heat’s exit is governed by the shingle-like louvers or a controlled transparency of the dome film (made opaque to infrared at night to lessen heat loss).

Variant 1 of artificial inflatable Dome for Mars is shown in Fig. 23.1. Dome has top thin double film 4 covered given area and single under ground layer 6. The space between layers 4 - 6 is about 0.3 - 1 meters and it is filled by air. The support cables 5 connect the top and underground layers and Dome looks as a big air-inflated beach sunbathing or swimming mattress. The Dome includes hermetic sections connected by corridors 2 and hermetic air lock chambers 3. The top film has control transparency (reflectivity). That allows control temperature affecting the dome. The top film also is double. When a meteorite punches a hole into the top double film, the lower layer closes the hole and puts temporary obstacles in the way of escaping air. This dome has a fruitful soil layer, irrigation system, and cooling system 9 for support of air levels at a given humidity. This results in a closed biosphere with closed life circle that produces oxygen and food for people. Simultaneously, it is a beautiful Earthlike garden to refresh the soul of those weary of both machine-packed spaceships and lifeless alien rocks. The proposed design has a minimum specific mass, about $7-12 \text{ kg/m}^2$ (air - 3 kg, film - 1 kg, soil - 3 to 8 kg). Mass of the area $10 \times 10 \text{ m}$ is about 1 ton. Special shelters save people from cosmos radiation.

Figure 23.2 illustrates the second thin transparent dome cover we envisage. The Dome has double film: semispherical layer (low pressure about 0.005 - 0.01 atm) and lower layer (high pressure, i.e. about 1 atm). The hemispherical inflated textile shell—technical “textiles” 1 can be woven or non-woven (films)—embodies the innovations listed: (1) the film is very thin, approximately 0.1 to 0.3 mm; (2) the film has two strong nets, with a mesh of about $0.1 \times 0.1 \text{ m}$ and $a = 1 \times 1 \text{ m}$, the threads are about 0.3 mm for a small mesh and about 1 mm for a big mesh.

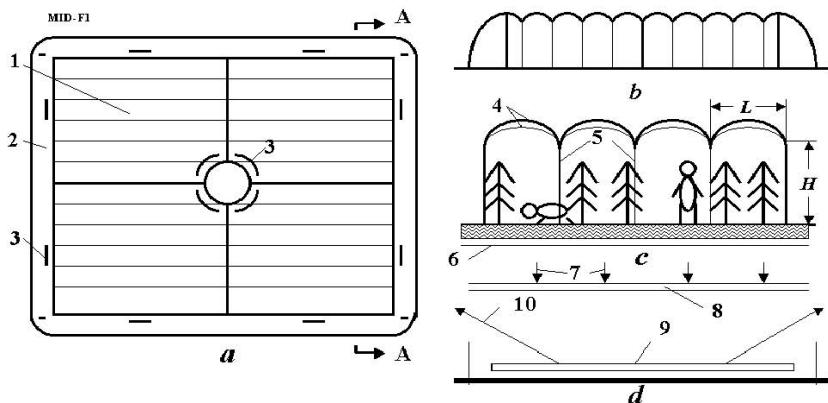


Fig. 23.1 Variant 1 of artificial inflatable Dome for Mars. (a) Top view of dome; (b) Cross-section AA area of dome; (c) Inside of the Dome; (d) Cooling system. Notations: 1 - internal section of Dome; 2 - gangways; 3 - doors; 4 - transparent thin double film ("textiles") with control transparency; 5 - support cables; 6 - lower under ground film; 7 - solar light; 8 - protection film; 9 - cooling tubes; 10 - radiation of cooling tubes.

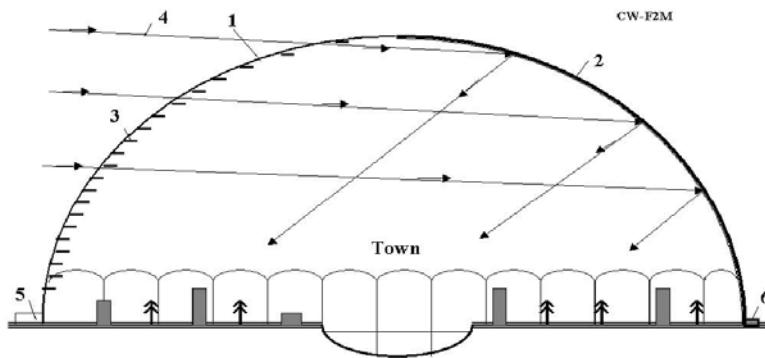


Fig. 23.2 Variant 2 of artificial inflatable Dome for Mars. Notations: 1 - transparent thin double film ("textiles"); 2 - reflected cover of hemisphere; 3 - control louvers; 4 - solar beams (light); 5 - enter (dock chamber); 6 - water extractor from air. The lower section has air pressure about 1 atm. The top section has pressure 0.01 - 0.1 atm.

The net prevents the watertight and airtight film covering from being damaged by micrometeorites (Fig. 23.3); the film incorporates a tiny electrically conductive wire net with a mesh about 0.001×0.001 m and a line width of about $100 \mu\text{m}$ and a thickness near $1 \mu\text{m}$. The wire net can inform the "Evergreen" dome repair supervisors (human or automated equipment) concerning the place and size of film

damage events (tears, rips, punctures); the film is twin-layered with the gap — $c \approx 1$ m and $b \approx 2$ m—between the layer covering. This multi-layered covering is the main means for heat insulation and anti-puncture safety of a single layer because piercing won't cause a loss of shape since the film's second layer is unaffected by holing; the airspace in the dome's twin-layer covering can be partitioned, either hermetically or not; and part of the covering may have a very thin shiny aluminum coating that is about $1\mu\text{m}$ for reflection of non-useful or undesirable impinging solar radiation.

Proposed inflatable Dome can cover a big area (town) and create terrestrial living conditions on Mars (Fig.23.4a). In future, the “Evergreen” dome can cover the whole planet (Fig.23.4b).

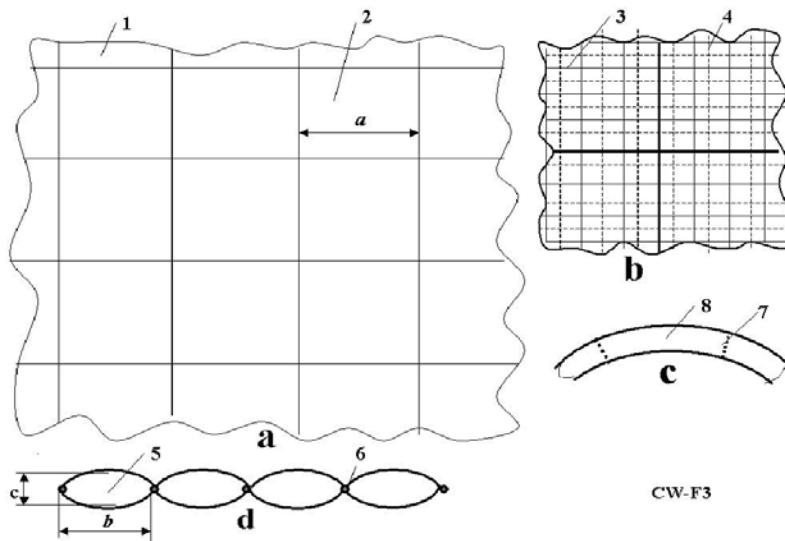


Fig. 23.3 Design of “Evergreen” cover. Notations: (a) Big fragment of cover; (b) Small fragment of cover; (c) Cross-section of cover; (d) Longitudinal cross-section of cover; 1 - cover; 2 - mesh; 3 - small mesh; 4 - thin electric net; 5 - sell of cover; 6 - tubes; 7 - film partition (non hermetic); 8 - perpendicular cross-section area.

23.2.4.1 Location, Illumination and Protection of Mars Settlements

The Solar radiation on Mars is approximately two times less than on Earth. If we want to have natural solar light, the settlement must have a magnetic control mirror suspended at high altitude in given (stationary) place (Fig. 23.5). To construct this mirror (reflector) we may use the theory of magnetic levitation developed in Bolonkin (2007b). If this reflector has a variable focus, as in Bolonkin (2006c, p. 306,

Fig. 16.3 there), then it may operate like a solar light concentrator and be used for getting energy during "night" (Mars-time).

The second important feature of the proposed installation is defending the settlement from solar wind and cosmic radiation. It is known that the Earth's magnetic field defends life from the high-energy particles (protons) of the solar wind.

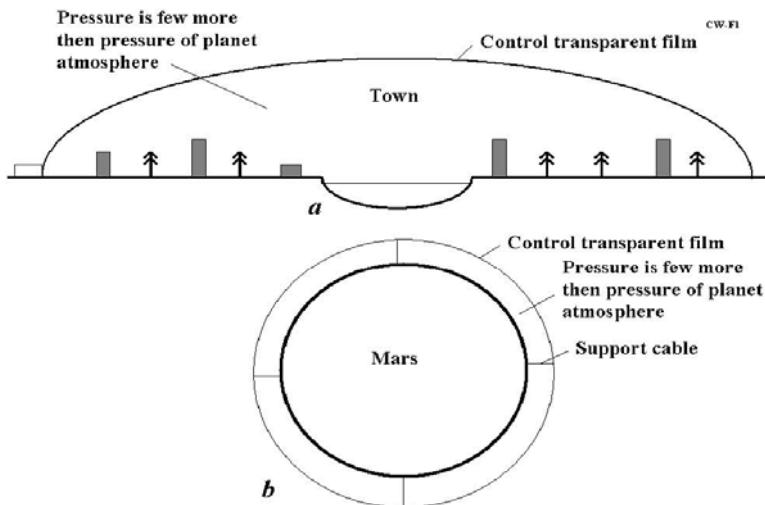


Fig. 23.4 (a) Inflatable film dome over a single town; (b) Inflatable film dome over whole planet Mars

The artificial magnetic field near Mars settlement is hundreds of times stronger than the Earth's magnetic field. It may help to defend humans from high energy particles. The pole location of this projected extraterrestrial settlement also decreases the intensity of the solar wind. Location of human settlements in polar zones or a crater also decreases the solar wind radiation. People can move to an underground dugout, a type of super radiation bunker, during periods of exceptionally high radiation activity (solar flares, coronal mass ejections).

The theory of this installation is presented below. The mass of the full reflector (rings, mirror, head screens) is about 70 - 80 kg. If a reflector is used also as powerful energy source the mass can reach 100 - 120 kg. Note, for levitating the reflector no rocket is necessary. The magnetic force increases near ground (see Eq. (23.3)). This force lifts the reflector to the needed altitude. The reflector also will be stable because it is located in the magnetic 'hole' of a more powerful ground ring magnet (it is centered in the magnetic field of it).

The artificial magnetic field may be used also for free flying of men and vehicles, as it is described in Bolonkin (2008). Mars has a lower gravity than the Earth. Then, electrostatic artificial gravity may be used to maintain a healthy constitution (Bolonkin, 2006c).

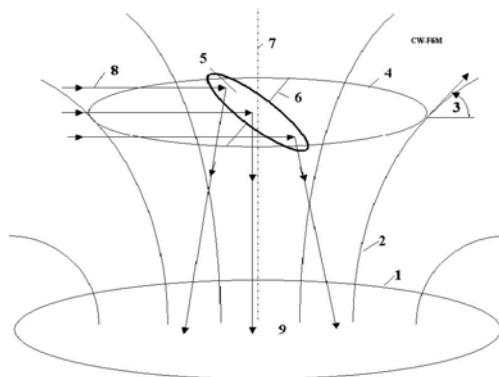


Fig. 23.5 Magnetic control mirror is suspended at high altitude over human Mars settlement. Notations: 1 -superconductivity ground ring; 2 - magnetic lines of ground superconductivity ring; 3 -angle (α) between magnetic line of the superconductivity ground ring and horizontal plate (see eq. (23.6)); top superconductivity ring for supporting the mirror (reflector) 5; 6 - axis of control reflector (which allows turning of mirror); 7 - vertical axis of the top superconductivity ring; 8 - solar light; 9.

Figure 23.6 illustrates a lightweight, possibly portable house, using essentially the same construction materials as the dwelling/workplace.

Inflatable Mars satellite hotel. The proposed inflatable Mars satellite hotel is shown in Fig. 23.7. That has the common walking area (garden) covered by a film having the controlled transparency (reflectivity), internal sections (living rooms, offices, restaurants, concert hall, storage areas, etc.). The hotel has electrostatic artificial gravity (Bolonkin 2006c) and magnetic field. The electrostatic artificial gravity creates usual terrestrial conditions while the magnetic field allows people to easily exit the hotel and still be effectively defended from the solar wind.

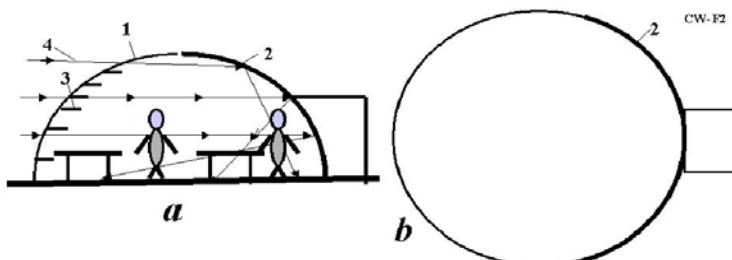


Fig. 23.6 Inflatable film house for Mars. Notation: (a) Cross-section area; (b) Top view. The other notations are same with Fig. 23.2

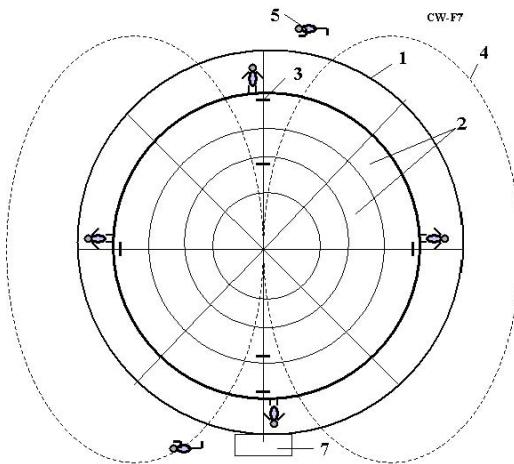


Fig. 23.7 Inflatable Mars satellite hotel. *Notations:* 1 - inflatable hotel (control transparency cover film); 2 - internal sections of hotel (living rooms, offices, café, music hall, storage, etc.); 3 - door and windows in internal sections; 4 - magnetic line; 5 - person outside the hotel (within hotel's magnetic field); 6 - common walking area (garden). 7 - dock chamber.

23.2.5 Theory of Inflatable Mars Dome

Specific mass of inflatable Dome. The mass (and relative mass) of the film is the sum of the masses of the top double layer and support cable. For a cylindrical section one has:

$$M = \frac{pS\gamma}{\sigma} H + \frac{\pi pS\gamma}{2\sigma} L \quad \text{or} \quad \bar{M} = \frac{M}{S} = \frac{p\gamma}{\sigma} \left(H + \frac{\pi}{2} L \right), \quad (23.1)$$

where M is film and cable mass, kg; p is additional air pressure, N/m²; S is cover area, m²; γ is specific mass of film and support cables, kg/m³; σ is safety tensile stress of film and the support cable, N/m²; H is height of Dome, m; L is distance between support cable, m. The needed thickness of the film, δ , is

$$\delta = \frac{\pi pL}{2\sigma} \quad (23.2)$$

Example. Let us take $p = 10^5$ N/m²; $\sigma = 10^9$ N/m² = 100 kgf/mm²; $\gamma = 1800$ kg/m³; $H = 3$ m; $L = 2$ m. Then, $\bar{M} = 1.1$ kg/m² and $\delta = 0.314$ mm.

For different values of the safety stress of film and support cables the relative mass of an inflatable local spherical section of dome may be estimated by:

$$\bar{M} = p \left(\frac{h\gamma_2}{\sigma_2} + \frac{\sqrt{2}}{4} \frac{L\gamma_1}{\sigma_1} \right), \quad (23.1')$$

where h is the length (height) of support cable, m; γ_1 and γ_2 is the specific mass of film and support cables, respectively, kg/m³; σ_1 and σ_2 is the safety tensile stress

of film and support cable, respectively, N/m^2 . Usage of Eq. (23.1') for $L = 2 \text{ m}$, $h = 2$, $\sigma_2 = 10^9 \text{ N/m}^2$, $\sigma_1 = 0.5 \times 10^9 \text{ N/m}^2$, $\gamma_1 = \gamma_2 = 1800 \text{ kg/m}^3$ and $p = 10^5 \text{ N/m}^2$ gives $\bar{M} = 0.6 \text{ kg/m}^2$. We must add about 3.6 kg/m^2 to take into account the air specific mass. The total specific mass is about 4.2 kg/m^2 . If the overpressure is only $p = 0.001 \div 0.01 \text{ atm}$, the specific mass of the Dome, including the air (or other gas) is only $0.0042 \div 0.042 \text{ kg/m}^2$.

The specific mass of Dome film for small pressure is very small, i.e. about $0.6 \div 6 \text{ g/m}^2$ or $6 \div 60 \text{ kg/hectare}$ ($1 \text{ hectare} = 10^4 \text{ m}^2$). Results of computations using Eq. (23.1') are presented in Fig. 23.8.

Magnetic stationary solar space reflector. The magnetic intensity from ground ring is given by:

$$B \approx \mu_0 \frac{iS}{2\pi H^3}, \quad S = \pi R^2, \quad (23.3)$$

where B is magnetic intensity, T; $\mu_0 = 4\pi \times 10^{-7}$ is magnetic constant; i is electric currency, A; S is the area of ground ring, m^2 ; R is ground ring radius, m; $H \gg R$ is the altitude of reflector, m. For example, using $R = 1000 \text{ m}$, $H = 1000 \text{ m}$ and $i = 10^5 \text{ A}$ one finds a magnetic intensity $B = 6.3 \times 10^{-5} \text{ T}$. The mass of superconductivity electric wire is

$$M_R = 2\pi R s \gamma_w, \quad (23.4)$$

where s is cross-section area of wire, m^2 and γ_w is specific mass of wire, kg/m^3 .

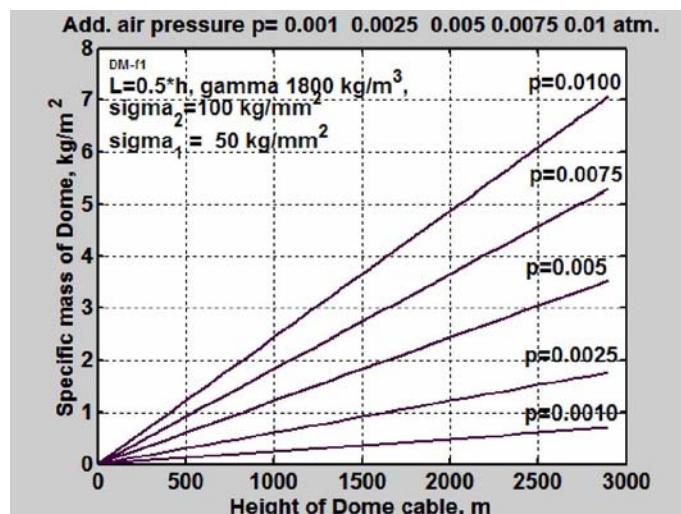


Fig. 23.8 Specific mass of the Dome having spherical sections, as a function of air (gas) pressure for distance of the support cable $L = 0.5 h$, specific mass of cable and film $\gamma_1 = \gamma_2 = 1800 \text{ kg/m}^3$, and safety stress $\sigma_1 = 100 \text{ kg/mm}^2$ (cable) and $\sigma_2 = 50 \text{ kg/mm}^2$ (film).

For density of electric currency $j = 10^5$ A/mm² and $\gamma_w = 8000$ kg/m³, the mass density of ground wire is about 50 kg. The mass of thin film heat screens which defend the wire from solar and Mars heat radiation is about 20 kg (Bolonkin 2007b). The mass of solar thin film reflector is

$$m_r = k_1 \pi r^2 \delta_r \gamma_r, \quad (23.5)$$

where r is reflector radius, m; δ_r is thickness of reflector film, m; γ_r is mass density of reflector, k_1 is coefficient of reflector mass increasing from additional support parts (for example, from inflatable ring). For example, using $r = 20$ m, $\delta_r = 5$ μ, $\gamma_r = 1800$ kg/m³ and $k_1 = 1.2$ one finds a reflector mass of 13.6 kg. The mass of top ring is

$$m = \frac{m_r}{(B j_r \cos \alpha / g_m \gamma_w) - k_2}, \quad (23.6)$$

where j_r is the density of electric currency, A/m²; g_m is Mars gravity, m/s² ($g_m = 3.69$ m/s²); α is angle between magnetic line and planet surface (Fig. 23.8); $k_2 > 1$ is coefficient of top ring mass increasing from heat radiation screens. The mass of top ring is small (less 0.5 kg).

The energy emitted by a body may be computed by using the Stefan-Boltzmann law:

$$E = \varepsilon \sigma_s T^4 \quad [\text{W/m}^2], \quad (23.7)$$

where ε is the coefficient of body blackness ($\varepsilon = 0.03 \div 0.99$ for real bodies), $\sigma_s = 5.67 \times 10^{-8}$ is Stefan-Boltzmann constant, T is temperature, K.

The daily average solar energy at Earth's orbit is calculated by:

$$Q = 86400 c q t, \quad (23.8)$$

where c is daily average coefficient, $c \approx 0.5$; t is relative daily light time, $86400 = 24 \times 60 \times 60$ is the number of seconds in an Earth day while $q = 1400$ W/m² is a rough value for the solar constant at Earth's orbit (1367 W/m²). To obtain the daily average solar energy at Mars's orbit one should multiply Eq. (23.8) by 0.43.

The overall (convection and conduction) heat loss flux q (W/m²) of the dome film cover is (Bolonkin 2007c):

$$q = k(t_1 - t_2), \quad \text{where} \quad k = \frac{1}{1/\alpha_1 + \sum_i \delta_i / \lambda_i + 1/\alpha_2} \quad (23.9)$$

Here k is the heat transfer coefficient, $t_{1,2}$ are temperatures of the initial and final multi-layers of the heat insulators, respectively, $\alpha_{1,2}$ are convection coefficients of the initial and final multi-layers of heat insulators ($\alpha = 30 \div 100$), δ_i are thickness of insulator layers; λ_i are coefficients of heat transfer of insulator layers (see Table 23.1), $t_{1,2}$ are temperatures of initial and final layers, respectively, °C. The radiation heat flux q (W/m²) for the service area is given by:

$$q = C_r \left[\left(\frac{T_1}{100} \right)^4 - \left(\frac{T_2}{100} \right)^4 \right], \text{ where } C_r = \frac{c_s}{1/\varepsilon_1 + 1/\varepsilon_2 - 1}, c_s = 5.67. \quad (23.10)$$

Here C_r is the general radiation coefficient, ε 's are blackbody emissivities of the plates (see Table 23.2) and T is temperatures of the plates, K. The *radiation* flux across a set of heat screens is computed by

$$q = 0.5 \frac{C'_r}{C_r} q_r, \quad (23.11)$$

where C'_r is given by Eq. (23.10), used between plate and reflector. Data for some materials is given in Tables 23.1, 23.2. The air layer (thickness δ) is the best thermal insulator (Table 23.1). Also, the shiny aluminum louver coating is excellent for offsetting radiation losses from the dome (Table 23.2).

Table 23.1 Heat Transferring and heat capacity (Nacshekin 1969, p 331)

Material	Density (kg/m ³)	Thermal conductivity λ (W/m°C)	Heat capacity (kJ/kg°C)
Concrete	2300	1.279	1.13
Baked brick	1800	0.758	0.879
Ice	920	2.25	2.26
Snow	560	0.465	2.09
Glass	2500	0.744	0.67
Steel	7900	45	0.461
Air	1.225	0.0244	1

Table 23.2 Blackness coefficient (Nacshekin 1969, p 465)

Material	Blackness coefficient	Temperature (°C)
Bright Aluminum	0.04 - 0.06	50 ÷ 500
Baked brick	0.88 - 0.93	20
Glass	0.91 - 0.94	20 ÷ 100

The total radiation heat flux Q may be computed by using the above equations. Equations (23.7) – (23.11) allow computation of the heat balance as well as a comparison between the incoming heat (gain) and the outgoing heat (loss). The heat from combusted fuel is given by:

$$Q = c_t m / \eta, \quad (23.12)$$

where c_t is heat rate of fuel [J/kg]; ($c_t = 40$ MJ/kg for liquid oil fuel); m is fuel mass, kg; η is efficiency of heater, $\eta = 0.5 - 0.8$.

The thickness of the dome envelope, its sheltering shell of film, is computed by using the following relations, derived from the equation of tensile strength:

$$\delta_1 = \frac{Rp}{2\sigma}, \quad \delta_2 = \frac{Rp}{\sigma} \quad (23.13)$$

where δ_1 is the film thickness for a spherical dome, m; δ_2 is the film thickness for a cylindrical dome, m; R is radius of dome (or dome section, or between support cables), m; p is additional pressure into the dome, N/m²; σ is safety tensile stress of film, N/m². For example, let us compute the film thickness for a cylindrical dome having radius $R = 100$ m, additional air pressure $p = 0.01$ atm at top section in Fig. 23.7 ($p = 1000$ N/m²), safety tensile stress $\sigma = 50$ kg/mm² ($\sigma = 5 \times 10^8$ N/m²). From Eq. (23.13) one finds

$$\delta = (100 \times 1000) / (5 \times 10^8) = 0.0002 \text{ m} \quad (23.14)$$

The dynamic pressure p_w from wind on Mars is:

$$p_w = \frac{\rho V^2}{2} \quad (23.15)$$

where ρ is atmospheric density, kg/m³; V is wind speed, m/s.

Increasing the heat protection reduces the heat losses and we can utilize inflated dome covers with more layers and more heat screens. One heat screen decreases the heat losses by a factor of 2, two screens can decrease the heat flux three times, three screens - 4 times, and so on. If the inflatable domes have a multi-layer structure, the heat transfer decreases proportional to the total thickness of its enveloping film layers.

23.2.6 Projects for Mars

The dome shelter innovations outlined here can be used for many practical applications. We suggest initial projects could be small (10 m diameter) houses (Fig. 23.6) followed by an “Evergreen” dome covering a land area 200×1000 m, with irrigated vegetation, homes, open-air swimming pools, playground, concert hall.

The house and “Evergreen” dome have several innovations: magnetic suspended Sun reflector, double transparent insulating film, controllable jalousies coated with reflective aluminum (or film with control transparency) and an electronic cable mesh inherent to the film for dome safety/integrity monitoring purposes. By undertaking to construct a half-sphere house, we can acquire experience in such constructions and explore more complex constructions. A 10 m diameter home has a useful floor area of 78.5 m^2 , airy interior volume of 262 m^3 covered by an envelope with an exterior (cover) area of 157 m^2 (design of Fig. 23.5). Its film enclosure material would have a thickness about 0.32 mm and a total mass of about 100 kg.

A city-enclosing “Evergreen” dome of 200×1000 m (Fig. 23.9 with spherical end caps) have the following characteristics: useful area = $2.3 \times 10^5 \text{ m}^2$, useful

volume $17.8 \times 10^6 \text{ m}^3$, exterior dome area of $3.75 \times 10^5 \text{ m}^2$ consisting of a film of 0.0003 m thickness and weighting about 250 tons. If the “Evergreen” dome were formed with concrete 0.25 m thick, the mass of the city-size envelope would be 200×10^3 tons, which is a thousand times heavier. Also, just for comparison, if we made a gigantic “Evergreen” dome with stiff glass, thousands of tons of steel, glass would be necessary and such materials would be very costly to transport from Earth on Mars, where they would be assembled. Our non-woven textile (film) is flexible and plastic can be relatively cheap. The single greatest boon to “Evergreen” dome construction is the protected cultivation of plants within a dome that generates energy from the available and technically harnessed sunlight.

23.2.7 Discussion

As with any innovative project proposal, the reader will naturally have many questions. We give here brief answers for two possible questions.

1. *Cover damage.* The envelope contains a rip-stopping cable mesh so that the film cannot be seriously damaged. Its section, double layering structure governs the escape of air inside the living realm. Electronic signals alert supervising personnel of all rupture problems and permit a fast repair by well-trained responsive emergency personnel. The cover consists of a double film.
2. *What is the expected lifetime of the film covering?* Depending on the materials used, it may be as much as a decade. In all or in part, the cover can be replaced periodically.

23.3 AB Method of Closed-Loop Water Cycle Agriculture

23.3.1 Brief Description of Problem

This section is based on Bolonkin (2007e). The Mars’ population will need food. It is very expensive—in fact, essentially economically impossible to deliver all required food from Earth. We will need to get a foothold at Mars, and that implies agricultural self-sufficiency. But the Mars’ surface is presently unsuitable for agriculture.

Irrigation is the artificial application of water to the soil, usually for assisting in growing crops. In crop production it is mainly used to replace missing rainfall in periods of drought, but also to protect plants against frost. Additionally irrigation helps to suppress weed growing in rice fields. In contrast, agriculture that relies only on direct rainfall is sometimes referred to as dry-land farming or as rain fed farming. It is often studied together with drainage, which is the natural or artificial removal of surface and sub-surface water from a given area. Various types of irrigation techniques differ in how the water obtained from the source is distributed within the field. In general, the goal is to supply the entire field uniformly with water, so that each plant has the amount of water it needs, neither too much nor too little. Irrigation gives high stability harvests which are ~ 3 – 5 times more than conventional agriculture would tend to provide. On Earth, irrigation is performed

only on 16% of all active agricultural area but it produces as much of the agriculture production as all non-irrigated cultivation.

A single person uses from 10 - 50 liters per day (the last value refers to the industrial countries and includes the watering of house garden). Growing plants on any scale requires a huge amount of water. For example, one hectare of wheat requests 2000 kL, cabbage - 8000 kL, a forest 12,000 - 15,000 kL per summer.

In natural conditions, soil acts as a mineral nutrient reservoir but the soil itself is not essential to plant growth. When the mineral nutrients in the soil dissolve in water, plant roots are able to absorb them. When the required mineral nutrients are introduced into a plant's water supply artificially, soil is no longer required for the plant to thrive. Hydroponics is a method of growing plants using mineral nutrient solutions instead of soil. Terrestrial plants may be grown with their roots in the mineral nutrient solution only or in an inert medium, such as perlite, gravel or Rockwool. Almost any terrestrial plant will grow with hydroponics, but some will do better than others. It is also very easy to do; the activity is often undertaken by very young children with such plants as watercress. Hydroponics is also a standard technique in biology research and teaching and a popular hobby. Hydroponics also saves an incredible amount of water; it uses as little as 1/20 the amount as a regular farm to produce the same amount of food. The water table can be impacted by the water use and run-off of chemicals from farms, but hydroponics may minimize impact as well as having the advantage that water use and water returns are easier to measure. The environment in a hydroponics greenhouse is tightly controlled for maximum efficiency and this new mindset is called Soil-less/Controlled Environment Agriculture (S/CEA). Growers monitor the temperature, humidity, and pH level constantly.

Hydroponics: advantages and disadvantages

- While removing soil-grown crops from the ground effectively kills them, hydroponically grown crops such as lettuce can be packaged and sold while still alive, greatly increasing the length of freshness once purchased.
- Solution culture hydroponics does not require disposal of a solid medium or sterilization and reuse of a solid medium.
- Solution culture hydroponics allows greater control over the root zone environment than soil culture.
- Over- and under-watering is prevented.
- Hydroponics is often the best crop production method in remote areas that lack suitable soil, such as colonies on Mars.
- In solution culture hydroponics, plant roots can be seen.
- Soil borne diseases are virtually eliminated.
- Weeds are virtually eliminated.
- Fewer pesticides may be required because of the above two reasons.
- Edible crops are not contaminated with soil.
- Water use can be substantially less than with outdoor irrigation of soil-grown crops.

- Hydroponics cost 20% less than other ways for growing strawberries. Many hydroponic systems give the plants more nutrition while at the same time using less energy and space.
- Hydroponics allow for easier fertilization as it is possible to use an automatic timer to fertilize the plants.
- It provides the plant with balanced nutrition because the essential nutrients are dissolved into the water-soluble nutrient solution.
- If timers or electric pumps fail or the system clogs or springs a leak, plants can die very quickly in many kinds of hydroponic systems.
- Hydroponics usually requires a greater technical knowledge than geponics.
- For the previous two reasons and the fact that most hydroponic crops are grown in greenhouses or controlled environment agriculture, hydroponic crops are usually more expensive than soil-grown crops.
- Solution culture hydroponics requires that the plants be supported because the roots have no anchorage without a solid medium.
- The plants will die if not frequently monitored while soil plants do not require such close attention.
- Hydroponics has been widely misconceived as miraculous. There are many widely held misconceptions regarding hydroponics, as noted by the following facts:
 1. Hydroponics will not always produce greater crop yields than with good quality soil.
 2. Hydroponic plants cannot always be spaced closer together than soil-grown crops (geponics) under the same environmental conditions.
 3. Hydroponic produce will not necessarily be more nutritious or better tasting than geponics.
 4. Hydroponics will grow 30% faster and cost less. They are also proven to be healthier and more productive.

With pest problems reduced, and nutrients constantly fed to the roots, productivity in hydroponics is high, plant growth being limited by the low levels of carbon dioxide in the atmosphere, or limited light. To increase yield further, some sealed greenhouses inject carbon dioxide into their environment to help growth (CO_2 enrichment), or add lights to lengthen the day, control vegetative growth etc.

This technology allows for growing where no one has grown before and will allow humanity to live on Mars.

23.3.2 Description and Innovations

Our project mainly consists of a closed dome covering a surface area on Mars by a thin film with controlled heat conductivity and (optionally) controlled clarity (reflectivity, albedo, carrying capacity of solar spectrum). The film is located at an altitude of $\sim (5 - 300 \text{ m})$. The film is support at this altitude by air pressure. The

film is connected to Mars' ground by cables. The cover may require double-layer film. We can control the heat conductivity of the dome cover by pumping an air (or other gas) between two layers of the dome cover and change the solar heating (solar radiation) by control of cover clarity. That allows selecting for different conditions (solar heating) in the covered area and by pumping air into dome. Envisioned is a cheap film having liquid crystal and conducting layers. The clarity is controlled by electric voltage. These layers, by selective control, can pass or blockade the solar light (or parts of solar spectrum) and pass or blockade the Mars radiation. The incoming and outgoing radiations have different wavelengths. That makes control of them separately feasible and therefore possible to manage the heating or cooling of the Mars surface under this film. In open atmosphere, the Sun heats the ground; the ground must heat the whole atmosphere before stable temperature rises are achieved. In our case the ground heats *only* the air into the dome (as in a hotbed). We have a literally greenhouse effect.

The building of a film dome is very easy. We spread out the film over Mars surface, turn on the pumping and the film is raised by air overpressure to the needed altitude limited by the support cables. Damage to the film is not a major problem because its double layer covers and protects it from loss of large amounts of air.

The first main innovation of the proposed dome (and main difference from a conventional hotbed, or greenhouse) is the inflatable *high* span of the closed cover (up to 300 m). The high height of the enclosed volume aids organizing of a *Closed Loop Water Cycle* - accepting of water vaporized by plants and returning this water in the nighttime when the air temperature decreases. That allows us to perform irrigation in the large portion of Mars land area that does not have enough water for agriculture. We can convert the Mars arid soil into gardens without expensive delivery of remote freshwater. Good soil is not necessity-- hydroponics allows us to achieve record harvests on any soil.

The second important innovation is using a cheap controlled heat conductivity, double-layer cover (controlled clarity is optionally needed for some regions). This innovation allows conserving solar heat and controlling temperature.

The third innovation is the control of the cover height, which allows adapting to local climatic seasons.

The fourth innovation is the usage of a cheap thin film for the high altitude cover. This innovation decreases needed construction mass and the construction cost by thousands of times in comparison with the conventional very expensive glass-concrete domes proposed for Earth's city use.

Our design of the dome is presented in Fig.23.9 that includes the thin inflated film dome. The innovations are listed here: (1) the construction is air-inflatable; (2) each dome is fabricated with very thin, transparent film (thickness is 0.3 to 1 mm) having controlled clarity and controlled heat conductivity without rigid supports; (3) the enclosing film has two conductivity layers plus a liquid crystal layer between them which changes its clarity, color and reflectivity under an electric voltage (Fig. 23.10); (4) the bound section of dome has a hemispheric form (#5, Fig.23.9). The air pressure is larger in these sections and they protect the central sections from the effects of the outer Martian wind.

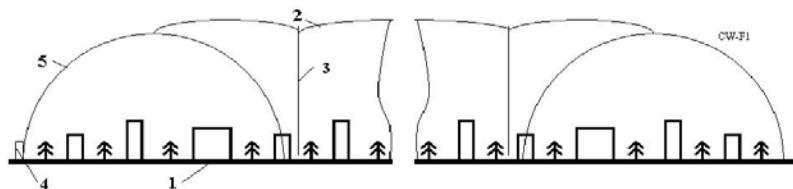


Fig. 23.9 Film dome over Martian agricultural area. *Notations:* 1 - area, 2 –double thin film cover with a control heat conductivity and clarity, 3 – control support, 4 - exits, 5 - semi-cylindrical border section

Figure 23.10 illustrates the thin transparent control dome we envisage. The inflated textile shell—technical “textiles” can be woven or non-woven (films)—embodies the innovations listed: (1) the film is thin, approximately 0.3 to 1 mm. A film this thin has never before been used in a major building; (2) the film has two strong nets, with a mesh of about 0.1×0.1 m and $a = 1 \times 1$ m, the threads are about 1 mm for a small mesh and about 2 mm for a big mesh. The net prevents the watertight and airtight film covering from being damaged by vibration; (3) the film incorporates a tiny electrically conductive wire net with a mesh about 0.1×0.1 m and a line width of about 100 μm and a thickness near 10 μm . The wire net is electric (voltage) control conductor. It can inform the dome maintenance engineers concerning the place and size of film damage (tears, rips, etc.); (4) the film may be twin-layered with the gap — $c = 1$ m and $b = 2$ m—between film layers for heat insulation. This multi-layered covering is the main means for heat isolation and puncture of one of the layers won’t cause a loss of shape because the second film layer is unaffected by holing; (5) the airspace in the dome’s covering can be partitioned, either hermetically or not; and (6) part of the covering can have a very thin shiny aluminum coating that is about 1 μm to collect additional solar radiation or reflect it in some particular cases.

Brief information about the cover film. The dome cover (film) has five layers (Fig. 23.10c): transparent dielectric layer, conducting layer (about 1 - 3 μm), liquid crystal layer (about 10 - 100 μm), conducting layer (for example, SnO_2), and transparent dielectric layer. Common thickness is 0.3 - 1 mm. Control voltage is 5 - 10 V. This film may be produced by industry relatively cheaply. The conventional control clarity (transparency) film reflects superfluous energy back to space if too much. If film has solar cells it may convert part of the superfluous solar energy into electricity.

Transparency. Transparency is the material property of allowing radiation to pass through. The degree of transparency varies according to the wavelength of the light. Only the vacuum is totally transparent, any other material has a certain absorption for electromagnetic waves. There are some materials that change their color (or other optical property) by the application of an electric charge, a technology known as electrochromics. As the color change is persistent and energy need only be applied to effect a change, electrochromic materials are used

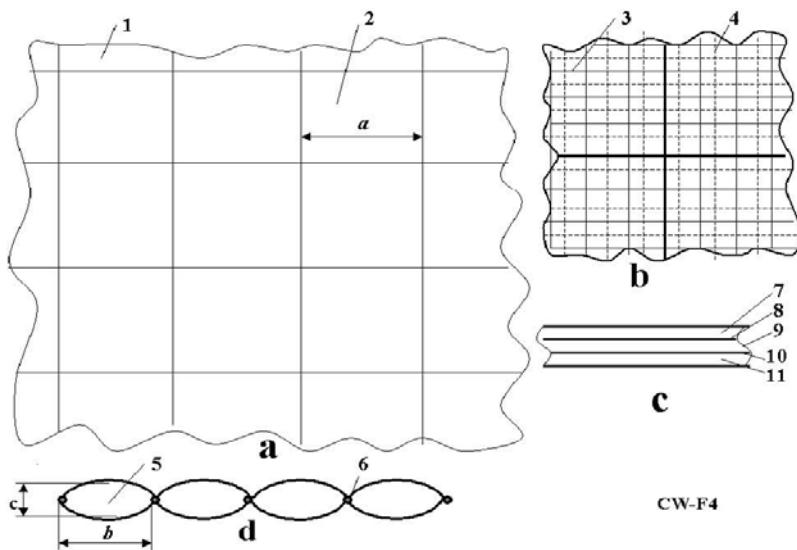


Fig. 23.10 Design of covering membrane. *Notations:* (a) Large fragment of cover with control clarity (reflectivity, carrying capacity) and heat conductivity; (b) Small fragment of cover; (c) Cross-section of cover (film) having five layers; (d) Longitudinal cross-section of cover; 1 - cover; 2 - mesh; 3 - small mesh; 4 - thin electric net; 5 - cell of cover; 6 - tubes; 7 - transparent dielectric layer, 8 - conducting layer, 9 - liquid crystal layer, 10 - conducting layer, and 11 - transparent dielectric layer.

to control the amount of radiation allowed to pass through windows ("smart windows"). One good example of an electrochromic material is polyaniline which can be formed either by the electrochemical or chemical oxidation of aniline. If an electrode is immersed in hydrochloric acid which contains a small concentration of aniline, then a film of polyaniline can be grown on the electrode. Depending on the redox state, polyaniline can either be pale yellow or dark green/black. Other electrochromic materials include tungsten oxide (WO_3), which is the main chemical used in the production of electrochromic windows or smart windows.

23.3.3 Theory of AB Dome

As the Martian wind flows over and around a fully exposed, nearly completely sealed inflated dome, the weather affecting the external film on the windward side must endure positive air pressures as the wind stagnates. Simultaneously, low atmospheric fluid pressure eddies will be present on the leeward side of the dome. In other words, external pressure gradients caused by the atmospheric fluid density differences on different parts of the dome's envelope is characterized as the "buoyancy effect". The buoyancy effect will be greatest during the coldest weather when the dome is heated and the temperature difference between its interior and exterior are greatest. In extremely cold conditions, such as on Mars, the buoyancy effect tends to dominate dome pressurization.

The theory presented in this section is derived using Nacshekin (1969). The solar constant at the level of Mars orbit is approximately 590 W/m^2 . The Mars atmosphere is very thin and (when dust storms are not activated) allows most of the incident solar radiation to reach the ground. This is enough for normal plant growth in equatorial regions.

The visible part of the Sun's wavelength spectrum is $\lambda = 400 - 800 \text{ nm}$ (0.4 to $0.8 \mu\text{m}$). Any warm body emits radiation. The emission wavelength depends on body's temperature. The wavelength of the maximum intensity is governed by Wien's law:

$$\lambda_m = \frac{2.9}{T}, \quad [\text{mm}], \quad (23.16)$$

where T is body temperature, K. For example, if a body has the temperature 20°C ($T = 293 \text{ K}$), the wavelength of maximum intensity is $\lambda_m = 9.9 \mu\text{m}$. The energy emitted by a body may be computed by using Stefan-Boltzmann law, Eq. (23.7).

The maximum solar energy flux on Mars surface is given by the approximate relationship:

$$q = q_o \cos(\varphi \pm \theta) [\text{W/m}^2], \quad (23.17)$$

where $\varphi < 90^\circ$ is the latitude of a location on Mars surface, $\theta \sim 25^\circ$ is the angle between the projection of Mars polar axis to a plane perpendicular on the ecliptic plane and containing the line Sun-Mars and the perpendicular on the ecliptic plane. The sign "+" in Eq. (23.17) signifies Summer and the "-" signifies Winter. Also, q_o in Eq. (23.17) can reach $300 - 600 \text{ W/m}^2$; during the dust storm its magnitude decreases down to 100 W/m^2 (for details see Chap. 2).

The heat flux per unit surface area of dome film cover by convection and conduction is given by Eq. (23.9). The radiation heat flux per unit surface of service area is given by Eqs. (23.10). The radiation flux across a set of the heat reflector plates is computed by Eq. (23.11). Data for some construction materials is given in Table 23.3.

Table 23.3 Materials optical properties (Nacshekin 1969, p. 465)

Material	Temperature, T $^\circ\text{C}$	Emittance, ϵ
Bright aluminum	50 \div 500	0.04 - 0.06
Bright copper	20 \div 350	0.02
Steel	50	0.56
Asbestos board	20	0.96
Glass	20 \div 100	0.91 - 0.94
Baked brick	20	0.88 - 0.93
Wood	20	0.8 - 0.9
Black vanish	40 \div 100	0.96 - 0.98
Tin	20	0.28

The relationships listed above allow computing the heat balance and comparing the incoming heat (gain) and the outgoing heat (loss). The thickness of the dome envelope, its sheltering shell of film, is computed by Eq. (23.13). The dynamic pressure from Martian wind is computed by Eq. (23.15). For example, we assume a wind speed $V = 20$ m/s and a Martian atmospheric density, which is about 100 times less than the Earth atmospheric density. Then, the dynamic pressure is $p_w = 2.45$ N/m². That is four thousands time less when internal pressure $p = 100,000$ N/m². When the need arises, sometimes the internal pressure can be voluntarily decreased, bled off.

In Bolonkin (2006a) we show the dome heat loss is about 14 times smaller than the heat loss of the buildings unprotected by an inflated dome. In this section we considered a two-layer dome film and one heat screen. If needed, better protection can further reduce the heat losses as we can utilize inflated dome covers with more layers and more heat screens.

23.3.4 Irrigation without Water: Closed-Loop Water Cycle

The relationships used in this section can be derived by using well-known physical laws (see Nacshekin 1969). Therefore, detailed explanations are not provided. Our assumption is that the gaseous content of the Mars' dome is similar to the Earth's atmosphere and additional solar radiation may be delivered by ground placed reflectors or by space mirrors orbiting around Mars.

Water in the Dome's air. The amount of water in the Dome's air depends upon temperature and humidity. For 100% relative humidity, the partial pressure of water vapor at 1 atm pressure is shown in Table 23.4.

Table 23.4 Saturation pressure p_s of water vapor in air, for various temperatures

Temperature (°C)	-10	0	10	20	30	40	50	60
Saturation Pressure (kPa)	0.287	0.611	1.22	2.33	4.27	7.33	12.3	19.9

Air pressure is 1 atm (1013 hPa).

The amount of water per cubic meter of air may be computed by:

$$m_w = 0.00625 [p(t_2)h - p(t_1)] \quad (23.18)$$

where m_w is the mass of water, kg per 1 m³ of air; $p(t)$ is vapor pressure from Table 23.4, while $h = 0 \div 1$ is relative humidity. Results obtained by using Eq. (23.18) are presented in Fig. 23.11. Typical values for the relative humidity in terrestrial atmospheric air is 0.5 - 1.

Computation of closed-loop water cycle. Assume the operating temperature is achieved during the daytime. When the dome reaches that temperature, the control system may fill with air the space 5 (Fig. 23.5) between the double-layer of the film cover. The control system may also decrease the solar radiation input, increasing reflectivity of the liquid crystal layer of the film cover. That way, we can keep a constant temperature inside the dome.

The heating of the dome in the daytime may be computed by:

$$\begin{aligned} q(t) &= q_0 \sin(\pi t / t_d), \quad dQ = q(t)dt, \quad Q = \int_0^{t_d} dQ, \quad Q(0) = 0, \\ M_w &= \int_0^{t_d} adT, \\ dT &= \frac{dQ}{C_{p1}\rho_1\delta_1 + C_{p2}\rho_2H + rHa}, \\ a &= 10^{-5}(5.28T + 2), \quad T = \int_0^{t_d} dT, \quad T(0) = T_{\min}, \end{aligned} \quad (23.19)$$

where q is heat flux density, W/m^2 ; q_0 is maximum Sun irradiance on Mars during the day time, $q_0 \approx 100 \div 600, \text{W/m}^2$; t is time, s; t_d is daily (Sun) time, s; Q is heat, J; T is temperature in dome (air, soil), $^{\circ}\text{C}$; C_{pl} is heat capacity of soil, $C_{pl} \approx 1000 \text{ J/(kgK)}$; $C_{p2} \approx 1000 \text{ J/(kgK)}$ is heat capacity of air; $\delta_1 \approx 0.1 \text{ m}$ is thickness of heated soil; $\rho_1 \approx 1000 \text{ kg/m}^3$ is an approximate value for soil density; $\rho_2 \approx 1.225 \text{ kg/m}^3$ is density of the air; H is thickness of air (height of cover), $H \approx 5 \div 300 \text{ m}$; $r = 2,260,000 \text{ J/kg}$ is evaporation heat, a is coefficient of evaporation; M_w is mass of evaporation water, kg/m^3 ; T_{\min} is minimum temperature inside the dome after night, $^{\circ}\text{C}$.

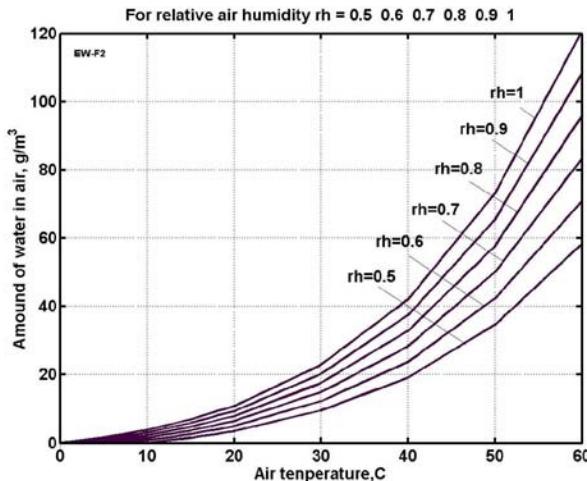


Fig. 23.11 Amount of water per cubic meter of air versus air temperature and relative humidity (rh). $t_l = 0^{\circ}\text{C}$

The convective (conductive) cooling of dome during night time may be computed as follows:

$$q_t = k(T_{\min} - T(t)), \quad \text{where} \quad k = \frac{1}{1/\alpha_1 + \sum_i \delta_i / \lambda_i + 1/\alpha_2}, \quad (23.20)$$

where q_t is the heat flux through the dome cover by convective heat transfer, W/m^2 ; see the other notation in Eq. (23.9). We take $\delta = 0$ during night time (through active control of the film).

The radiation heat flux q_r (from dome to night sky, radiation cooling) may be estimated by Eqs. (23.10):

$$q_r = C_r \left[\left(\frac{T_{\min}}{100} \right)^4 - \left(\frac{T(t)}{100} \right)^4 \right], \quad [\text{W}/(\text{m}^2 \text{K}^4)] \text{ where} \quad (23.21)$$

$$C_r = \frac{c_s}{1/\varepsilon_1 + 1/\varepsilon_2 - 1}, \quad c_s = 5.67$$

For various notations see Eq. (23.10). We take $\varepsilon = 1$ in night time (through the active control of the film). Other relationships similar to Eqs. (23.19) are also used:

$$dQ = [q_t(t) + q_r(t)]dt, \quad Q = \int_0^{t_d} dQ, \quad Q(0) = 0, \quad M_w = \int_0^{t_d} adT, \quad (23.22)$$

$$dT = \frac{dQ}{C_{p1}\rho_1\delta_1 + C_{p2}\rho_2H + rHa},$$

$$a = 10^{-5}(5.28T + 2), \quad T = \int_0^{t_d} dT, \quad T(0) = T_{\min},$$

As an example, let us use the following values: $H = 135 \text{ m}$, $\alpha = 70$, distance $\delta = 1 \text{ m}$ between cover layers (see #5 in Fig. 23.2), $\lambda = 0.0244$ for air. Results of computation are presented in Figs. 23.12 and 23.13. For a dome cover height $H = 135 \text{ m}$

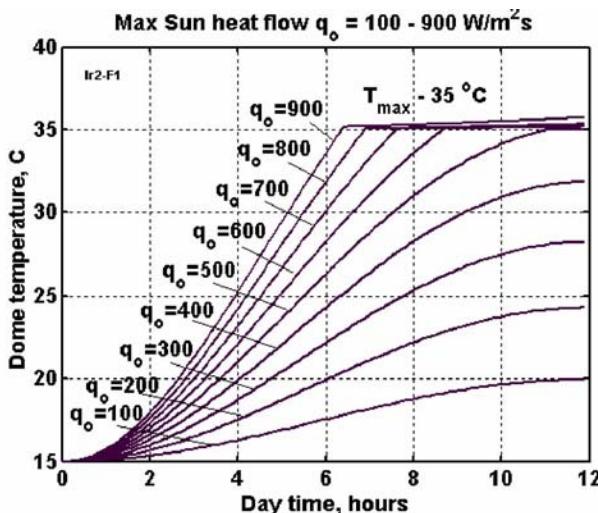


Fig. 23.12 Heating of the dome by solar radiation from the night temperature of $15 \text{ }^{\circ}\text{C}$ to $35 \text{ }^{\circ}\text{C}$ for various day time durations and daily maximum solar irradiance (W/m^2). Height of dome film cover equals $H = 135 \text{ m}$. The control system limits the maximum internal dome temperature to $35 \text{ }^{\circ}\text{C}$.

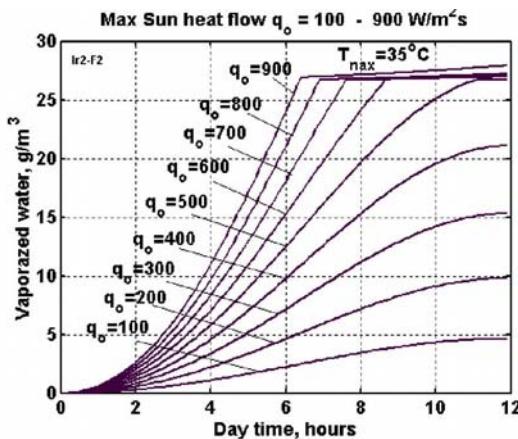


Fig. 23.13 Water vaporization for 100% air relative humidity for different values of maximum solar radiance (W/m^2) and various day time durations. Height of dome film cover equals $H = 135 \text{ m}$. The temperature control system limits the maximum internal dome temperature to 35°C .

the night precipitation is $0.027 \times 135 = 3.67 \text{ kg}$ or 3.67 mm/day (Fig. 23.13). If this quantity is not enough, we can increase the height of dome cover.

As one can see, the same operating temperature may be obtained in a wide range of latitudes, during both summer and winter time. That means that the climate of the covered areas does not depend on their location upon the Mars' surface (up to latitude 20° - 30°), nor on the Mars' seasons or on the outside weather. Note that the present Dome design is not optimal.

Table 23.5 Material properties (Bolonkin 2007a)

Material	Tensile strength (kg/mm ²)	Density (g/cm ³)		Tensile strength (daN/mm ²)	Density (g/cm ³)
Whiskers		Fibers			
AlB ₁₂	2650	2.6	QC-8805	620	1.95
B	2500	2.3	TM9	600	1.79
B ₄ C	2800	2.5	Allien 1	580	1.56
TiB ₂	3370	4.5	Allien 2	300	0.97
SiC	1380-4140	3.22	Kevlar or Twaron	362	1.44
Materials		Dynecta or Spectra			
Steel strands	186	7.8	Vectran	283-334	0.97
Steel Piano wire	220-248	7.8	E-Glass	347	2.57
Steel A514	76	7.8	S-Glass	471	2.48
Aluminum alloy	45.5	2.7	Basalt fiber	484	2.7
Titanium alloy	90	4.51	Carbon fiber	565	1.75
Polypropylene	2-8	0.91	Carbon nanotubes	6200	1.34

23.3.5 Discussion

As with any innovative project proposal, the reader will naturally have many questions. We give here brief answers for three possible questions.

1. *How can dust be removed from the dome?* By outer wind or dedicated crawling vacuum cleaner.
2. *Storm wind.* The storm wind can only be on the bounding (outside) sections of dome. They are special semi-cylindrical form (Fig. 23.6) and stronger than central sections.
3. *Cover damage.* The envelope contains a rip-stop cable mesh so that the film cannot be damaged greatly. Electronic signals alert supervising personnel of any rupture problems. The needed part of cover may be reeled down by control cable and repaired. Dome has independent sections.

23.4 Conclusions

“Evergreen” domes can foster the fuller economic development of Mars - thus, increasing the effective area of territory dominated by humans on three celestial bodies. Normal human health can be maintained by ingestion of locally grown fresh vegetables and healthful “outdoor” exercise.

The Mars environment is not suitable for unshielded human life. The artificial agriculture areas, crop capacity, carrying capacity addable to humanity’s domain by means of converting Mars into gardens are a way to open an important escape hatch from some of humanity’s most pressing problems. The proposed cheapest AB method may dramatically increase the potentially realizable sown area, crop capacity; indeed the range of territory suitable for human living. In theory, converting some regions of Mars into prosperous garden would be the equivalent of colonizing an entire new planet. The suggested method is relatively inexpensive. We can start from small areas and extended the practice over a large area.

Let us summarize some advantages of using the AB Dome on Mars:

- No need for large amounts of water for irrigation;
- Low cost of inflatable film Dome per area reclaimed;
- Control of inside temperature;
- Covered area is not at risk from dust storms and changing weather;
- Possibility of flourishing crops even with a sterile soil (hydroponics);
- 4 – 6 harvests in one Martian year;
- Covering entire Martian colonies by the proposed domes.

Acknowledgements

The author thanks Joseph Friedlander (Israel) for language editing and useful technical advices.

References

- Bolonkin, A.A.: Control of Regional and Global Weather (2006a), <http://arxiv.org>
- Bolonkin, A.A.: Cheap Textile Dam Protection of Seaport Cities against Hurricane Storm Surge Waves, Tsunamis, and Other Weather-Related Floods (2006b), <http://arxiv.org>
- Bolonkin, A.A.: Non-Rocket Space Launch and Flight. Elsevier, London (2006c)
- Bolonkin, A.A.: New Concepts, Ideas, and Innovations in Aerospace, Technology and Human Life. Nova Science, New York (2007a)
- Bolonkin, A.A.: AB Levitation and Energy Storage. Paper AIAA-2007-4613 at the 38th AIAA Plasma Dynamics and Lasers Conference in conjunction with the 16th International Conference on MHD Energy Conversion on June 25-27, 2007, Miami, USA (2007b)
- Bolonkin, A.A.: AB Method of Irrigation without Water (Closed-loop water cycle) (2007c), <http://arxiv.org>
- Bolonkin, A.A.: Inflatable Dome for Moon, Mars, Asteroids and Satellites. Paper AIAA-2007-6262 presented at AIAA Conference Space-2007, September 18-20, 2007, Long Beach, CA, USA (2007b)
- Bolonkin, A.A.: Cheap artificial AB-Mountains, Extraction of Water and Energy from Atmosphere and Change of Country Climate (2007e), <http://arxiv.org>
- Bolonkin, A.A., Cathcart: A Cable Space Transportation System at the Earth's Poles to Support Exploitation of the Moon. J. British Interplanet Soc. 59, 375–380 (2006a)
- Bolonkin, A.A., Cathcart, R.B.: Inflatable 'Evergreen' Polar Zone Dome (EPZD) Settlements (2006b), <http://arxiv.org>
- Bolonkin, A.A., Cathcart, R.B.: The Java-Sumatra Aerial Mega-Tramway (2006c), <http://arxiv.org>
- Bolonkin, A.A., Cathcart, R.B.: Inflatable 'Evergreen' Dome Settlements for Earth's Polar Regions. Clean Technol. Environmental Policy 9, 125–132 (2006d)
- Naschekin VV, Technical thermodynamic and heat transmission (in Russian). Publishing House for High Universities (Vishshiaia Shkola), Moscow (1969)

Chapter 24

Self-sustaining Martian Colonies Utilizing the North Polar Ice Cap and the Atmosphere

James Powell, George Maise, Jesse Powell, and John Paniagua

Plus Ultra Technologies, Inc., Stony Brook, New York

In Xanadu did Kubla Khan
A stately pleasure dome decree
Where Alph the sacred river ran
Through caverns measureless to man
Down to a sunless sea...
It was a miracle of rare device.
A sunny pleasure dome with caves of ice!
Samuel Taylor Coleridge

24.1 Introduction

The history of Polar exploration on Earth provides two important lessons for future manned missions to Mars:

1. Use the best technology available
2. Live off the land

Sadly, British Polar expeditions tended to ignore both rules. Scott's South Pole expedition violated Rule #1. They man hauled their sleds because they considered it more noble technology than using dogs. When their supplies ran out, they froze to death. Meanwhile, Amundsen's party raced to the Pole by dog sled, throwing away surplus food on their way back. Sir John Franklin's expedition to find the Northwest Passage perished when they violated Rule #2. They lived on the tinned food they had brought from England, ignoring the nearby healthy Eskimos who lived off the land. Eventually they starved to death helped along by lead poisoning from their tinned containers.

Following Rule #1, the manned Mars missions should use high performance Nuclear Thermal Propulsion (NTP) for the journey to Mars and back, instead of chemical rockets. The trip times will be much shorter, reducing cosmic radiation dose to the crew, as well as physical and physiological stress.

Even more important is Rule #2, "Live off the land". The critical material needed for living off the land on Mars is water. There has been an intense search

for water on Mars. But the question of water on Mars is really a non-question. There are many cubic kilometers of water ice available in the North Polar Cap, with wide areas where the ice is exposed at the surface. The assumption appears to be that the dry areas on Mars are the only suitable ones for exploration and bases – that exploration of, and bases on, the Polar Caps would be of little value.

In fact, the reverse is true. The North Polar Cap, with readily available H₂O ice and the ability to easily carve out comfortable shielded habitats and storage chambers is the best site for manned bases and eventually, colonies. Not only can the North Polar Cap serve as a site for robust bases and a colony, but it is of great scientific value by itself. The multi-kilometer thick ice sheet is a repository for millions of years of data on Martian meteorology (atmospheric composition, dust loading, temperature, isotope ratios, etc.), climate, solar and cosmic ray activity, meteorite bombardment, and so on.

Moreover, the ice sheet is a very promising place to look for past life forms, both as fossils formed from the ancient ocean that covered much of the Northern Hemisphere of Mars, as well as microfossils carried by the wind from other Martian regions and accumulated in the ice. The ice sheet would also accumulate traces of biologic chemicals carried by the wind.

As described in the following sections, using the water ice, Martian atmosphere, and minerals from the accumulated dust loadings in the North Polar Cap, virtually all of the propellants, fuels, life support materials, including food and construction materials for manned bases on Mars – and eventually colonies – can be manufactured there, without having to bring them from Earth. Going even beyond self-sufficiency, colonists on Mars could manufacture large quantities of supplies to be transported back to and stockpiled in depots in high Earth orbit, taking advantage of the much lower speed needed to lift off from the Martian surface. The stockpiled materials could sustain space hotels for Earth tourists, zero gravity manufacturing facilities, supplies for lunar bases, exploration missions to the Outer Solar System, and a robust Earth-Mars transport infrastructure.

With the North Polar Cap as a permanent base, astronauts could explore all of Mars in detail using robotic and manned flyers and rovers, constructed with and powered by the materials manufactured on Mars. Robotic flyers could lay strings of depots with food, water, fuels, and breathable air for the explorers, enabling them to make long journeys safely and comfortably.

To help ensure this capability, a compact robotic factory concept is proposed that would land on the North Polar Cap two years before the first manned mission took off from Earth. In that period, the robotic factory unit would manufacture and stockpile a very large and ample supply of materials for the astronauts stay on Mars and their return journey, before they left Earth. This concept, termed ALPH, maximizes mission capabilities and minimizes mission risk, and is described in the following sections.

24.2 The ALPH Concept

ALPH is a compact nuclear powered robotic factory unit that would land on the North Polar Cap in advance of the first manned mission to Mars (Powell et al. 1998, 2000). The ALPH unit would consist of a very small, light water cooled

reactor and a set of small chemical manufacturing units. The total mass of the ALPH unit, including reactor, process units, controls, and the lander structure is ~5 metric tons. The ALPH reactor would produce warm water by melting the surrounding H₂O ice, plus electrical power from its small steam turbine [the thermal energy for melting H₂O ice would be the waste heat from the Standard Rankine Steam cycle]. Using the thermal energy and electric power generated by the ALPH reactor, the set of chemical manufacturing units on the landed spacecraft would carry out the following functions:

1. Extract carbon dioxide and nitrogen from the Martian atmosphere.
2. Electrolyze melt water to produce hydrogen and oxygen gas.
3. Liquefy and store H₂ and O₂ for use as propellants.
4. Mix O₂ with N₂ (from 1) and store for use as breathable air.
5. React H₂ with CO₂ to make methane and methanol, and store for future use as fuels.
6. React H₂, CO₂ and N₂ to form plastics and store for future use as construction materials.
7. Grow algae and yeast using water, methanol, and light emitting diodes, and store for future use as food for humans and fish farming.

Figure 24.1 outlines the development stages of a Mars exploration and colonization program based on the use of ALPH. In the first stage, robotic units like the recent Mars rovers would continue to explore Mars. However, specific exploration of potential landing sites on the North Polar Cap would be undertaken to determine ice composition and internal structure, dust content, surface smoothness, etc. In the second stage, two or more ALPH units would land on the Polar Cap and begin to manufacture and stockpile supplies and materials to maximize the probability of mission success. It would appear desirable to have at least 2 ALPH units land at different points on the Polar Cap. The manufacturing and stockpile activity would be continuously monitored from Earth by video and communication links to ensure proper quantity and quality of the supplies and materials, and that they were properly stockpiled to be available when the astronauts landed.

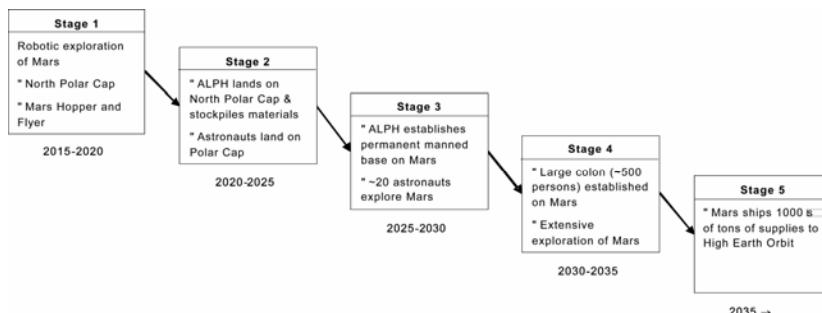


Fig. 24.1 Stages in the exploration and colonization of Mars based on the ALPH factory concept

After the first manned landing, there could be additional manned missions to the same site to continue the exploration there, or there could be landings at other sites established by ALPH units. At some point, probably after several individual missions, the decision would be made to set up a permanent manned base, leading to the third stage of the exploration of Mars. The number of astronauts at the permanent base could be as little as 5 to 10, and perhaps as great as 20 to 30. While a single ALPH unit could still supply even the larger base, it would be desirable to have 2 or 3 operating ALPH units to ensure reliability and robustness of supply.

The fourth stage would be an actual colony on Mars, termed Xanadu after Coleridge's poem, "Kubla Khan", of perhaps as many as 500 persons. Colonists could reside permanently on Mars or spend several years there and then return to Earth.

Whether a Mars colony will actually occur will depend on many factors, including the willingness of colonists to leave Earth permanently or for years, whether there is economic and practical stability on Earth, and whether Earth is willing to support a colony on Mars. To a great extent that will depend on what is discovered on Mars, and the value of the products brought back from Mars.

In the fifth stage, it appears practical using ALPH type systems to return thousands of tons per year of propellants, fuels, breathable air, plastics and other construction materials, plus food – all manufactured on Mars – to high Earth orbit for a large scale space infrastructure, at much lower cost than if the materials were launched from Earth.

24.2.1 Stage 1: Robotic Exploration of the Polar Cap

Before sending manned missions to Mars, it will be very desirable to gather much more data about the planet to help ensure mission safely, and to optimize the exploration programs for the astronauts, so that their efforts will yield the maximum scientific return. In particular, since the ALPH base will be on the North Polar Cap, robotic exploration units should investigate the interior structure, in detail, determining the degree and kind of dust loading, how it varies with depth at potential landing sites, etc.

Stage 1 robotic exploration of the North Polar Cap and the other areas of Mars would be carried out using two kinds of robotic explorers:

- MICE (Mars Ice Cap Explorer)
- The Mars Hopper

MICE would explore the North Polar Cap, while the Mars Hopper would explore other areas of Mars. The North Polar Cap is very attractive for scientific research. MICE can obtain detailed data at many locations inside the ice sheet on the composition and meteorology of the Martian atmosphere over millions of years, on the geology of the wind borne dust collected by the ice cap, on ancient solar wind and cosmic ray activity, and possibly, evidence of life on Mars, through microfossils, DNA traces, etc. plus detailed information on the internal structure of the ice cap, and of possible landing sites. There is a relatively small region of pure

H_2O ice in the cap, estimated to be 837 km^2 in area and at least 1 km thick (Keiffer and Zent 1992) which is surrounded by hundreds of thousands of square kilometers of layered polar terrain. Eroded channels in this region show alternating layers of ice and dust. The ice layers are 10 to 30 meters thick, separated by thinner layers of an ice-dust composite. Quantitative data on the dust content in this portion of the Ice Cap is lacking. Carr (Carr 1966) assumes for that the average dust/ice content is 50%/50%. However, visual indications suggest that the dust content is substantially less than 50%; also, the more northerly regions probably have a smaller dust content, since most of the dust would be trapped at lower latitudes. Having a dusty ice cap rather than pure ice is not serious, as long as it did not hinder access or processing. Dust contents as high as 20 to 30% by volume should not be a problem, since melt channels and cavities could still be created inside the ice sheet. The dust would be filtered or centrifuged out, and the purified water then used to make propellants, food, oxygen, etc. The presence of dust would actually be beneficial, since it would be a feed material for making aluminum, iron, and other metals, as well as ceramics based on silica, magnesia, etc. This would eliminate long trips to collect feed material from lower latitudes. Compact, ultra lightweight mobile robotic probes would gather detailed data on the internal structure of the cap that would be transmitted back to Earth in real time, except for the delay due to the finite speed of light, which could be a maximum of ~1 hour. Scientists on Earth could control the path and speed of the MICE probe. The MICE probe (Fig. 24.2) would use a compact, ultra lightweight nuclear reactor as a source of heat and electric power, to melt its way down through the ice, collecting data on the dust content and its properties, as well as a wide range of data on the composition of the ancient Martian atmosphere, solar and cosmic ray activity, and possible evidence of Martian life forms, including wind borne microfossils and

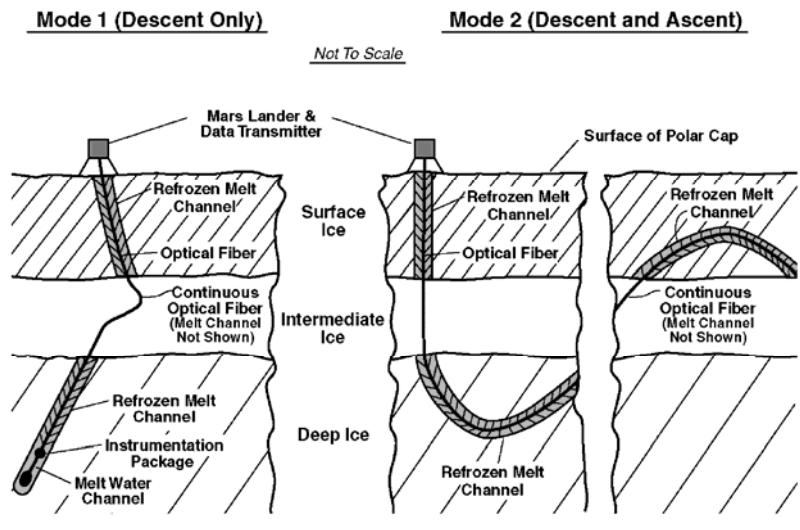
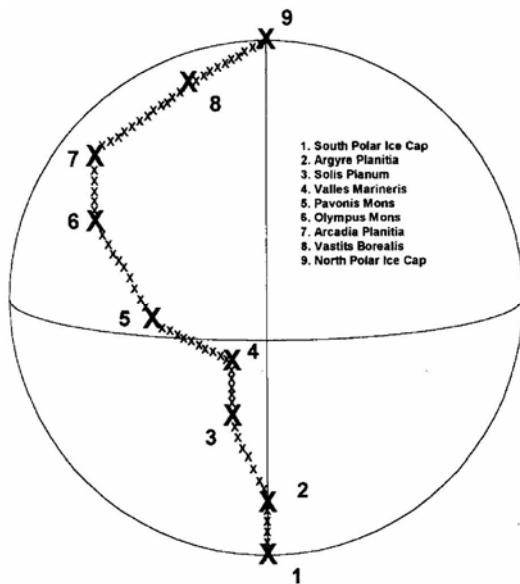


Fig. 24.2 MICE operational modes

DNA traces. Details of the MICE concept are described elsewhere (Powell et al. 1999). MICE probes can explore the ice cap vertically and horizontally, descending to depths of kilometers, and travel horizontally for distances considerably greater than a kilometer.

In addition to the MICE probe, the mobile Mars Hopper robotic probe could explore many, e.g., dozens, of widely separated sites on Mars during its mission (Fig. 24.3). During its 18 month stay on Mars, the Hopper (Powell et al. 2001) would explore almost a hundred sites. The Hopper would use CO₂ propellant obtained from the Martian atmosphere and a compact ultra lightweight nuclear thermal propulsion engine to carry out rocket powered hops of 200 to 300 kilometers between sites. At each site the Hopper would obtain data on local surface and subsurface geology and potential life forms, collect samples, take pictures, and so on. The Hopper could land in alluvial type valleys, canyons, volcanoes, impact craters, etc. Since the Hopper would not travel on the surface, it could land at sites not accessible to wheeled rovers. The Hopper would also land at several sites on Mars Polar Caps, using a warm water heated probe to collect data about the subsurface structure of the Polar Cap, the composition and properties of the trapped ancient Martian atmosphere, and fossil traces of ancient life. With only a few Mars Hoppers, virtually the entire planet could be explored in great detail, providing data to help answer questions about whether life now exists on Mars, or ever existed in the past, whether and where ancient oceans, lakes, and rivers flowed, how the Martian atmosphere has changed over time, and many other unknowns.

Fig. 24.3 Sites to be explored on Mars. Note: This is a partial list of the 70 sites along the flight path trajectory shown above



The enormous scientific results obtained from the MICE probes and the Mars Hopper missions will provide the basis for proceeding to Stage 2 of the Xanadu concept. They will enable a decision on the best location for the colony site on the

North Polar Cap, what conditions will be encountered when the robotic factory is landed, the properties of the interior structure of the ice sheet - important for the formation of the habitats and storage cavities what amounts and kinds of dust loads to be expected in processing the ice and so on. Moreover, the amount of data obtained on the ancient meteorology and geology of Mars, as well as, the issue of whether life ever existed on Mars, will be immense.

24.2.2 Stages 2 and 3: Deployment of the Robotic ALPH Factory and Permanent Manned Base on Mars

After landing on the polar ice sheet, the spacecraft would lower the ALPH reactor, which would then turn on and start to melt a channel through which it would sink into the ice. At a depth of approximately 10 meters it would cease its descent, thereafter to remain in place as it continued to operate. The melt channel above it would refreeze to ice, leaving a small melt pool around the reactor. The ice above the reactor would provide complete shielding from all of the neutron and gamma radiation that it emitted. Astronauts would be able to walk on the surface above it, and to live in ice cavities in the ice sheet without receiving any appreciable radiation dose.

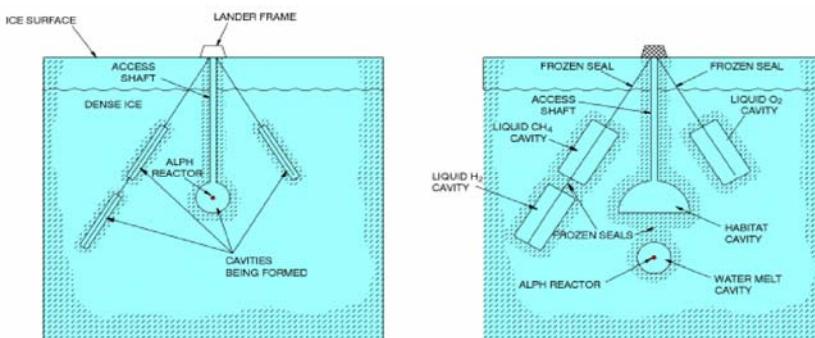


Fig. 24.4 ALPH angled branch configuration

In addition to deploying the ALPH reactor, the spacecraft would also deploy horizontal flexible lines that were connected to the reactor coolant circuit. Using warm water from the reactor, the lines would sink down into the ice, as illustrated in Fig. 24.4. After reaching the desired depth, warm water would then be admitted into an expandable-water filled balloons (Fig. 24.5) attached to the lines. The warm water balloon would then create cavities for stockpiling the supplies produced by ALPH, and for habitats for the astronauts after their landing.

Figure 24.5 shows two possible methods for melting a cavity inside the ice sheet. In both approaches, a flexible, thin wall, collapsed plastic balloon surrounds the warm water melt line at an appropriate point along its length. When the warm water line has melted its way down to the desired depth, additional warm pressurized water is pumped into the collapsed balloon, causing it to expand against, and

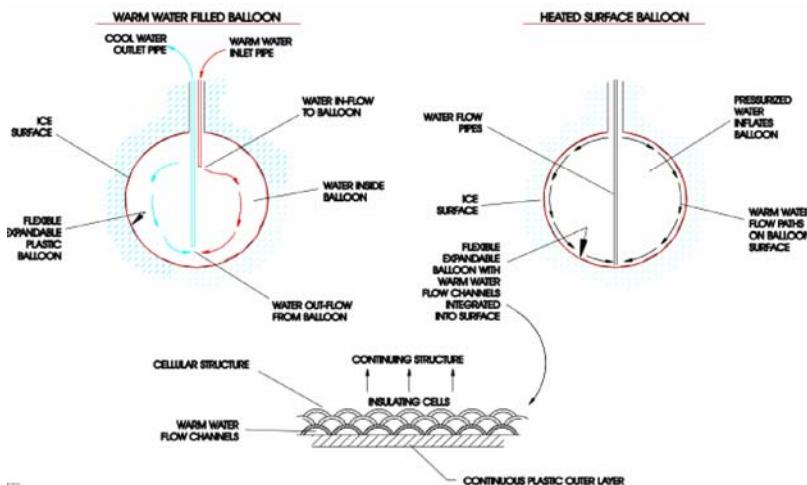


Fig. 24.5 Warm inflated balloon design types

start to melt, the surrounding ice. In the warm water filled balloon approach (left picture in Fig. 24.5), the inlet water circulates throughout the whole balloon, while in the heated surface balloon approach (right side of Fig. 24.5), the warm water only circulates through the network of flow channels on the surface of the balloon. The interior of the balloon is filled with pressurized water, which keeps the balloon inflated, and its hot surface pressed against the surrounding ice. The heated skin approach has two advantages. First, it reduces the volume of warm water required; second, after the cavity has been created, the water in the network of surface channels is flushed out and replaced with a thermally insulating gas blanket. The gas would be air, N₂, CO₂, or H₂, depending on the nature and temperature of the cavity contents. In type 1 cavities, materials produced by ALPH would be stored at a lower temperature than the surrounding ice sheet - for example, liquid H₂, liquid O₂, liquid air, and liquid methane. In type 2 cavities, materials would be stored at a higher temperature than the surrounding ice - for example, water, methanol, plastics, and food. The habitats for the colonists would also be kept at a higher temperature than the surrounding ice. In the first type, the heat leak into the cavity from the ice is minimized by the cellular gas blanket in order to keep the power demand on the cryogenic refrigerators as small as possible. In the second type, the heat leak out of the cavity is kept small, in order to minimize space heating requirements, and to prevent local melting of the surrounding ice.

Large cavities can be created in a short time using the waste heat from the ALPH nuclear reactor. For example, creating a cylindrical cavity 10 meters in diameter and 15 meters long would take only 25 days and require about 500 KW(th) of heat. This corresponds to ~12% of ALPH's reject heat from the steam cycle, so that a total of 8 such cavities could be created in a little more than 3 weeks. After formation, the cavities can be sealed using a freeze seal technique, in which the entrance to the cavity is closed off with an ice plug created by injection of water and subsequent freezing by the surrounding ice.

The ALPH reactor generates electrical and thermal power which is fed back to the process units on the spacecraft lander, to make the supplies that will be stock-piled in the ice cavities to be utilized when the astronauts land.

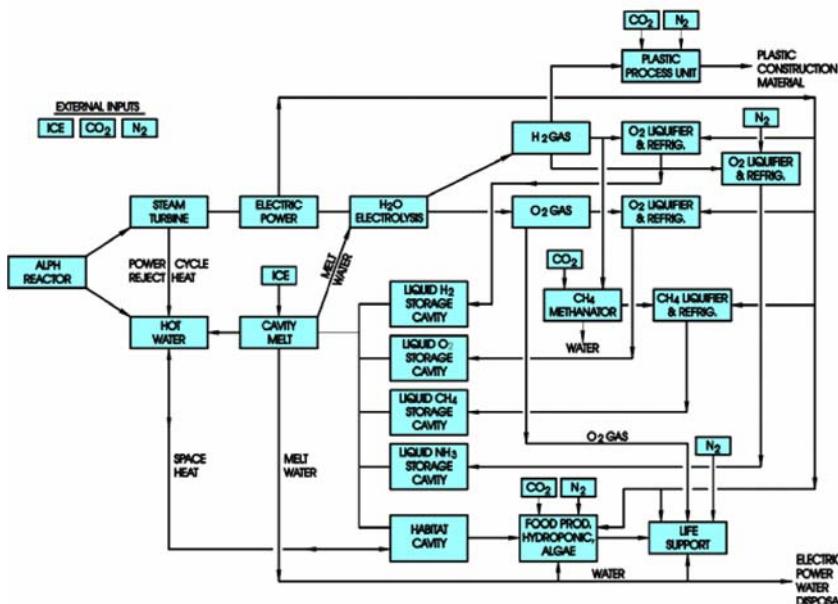


Fig. 24.6 ALPH production flowsheet

Figure 24.6 shows the ALPH production flowsheet. All of the process units utilize existing technology, with water, CO₂ and N₂ as the basic feed materials. The water comes from melting of the local ice, while CO₂ and N₂ came from the Martian atmosphere. The separation into CO₂ and N₂ gas streams can be done using either a simple temperature swing absorption (TSA) process with a fixed bed of molecular sieve beads, or by a compressor/liquefaction cycle. The water electrolysis process would use a high performance solid polymer electrolyte (SPE) with an electrical efficiency in the range of 80 to 90%. The O₂ would be liquefied for future use in propulsion engines, fuel cells, etc., or mixed with N₂ and stored as liquid air for future use as breathable air. Hydrogen and carbon dioxide would be reacted to form methane, methanol, and plastics. Using methane or methanol as substrates for growth of yeasts and bacteria in fermentation vessels would yield food supplies that contain protein, fats, and carbohydrates. Ethanol can also be synthesized for use as a substrate for food production. Table 24.1 lists the principal parameters of the ALPH factory, together with an inventory of the hundreds of tons of supplies that it would produce and stockpile in the 2-year period before the first colonists landed. The total weight of the factory is approximately 5 metric tons, which is tiny compared to the amount of supplies that it would produce.

Table 24.1 First ALPH System Parameters

Time interval between ALPH landing and colonists landing (months)	20
Reactor thermal power (MW(e))	5.3
Power cycle type	Steam turbine (20% eff.)
Output electric power per unit (kW(e))	1070
Outlet steam conditions (deg C, Atm)	250, 37
Stockpiled supplies after 20 months of operation, metric tons	
Liquid hydrogen	160
Liquid oxygen	1680
Liquid methane	60
Methanol	30
Plastics	30
Food	10
Number of storage units	8
Diameter of storage units (m)	9
Height of sub-surface habitat for colonists (m)	5

Basis: ALPH unit lands at colony site to produce supplies for initial colonists

The ALPH reactor would use existing Zr/UO₂ cermet fuel, which has operated with a perfect safety record. All fission products are fully contained inside the cermet fuel with zero release to the water coolant. A burnable neutronic poison can be incorporated into the cermet fuel, allowing the reactor to operate for many years on a given fuel loading.

24.2.3 Stage 4: The Xanadu Mars Colony

In Stage 4, the Mars Colony would grow to a mature size of several hundred people. The ultimate size of the colony will depend on many unknown factors – colonists' interest and commitment to live on Mars, far away from Earth; the degree of financial support Earth is willing to spend; the value of products (materials, knowledge, movies and videos, etc.) that colonists could send back to Earth, and so on.

A nominal size of 500 persons appears to be a reasonable projection. It is large enough that the colonists could have the numerous and diverse skills needed for the detailed exploration of Mars, as well as the doctors, psychologists, teachers, plumbers, engineers, and so forth, who could sustain a good quality of life in the colony.

The life support requirements for the colonists can be readily provided by a set of six ALPH units, each of 2 MW(e) electric output – twice the output of the first ALPHs that landed to set up the initial manned bases on the North Polar Cap. Using the ALPH's thousands of tons of supplies could be produced that would sustain the 500 colonists, permit them to carry out detailed exploration of a wide range of sites all over Mars, and ship propellants and supplies back to Earth orbit to sustain the Earth-Mars transport infrastructure (Table 24.2).

Table 24.2 ALPH production rates for Xanadu colony

Total reactor thermal power	60 MW (th)
Total electric generation	12 MW (e)
Number fo ALPH reactors (10 MW[th])	6
<u>H₂ production, metric tons per year</u>	
Liquid H ₂ for Earth-Mars-Earth transport	835
H ₂ for CH ₄ and CH ₃ OH fuel	40
H ₂ for plastics	80
H ₂ for food substrates	90
H ₂ for stockpile (30%)	<u>300</u>
Total	1345
<u>O₂ production, Metric tons per year</u>	
Liquid O ₂ for Earth-Mars-Earth transport	1430
O ₂ for breathable air	80
O ₂ for combustible fuels	800
O ₂ for plastics and food	100
O ₂ for stockpile (50%)	1200
O ₂ surplus (discarded)	<u>7160</u>
Total used	3610
<u>General supplies, metric tons per year</u>	
Breathable air (80% N ₂ , 20% O ₂)	400
Water for life support	18,000
CH ₄ for exploration	15
Fuels for colony	180
Plastics	500
Food (including 50% stockpile)	270
Electric power for life support	2 MW(e)
Space heat for life support	12 MW(th)

Basis: Average yearly production rate; Steady state colony of 500 persons; 20% turnover every 2 years

While some colonists would probably choose to stay on Mars for the rest of their lives, others probably would want to return to Earth after a few years on Mars. Assuming that there would be a turnover rate of 20% of the colony every 2 years (100 persons back to Earth, and 100 new colonists to Mars), the 2-way periodic trips could be accommodated by transporting propellant and supplies that were manufactured on Mars back to depots in Earth orbit. From there the returning colonists would be transported back to Earth's surface, while outbound colonists would take off for Mars. Table 24.3 shows the propellant manufacture requirements for this 2-way infrastructure.

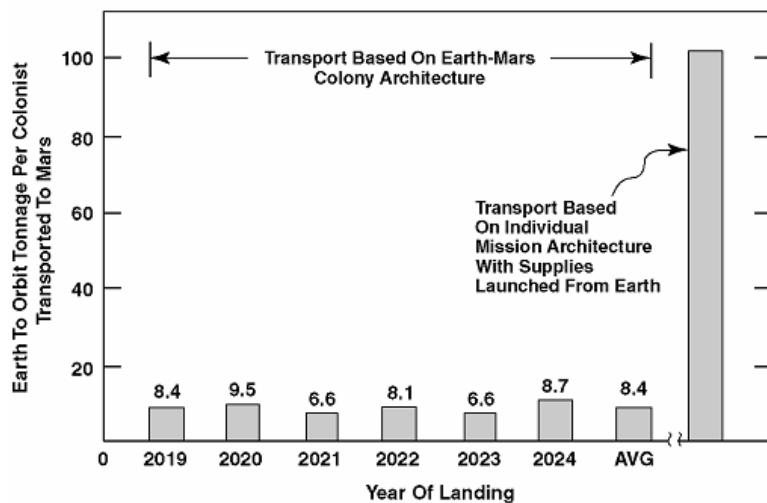
With this 2-way transport infrastructure and the manufacture of supplies on Mars by ALPH, the Earth to orbit launch mass requirements for a person on Mars can be reduced by more than a factor of 10, as compared to the requirements if all supplies were brought from Earth (Fig. 24.7). Moreover, this comparison greatly underestimates the launch mass advantage, since a colonist supported by ALPH

Table 24.3 Mars propellant flow: requirements for a steady state Mars colony

	Liquid H ₂ (Metric Tons/Yr)	Liquid O ₂ (Metric Tons/Yr)
1. Liquid H ₂ transported to GEO depot for 5 habitat trips per 2 years	175	---
2. Liquid H ₂ required by NTCPV to transport liquid H ₂ to GEO depot (2 NTCPV trips per 2 years)	200	---
3. Liquid H ₂ in Mars depot to fuel cyclers for habitat trips back to Earth	175	---
4. Fuel for H ₂ /O ₂ engines to lift habitats from colony to Mars depot (5 habitat trips per 2 years)	45	230
5. Fuel for H ₂ /O ₂ engines on fueling shuttle to lift liquid H ₂ to Mars depot	240	1200
Total produced at colony	835	1430

Basis: Mature colony of 500 persons; Turnover rate of 20% every 2 years; (100 persons to Mars, 100 persons to Earth); 5 habitat trips out every 2 years; 5 habitat trips back every 2 years; Propellant requirements normalized to an annual basis; 5/1 mixture ratio for H₂/O₂ engines

would have a much longer stay on Mars compared to the Earth only supply person – as much as a factor of 10 – and would have a much safer, much more productive, and much more comfortable stay than the Earth only supply person.

**Fig. 24.7** Earth to orbit launch mass required to transport a colonist and supplemental supplies to Mars

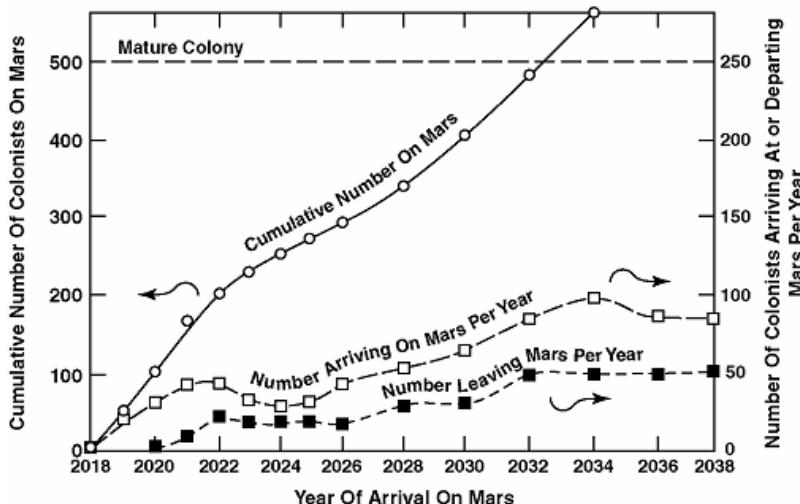


Fig. 24.8 Number of colonists on Mars as a function of time

Figure 24.8 shows an illustrative example of the buildup of the Xanadu colony with time. Assuming a probably optimistic start date of 2018, the colony would reach a mature state in approximately 15 years.

24.2.4 Stage 5: Transport of Materials from Mars to Earth Orbit

A Mars Colony that transported propellants and supplies back to Earth orbit would enable a large scale space infrastructure that could support a wide range of important activities, including hotels for space tourists, zero g manufacturing facilities, Lunar bases, major missions to the Outer Solar System, and so on. The benefits of this large scale space infrastructure would be enormous, and would more than justify the establishment of the Xanadu Colony on Mars.

The flowsheet for the transport system architecture, termed MERIT (Mars-Earth Rapid Interplanetary Transport) is illustrated in Fig. 24.9. While each ALPH factory unit would be designed to produce a number of different products, e.g., liquid H₂ and O₂, breathable air, plastic, food, etc., each of the MERIT units would be designed to produce one specific product. This reflects the much greater production level required for a large scale infrastructure, and the potential improvement in efficiency by specializing on a single product.

Each MERIT unit combines an ALPH nuclear reactor and the appropriate chemical process equipment with the capability to shuffle between the Mars surface and Mars orbit. After the MERIT shuttlecraft lands on the surface of the North Polar ice sheet, the ALPH reactor and process unit operate to fill the shuttle tank with the appropriate product. When the tank is full, the shuttle takes off to rendezvous with a waiting orbiting tanker. After unloading its cargo the MERIT shuttle returns to the Polar Cap to manufacture its next load. When the tanker is full, it takes off for Earth orbit to deliver its cargo.

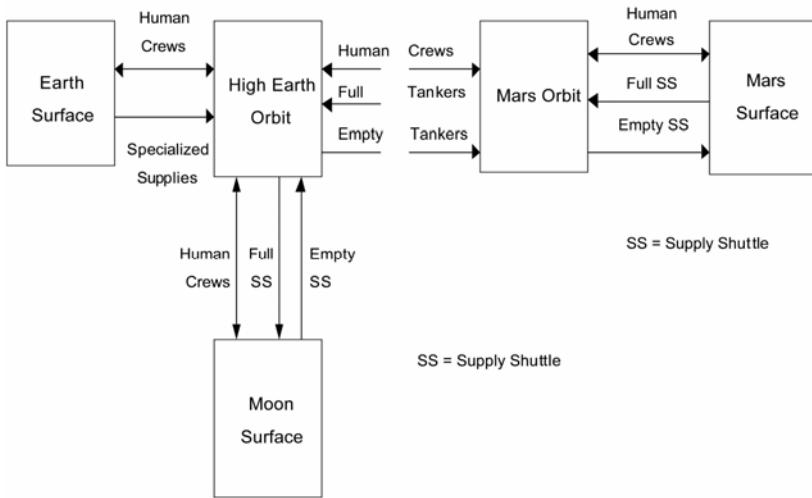


Fig. 24.9 MERIT architecture for the human exploration of Mars and the Moon

To efficiently deliver the cargo made by the MERIT units, each MERIT shuttle has a Nuclear Thermal Propulsion (NTP) engine using H₂ propellant. Accordingly, each MERIT unit must produce, in addition to its main product, enough liquid H₂ to propel the unit with its cargo, nominally 20 tons for each shuttle load. To maximize the delivery capability of the MERIT factory units, a new type of very lightweight tankage is proposed.

The amount of material flows in a MERIT architecture will depend on the number of persons working at, and visiting, facilities in Earth orbit, and the various activities. Such activities could include space manufacturing, constructing and maintaining space solar satellites for power beaming to Earth, space tourists, maintenance and source activities in support of space tourism, scientific research, transfer to bases on the moon, exploration of Mars, and other bodies, asteroid mining, etc.

Assuming an average in-orbit population of 1000 persons, an illustrative MERIT architecture can be derived. To support 1000 persons, 500 which are space tourists plus 500 being long term staff to carry out manufacturing activities, maintenance, and service, scientific research, etc., a total of approximately 3000 tons per year of supplies would be required. Table 24.4 shows the assumed breakdown for the various supplies.

The total number of MERIT units and their power level to supply 3000 tons of materials annually to Earth orbit is 14 units (Table 24.5).

A minimum of 2 units for each category is assumed, to ensure that if a factory unit failed the remaining unit could supply at least 50% of the desired supply until a replacement were landed on Mars. The power level is selected to supply the desired amount of material per unit. In practice, the units could have higher power capability without substantially increasing total mass, so that if a particular unit were to fail, the remaining unit or units could supply the full amount of material by increasing their power.

Table 24.4 Annual mass flows for an illustrative MERIT architecture

• Annual life support materials
400 tons food (1 kg per person day)
400 tons water (1 kg per person day)
200 tons makeup air (0.5 kg per person day)
• Annual propellants and construction materials
1000 tons of liquid H ₂ (for both nuclear thermal propulsion and H ₂ /O ₂ engines)
500 tons of liquid O ₂ (for H ₂ /O ₂ engine)
500 tons of plastics
• Tanker parameters
2 year round trip time
1000 tons cargo capacity for each tanker
7 tanker fleet (1 extra)
• Each tanker carries the following cargo
135 tons food, 135 tons water, 65 tons air, 365 tons of liquid H ₂ , 180 tons of liquid O ₂ , and 180 tons of plastics

Basis = 1000 total persons in orbit of which; 500 are space tourists (1 week stay, 25,000 annual visitors); 500 are long term staff (service personnel, scientists, etc.)

Table 24.5 Number and power of MERIT factory units operating on Mars North Polar Cap to deliver 3000 tons/year of cargo to high Earth orbit

Type of MERIT Production Unit	Operating Power Used	# on Mars
Liquid H ₂ (1000 tons/yr)	3 MW()	4
Liquid O ₂ (500 tons/yr)	1.2 MW(e)	2
Water (400 tons/yr)	1 MW(e)	2
Liquid Air (200 tons/yr)	0.5 MW(e)	2
Food (400 tons/yr)	2 MW(e)	2
Plastics (500 tons/yr)	2 MW(e)	2

24.3 Technology Development for ALPH and Xanadu

The technology development for the exploration and colonization of Mars depends on the stage involved. For the first stage, the robotic exploration of the North Polar Cap, the required nuclear and materials hardware technology for the MICE probe already exists. Design and testing of the MICE probe based on this known technology is necessary, but there are no “show stoppers” and no breakthroughs are required. Development of the control and communications systems for MICE is needed, but there do not appear to be any barriers to such development.

Significant new technology development will be needed for the Mars Hopper, particularly for its compact Nuclear Thermal Propulsion (NTP) engine. There has been substantial work on the US and Former Soviet Union (FSU) on NTP, including

the ground testing of the US-NERVA engine. However, the NERVA engine is much too heavy to be used for the Mars Hopper. Subsequent to the work on NERVA, the US carried out substantial R&D in the 1980's and early 90's on a much smaller NTP engine, the PBR (Particle Bed Reactor). A derivative of the PBR, termed MITEE (MIniature ReacTor EnginE) has been proposed for robotic space exploration (Powell et al. 1998a). The total weight of the MITEE engine is less than 200 kg, with a thrust of 14,000 newtons. With further development, a MITEE engine could be tested and ready for Mars missions early in the next decade.

The ALPH factory unit would have same reactor system as the MICE probe, using existing nuclear fuel and water cooled reactor technology. Similarly, the various chemical process units would use existing commercial industrial technology for water electrolysis, methanation, and the production of plastics and foods. Design of the individual sub-systems for application to the Mars mission would be necessary, together with their integration into full system. The testing of the full ALPH system could be carried out on Earth on existing ice sheets in Alaska, Greenland, Antarctica or other suitable site, using a non-nuclear electric/thermal power source. The deployment, production, and storage phases of the ALPH unit could be fully tested and demonstrated. In a parallel effort, the ALPH reactor power system would be designed and tested at an existing nuclear laboratory. The two parallel efforts could then be integrated into the final system for a Mars mission. Since the basic technology already exists for all of the sub-systems in ALPH, with the main effort being on designing for Mars conditions, and integration into a full system, ALPH could be ready for application within 10 years, by 2019.

For Xanadu, the basic ALPH technology would remain the same, with the only change being a scale-up in output capability for each unit. The nominal scale-up would be a factor of 2. A much greater development effort would be required with regard to the social, environmental, and technical structure of the Xanadu colony. What types of people went, what were the explorations to be carried out and how, what the length of stay on Mars would be, what colony size would be desired, how would colonists be transported to and from Mars, and so on.

Stage 5, the transport of large amounts of materials from Mars back to Earth orbit, would require major development of new technologies, including the shuttle vehicle to transfer supplies from the Martian surface to tankers in Earth orbit, large Earth-Mars tankers to deliver the supplies back to Earth orbit, and depots and space stations in High Earth Orbit (HEO). Such development probably would not start until after ALPH units and astronauts had successfully landed on Mars.

Acknowledgments

The authors wish to thank the NASA Institute for Advanced Concepts (NIAC) and its director Dr. Bob Cassanova for generous support of this work, under Research Grant 07600-053. The authors wish to express their deep appreciation to Ms Barbara Roland for preparation of the manuscript and to Ms Judy Otto for preparation of the drawings.

References

- Carr, M.H.: Water on Mars. Oxford University Press, Oxford (1966)
- Keiffer, H.H., Zent, A.P.: Quasi – Periodic Climate Change on Mars. In: Mars, K.H. (ed.), pp. 1180–1233. University of Arizona Press (1992)
- Powell, J., Paniagua, J., Maise, G., Ludewig, H., Todosow, M.: ALPH – A Robotic Precursor to Produce Large Amounts of Supplies for Manned Outposts on Mars, Paper IAF-98-Q-3-08. In: 49th International Astronautical Congress, Melbourne, Australia (1998)
- Powell, J., Paniagua, J., Maise, G., Ludewig, H., Todosow, M.: An Ultra Lightweight Nuclear Engine for New and Unique Planetary Missions, Paper IAF-98-R.1.01, In: 49th International Astronautical Congress, Melbourne, Australia (1998a)
- Powell, J., Maise, G., Paniagua, J.: MICE: A Compact Near-Term Mobile Robot for Exploration of the Martian Polar Ice Cap, Paper IAA-Q.3.08. In: 50th International Astronautical Congress, Amsterdam, The Netherlands (1999)
- Powell, J., Maise, G., Paniagua, J.: Development of Self-Sustaining Mars Colonies Utilizing the North Polar Cap and the Martian Atmosphere, Final Report, NIAC Research Grant 07600-053 (2000)
- Powell, J., Maise, G., Paniagua, J.: The Mars Hopper – A Mobil Lightweight Probe to Explore and Return Samples from Many Widely Separated Locations on Mars, Paper IAA-01-13.3.08, In: 52nd International Astronautical Congress, Toulouse, France (2001)
- Powell, J., Maise, G., Paniagua, J.: SUSEE: Ultra Light Nuclear Space Power Using the Steam Cycle, In: IEEE Aerospace Conference, Big Sky, Montana (2003)

Chapter 25

Mars Climate Engineering Using Orbiting Solar Reflectors

Colin R. McInnes

University of Strathclyde, Glasgow, UK

25.1 Introduction

Terrestrial geo-engineering is currently being explored as a large-scale venture to mitigate against rapid terrestrial climate change due to anthropogenic carbon emissions. A range of schemes have been proposed, including the use of orbiting solar reflectors to reduce solar insolation to compensate for increased radiative forcing of the climate (Early 1989; Angel 2006). While the scale of endeavour required to deploy geo-engineering schemes is impressive, on an even more ambitious scale the same technologies which can be envisaged to engineer the Earth's climate can be scaled to engineer the climate of Mars. Such terraforming schemes (engineering an Earth-like climate) have long been discussed, although the concept became somewhat more mainstream with the work of Sagan and others (Sagan 1961, 1973). Bioengineering schemes have been proposed, including the delivery of customized organisms to convert carbon dioxide to oxygen in the atmosphere of Mars, and darkening the Martian polar caps to reduce their albedo, again using customized organisms. Halocarbons synthesised on Mars have also been considered as a tool to quickly raise the surface temperature and so liberate trapped carbon dioxide (Gerstell et al. 2001; Badescu 2005). For other details see Chap. 26.

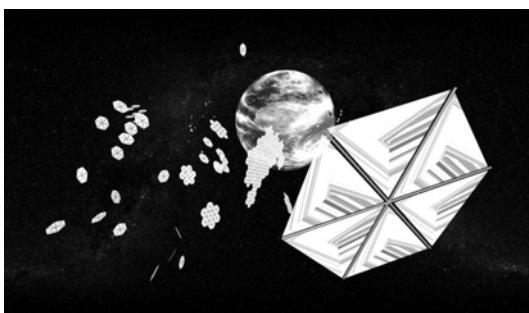
More direct terraforming schemes have been proposed for Mars, including the use of orbiting solar reflectors to increase total solar insolation. Such reflectors can be used to increase the mean surface temperature through direct radiative forcing, or can attempt to catalyze favourable climatic transitions by preferentially heating the polar caps in an attempt to release frozen carbon dioxide. Zubrin and McKay estimate that a rise in polar temperature of order 4 K may be sufficient to drive a transition to a new climate state in which the polar caps have evaporated, delivering an atmosphere with a surface pressure of 50-100 mb, as the first step of a longer term terraforming process. It is estimated that sufficient carbon dioxide exists in both the polar caps and regolith to ultimately raise the surface pressure to 300-600 mb. The solar reflectors to produce such a transition are envisaged to be in static equilibria high above the poles of the planet (Zubrin and McKay 1997), or

in displaced polar orbits (McInnes 2002). The reflector system proposed by Zubrin was to be used solely to heat the polar caps, while the much larger systems discussed by both Birch (1992) and Fogg (1992) were seen as providing an increase in total insolation, averaged over the planetary surface. These reflectors are envisaged as part of a larger terraforming effort, with the reflectors used to increase the total planetary insolation by as much as 30%.

As has been noted elsewhere (Maunter and Parks 1990), the main advantage of using large solar reflectors for planetary engineering (geo-engineering or terra-forming) is the vast energy leverage obtained in a relatively short time. The total energy intercepted by the reflector quickly grows beyond the energy required for its fabrication, leading to a highly efficient tool for climate modification. The use of reflectors with a deployed area of order 10^{12} m^2 and a mass of order 10^{10} kg at the Earth-Sun L₁ point has been explored (McInnes 2007) as a means of mitigating against anthropogenic climate change (Fig 25.1). As will be seen, the required reflector area for Mars terraforming is of order 10^{13} m^2 (and mass of order 10^{10} kg). A single large reflector is clearly impractical so that a large number of smaller reflectors are envisaged, fabricated from asteroid resources and distributed azimuthally in a uniform ring around a displaced polar orbit to provide the same total reflective area.

Families of displaced polar orbits suitable for Mars terraforming have been investigated in some detail for solar sails and a range of relatively near-term mission applications proposed (McInnes 1991, 1999; McInnes and Simmons 1992). Such displaced orbits can also be exploited for large solar reflectors, where the reflector orientation is defined *a priori* to ensure back reflection of solar radiation onto Mars, as discussed by Fogg (1995). The principal advantage of these orbits is that the reflector is never in eclipse, and it can be shown that the effective transfer of energy per orbit is greater than for circular orbits in the ecliptic plane of Mars. In addition, the strong solar radiation pressure perturbation on the reflector is managed by the out-of-plane component of gravitational force. In this chapter the orbital dynamics of displaced solar reflectors is investigated in detail, with a complete family of orbits suitable for Mars terraforming identified. The linear stability characteristics of this family of orbits are also investigated, and a stable sub-group of orbits identified with the resulting orbits locked in a Sun-synchronous state. This family of passively stable, Sun-synchronous orbits appears to offer significant benefits for large Mars solar reflectors.

Fig. 25.1 Swarm of reflectors for terrestrial geo-engineering (Dario Izzo ESA).



In addition to displaced orbits, the use of static equilibria for Mars solar reflectors proposed by Zubrin is also discussed (Zubrin and McKay 1997). At these locations on the night-side of Mars, the solar radiation pressure force exerted on the solar reflector exactly balances the local gravitational force. The efficiency of these locations for increasing planetary insolation is discussed. It has been shown that even for an ideal reflector, the reflector orientation which is required to ensure static equilibrium leads to only half of the radiation reflected being intercepted by the planetary body, while the remainder is lost. For a more realistic reflector with non-perfect optical reflectivity it is found that the problem is significantly worse (McInnes, 2000). In fact, the orientation required for a non-ideal reflector to be maintained in static equilibrium can lead to all of the reflected radiation being directed away from the planetary body. Therefore, there are fundamental difficulties associated with the use of static equilibria for Mars terraforming.

25.2 Mars Climate Engineering

As discussed earlier, solar reflectors can be envisaged as planetary engineering tools which can be used to drive the mean surface temperature of Mars upwards by direct radiative forcing, or by attempting to liberate frozen carbon dioxide from the polar caps (Oberg 1981; McKay et al. 1991; Birch 1992; Fogg 1995; Zubrin and McKay 1997). Here, it will be assumed that solar reflectors are used to increase the planetary insolation by back reflecting solar radiation, averaged across the entire planetary surface. The approximate requirements for a solar reflector to increase the total insolation of Mars can therefore be determined using a simple thermal balance model for the mean planetary surface temperature. For an incident solar flux S , the planetary black-body temperature T_B can be obtained from a heat balance such that $(1-\alpha)\pi R^2 S = 4\pi R^2 \sigma T_B^4$ so that the mean surface temperature T_M may be written as

$$T_M = \Delta T_G + \left[\frac{(1-\alpha)}{4\sigma} S \right]^{1/4} \quad (25.1)$$

where α is the Bond albedo of Mars, R is the planetary radius (3390 km) and σ is the Stefan-Boltzmann constant ($5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$). The term ΔT_G accounts for any modest increase in the mean surface temperature over the black-body temperature due to current atmospheric greenhouse warming. For an incident solar flux of 589 W m^{-2} and an albedo of 0.25, the black-body temperature is found to be 210 K. This is only slightly cooler than the mean surface temperature of 217 K, the discrepancy being due to atmospheric carbon dioxide.

In order to directly increase the mean surface temperature to that required for terraforming by radiative forcing would require a significant increase in planetary insolation. However, feedback processes are likely to be of benefit and may allow a more modest increase in solar insolation to achieve a large rise in mean surface temperature. For example, as the mean surface temperature rises due to solar heating, carbon dioxide from the polar caps may be liberated, leading to a positive

feedback as the Martian atmosphere is thickened with greenhouse gases. In a synergetic scenario for terraforming Mars an increase in solar insolation of order 30% is envisaged (Fogg 1992), which is coupled with other activities to drive the mean surface temperature upwards.

The total planetary insolation of Mars is simply $\pi R^2 S$ which corresponds to a thermal input of 2.13×10^{16} W. To increase the planetary insolation by some fraction λ , the heat balance is modified such that $(1-\alpha)\pi R^2 S + \lambda\pi R^2 S = 4\pi R^2 \sigma T_B^4$. The new mean surface temperature can therefore be written as

$$T_M = \Delta T_G + \left[\frac{(1-\alpha+\lambda)}{4\sigma} S \right]^{1/4} \quad (25.2)$$

An enhancement of total insolation by 30% requires an additional 6.38×10^{15} W to be delivered by the solar reflectors. With this additional thermal input Eq. (25.2) shows that the black-body temperature of the planet rises to 228 K, providing only a modest direct increase in mean surface temperature, although again climatic feedback is anticipated to drive the mean surface temperature upwards (Fogg 1992).

25.3 Solar Reflector Orbit Selection

25.3.1 Circular Orbit

In order to assess candidate orbits for solar reflectors, an evaluation of three possibilities will be conducted: a circular orbit in the ecliptic plane of Mars, a polar orbit normal to the ecliptic plane of Mars and static equilibria high above the poles of Mars. For a reflector on a circular orbit, with the reflector position defined by polar angle θ relative to the Sun-line, the reflector is oriented such that $\alpha = \theta/2$ in order that reflected solar radiation is directed towards the surface of Mars, as shown in Fig. 25.2. For a reflector of area A , on a circular orbit of period T , the mean energy transfer ΔE per unit time is therefore given by

$$\left\langle \frac{\Delta E}{T} \right\rangle_c = \frac{1}{T} \int_0^T S A \cos(\theta/2) dt \quad (25.3)$$

However, the orbital angular velocity $\omega = d\theta/dt$ is related to the orbit period T such that $\omega = 2\pi/T$ so that

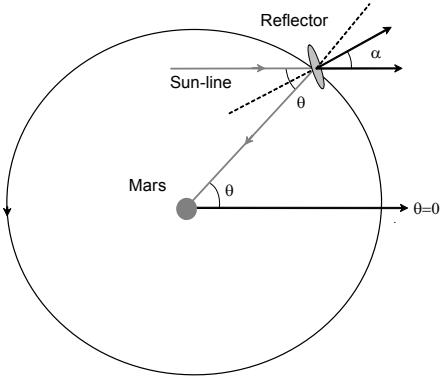
$$\left\langle \frac{\Delta E}{T} \right\rangle_c = \frac{S A}{\omega T} \int_{-\pi}^{+\pi} \cos(\theta/2) d\theta \quad (25.4)$$

Therefore, performing the integration, the mean energy transfer ΔE per unit time is given by

$$\left\langle \frac{\Delta E}{T} \right\rangle_c = \frac{2}{\pi} SA \quad (25.5)$$

It can be seen from Eq. (25.5) that approximately 64% of the maximum available solar energy SA is transferred to the surface of Mars by a reflector on a circular orbit.

Fig. 25.2 Solar reflector on a circular orbit in the ecliptic plane of Mars



25.3.2 Polar Orbit

For comparison, a circular polar orbit normal to the ecliptic plane of Mars will now be considered, as shown in Fig. 25.3. Then, the reflector orientation can be fixed as the reflector orbits normal to the Sun-line. This fixed orientation provides an additional benefit over orbits in the ecliptic plane. The mean energy transfer ΔE per unit time for a polar orbit is therefore given by

$$\left\langle \frac{\Delta E}{T} \right\rangle_p = \frac{SA}{\omega T} \int_0^{2\pi} \cos(\alpha) d\theta \quad (25.6)$$

In order that the reflected solar radiation is directed toward the surface of Mars (along the terminator) the reflector attitude α is fixed at $\alpha = \pi/4$ so that

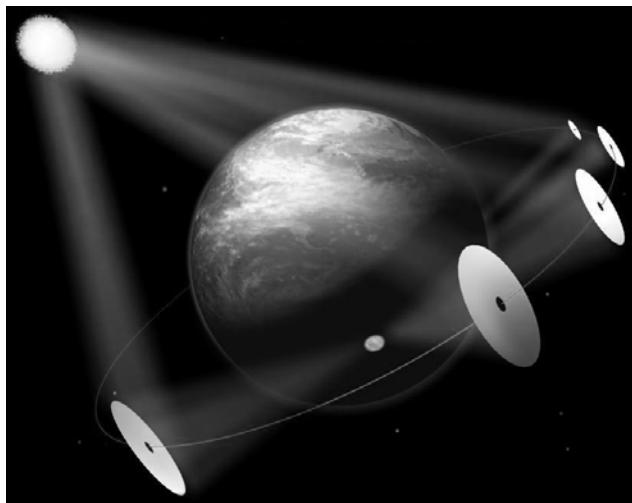
$$\left\langle \frac{\Delta E}{T} \right\rangle_p = \frac{1}{\sqrt{2}} SA \quad (25.7)$$

It can be seen that approximately 71% of the maximum available solar energy SA is transferred to the surface of Mars. Evaluating the effectiveness of the circular and polar orbits it can be seen that

$$\left\langle \frac{\Delta E}{T} \right\rangle_c < \left\langle \frac{\Delta E}{T} \right\rangle_p \quad (25.8)$$

so that polar orbits are in principle more efficient. A family of displaced polar orbits will be considered later in Sect. 25.4. These orbits are able to manage momentum transfer to the reflectors from solar radiation pressure, which would otherwise lead to secular perturbations on the reflector orbit.

Fig. 25.3 Solar reflectors on polar orbits (Aerospace Systems Ltd.).



25.3.3 Static Equilibria

An alternative to a Keplerian orbit about Mars is non-orbiting static equilibria, as discussed by Zubrin and McKay (1997), based on a concept by Forward (1991). These equilibria require that the solar radiation pressure force exerted on the reflector is directly balanced by the local gravitational force leading to a static equilibrium high above the poles. While a detailed three-body analysis of such equilibria has been performed (McInnes et al. 1994), for static equilibria close to Mars (approximating a 2-body rather than 3-body problem) the reflector normal must be directed along the Mars-reflector line to ensure the correct force balance, assuming an ideal reflector. However, the orientation required for a non-ideal reflector to be maintained in static equilibrium leads to most of the reflected radiation being directed away from the planetary body. Therefore, static equilibria do not appear to be a suitable location for Mars solar reflectors (McInnes, 2002).

25.4 Families of Displaced Polar Orbits

In order to re-direct solar radiation to increase the total planetary insolation, reflectors must be deployed on suitable orbits, as discussed in Sect. 25.3. One of the key issues is management of the momentum accumulated by the reflectors due to solar radiation pressure, which poses difficulties for reflectors deployed directly in Keplerian orbit about Mars. In addition, since the reflectors act as mirrors, they will project an image of the solar disk onto the surface of Mars. At the distance of

Mars from the Sun, the solar disk subtends an angle of order 0.01 rad. Therefore, a reflector at distance d from the surface of Mars will project a spot of diameter of order $D \sim d\theta$, assuming a perfectly flat reflecting surface. In order that the spot size is less than the diameter of Mars, the distance d must be less than 1.4×10^6 km, a distance comparable to that to the L_1 point. However, locations near the L_1 point are not attractive due to the long path length the reflected radiation must traverse and consequently the extremely high pointing accuracy required for the solar reflectors. In addition, in order to establish an artificial equilibrium point using solar radiation pressure near L_1 , the aspect angle of the reflectors is rather large, leading to a small projected reflector area (McInnes et al. 1994).

Artificial equilibria between Mars and the L_2 point have been considered by Zubrin, for heating the poles of Mars (Zubrin and McKay 1997). However, as discussed in Sect. 25.3.3, when a reflector with non-ideal reflectivity is considered the required aspect angle of the reflector is such that the re-direct radiation is in fact largely unable to illuminate the poles (McInnes 2002).

As noted earlier, reflectors in orbit about Mars will be strongly perturbed by solar radiation pressure. An attractive family of orbits for solar reflectors are displaced orbits which can be generated by orienting the reflector such that a component of the solar radiation pressure induced force exerted on the reflector is directed normal to the orbit plane (McInnes and Simmons 1992). These orbits are therefore circular and near polar, but are displaced behind Mars in the anti-Sun direction, as shown in Fig. 25.4 for a reflector orbit radius ρ and displacement distance z . The momentum accumulated by the reflector due to solar radiation pressure is offset by the z -component of the local gravitational force. Using a two-body analysis, the requirements for such displaced orbits can be investigated. However, it is found that such orbits can also be established using a full three-body analysis of the dynamics of the problem (Bookless and McInnes 2004).

Using cylindrical polar coordinates, the equations of motion for an ideal specular reflector can be written as

$$\ddot{\rho} - \rho \dot{\theta}^2 = -\frac{\rho}{r} \left(\frac{\mu}{r^2} \right) + a \cos^2 \alpha \sin \alpha \quad (25.9a)$$

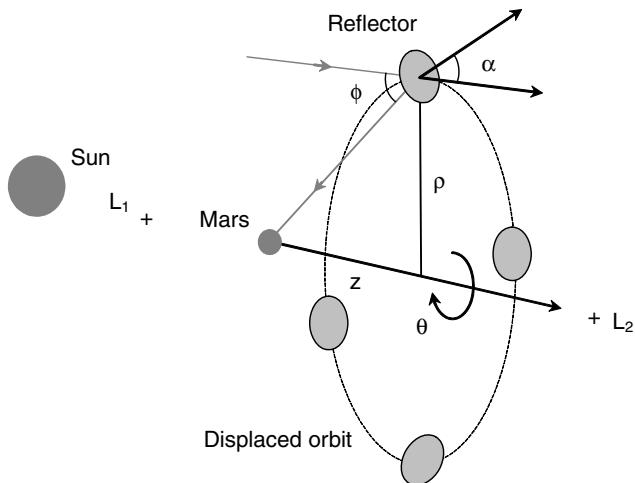
$$\rho \ddot{\theta} + 2\dot{\rho} \dot{\theta} = 0 \quad (25.9b)$$

$$\ddot{z} = -\frac{z}{r} \left(\frac{\mu}{r^2} \right) + a \cos^3 \alpha \quad (25.9c)$$

where μ is the product of the gravitational constant and the mass of Mars.

The reflector orientation is defined by a pitch angle α relative to the Sun-line and the acceleration a induced by solar radiation pressure is defined by $a = 2P/\sigma$ where P is the solar radiation pressure at the distance of Mars from the Sun (1.964×10^{-6} Nm $^{-2}$) and σ is the mass per unit area of the reflector. It is assumed that a is constant (although the magnitude of the solar radiation pressure induced

Fig. 25.4 Solar reflectors on a displaced polar orbit along the Sun-line



acceleration scales as $\cos^2 \alpha$). For a circular displaced orbit it is required that $\ddot{\rho} = 0$ and $\ddot{z} = 0$. Therefore, defining the orbital angular velocity $\omega = \dot{\theta}$ it can be shown that the required reflector pitch angle and acceleration can be written as

$$\tan \alpha = \frac{\rho}{z} \left[1 - \left[\frac{\omega}{\tilde{\omega}} \right]^2 \right], \quad \tilde{\omega}^2 = \frac{\mu}{r^3} \quad (25.10a)$$

$$a = \tilde{\omega}^2 \left[1 + \left[\frac{\rho}{z} \right]^2 \left[1 - \left[\frac{\omega}{\tilde{\omega}} \right]^2 \right]^2 \right]^{3/2} z \quad (25.10b)$$

For the reflector to direct reflected light towards Mars, it is clear from Fig. 25.3 that the reflector pitch angle α is related to ϕ such that $\alpha = \phi/2$. Assuming a perfect specular reflector, an image of the solar disk will then be projected onto Mars, with the image centre on the Mars-reflector line. The required reflector pitch angle is therefore defined as

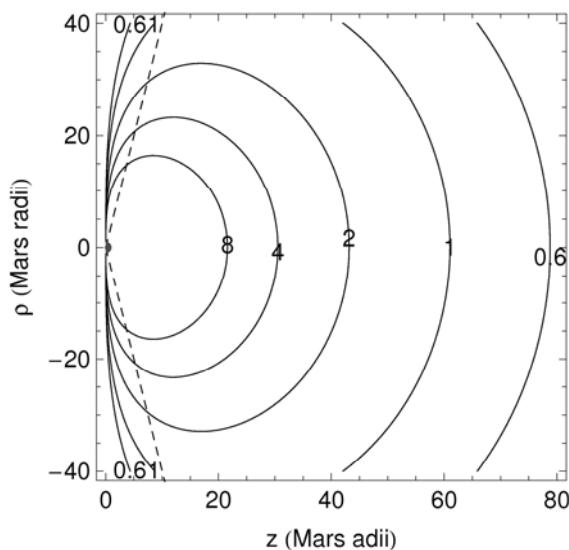
$$\tan \alpha = \tan \left[\frac{1}{2} \tan^{-1} \left[\frac{\rho}{z} \right] \right] \quad (25.11)$$

Since this orientation is defined a priori, the orbital angular velocity of the reflector must now be a variable of the problem in order to ensure correct illumination of Mars by the reflector. Using Eq. (25.10a) it can be seen that the orbital angular velocity required for this desired reflector orientation is given by

$$\omega = \tilde{\omega} \left[1 - \left[\frac{z}{\rho} \right] \tan \alpha \right]^{1/2} \quad (25.12)$$

The corresponding reflector acceleration can now be obtained from Eq. (25.10b). A section of the surfaces of constant reflector acceleration generated by this family of orbits is shown in Fig. 25.5 in the ρ - z plane. Each point on a surface corresponds to a displaced orbit with some radius ρ and displacement distance z in the anti-Sun direction. It can be seen that the surfaces of constant reflector acceleration have a rotational symmetry about the Sun-line, and that for a given reflector acceleration a an orbit with a large radius and small displacement is possible ($\alpha \sim 45^\circ$), or an orbit with a small radius and large displacement ($\alpha \sim 0^\circ$). Clearly, displaced orbits with a small pitch angle α are desirable to maximise the projected reflector area. However, orbits with large displacements are found to be unstable, as will be seen later in Sect. 25.5. An orbit with a relatively modest displacement is still able to offset the momentum accumulated by the reflector due to solar radiation pressure by the z -component of the local gravitational force, while the orbit remains Sun-synchronous with the orbit plane always normal to the Sun-Mars line. While not considered further here, it is noted that if the reflector orbit period is chosen to be an integer number of sidereal days, the spot projected onto the surface of Mars by the solar reflector will follow a repeat ground track. This may have some importance if certain regions are to be preferentially heated.

Fig. 25.5 Reflector requirements for a displaced polar orbit. Contours indicate radiation pressure acceleration a (mm s^{-2}) and the dashed line indicates the stability boundary (stable orbits if $\rho > 4z$).



25.5 Displaced Polar Orbit Stability

Now that a family of orbits with the appropriate reflector orientation has been defined, its stability characteristics will be investigated. This is achieved by linearising the equations of motion about the nominal orbit to obtain a variational equation. The variational equation then describes the motion of the reflector in the neighbourhood of the nominal orbit. The stability of the resulting trajectory may be determined by examining the eigenvalues of the variational equation. If any of

the eigenvalues are real and positive, the motion of the reflector is exponentially divergent so that the orbit will be unstable. However, if all the eigenvalues are purely imaginary, the reflector is expected to remain bound in a neighbourhood of the nominal orbit. It should be noted that a linear analysis is performed, which in the present case provides only necessary conditions for stability, but sufficient conditions for instability.

The non-linear equations of motion will be linearised by perturbing the reflector from its nominal orbit with the reflector attitude fixed in the rotating frame of reference (McInnes 1999). In addition, since the radiation field is uniform over the length-scale of the problem, the solar radiation pressure acceleration experienced by the reflector remains constant during the perturbation. Firstly, a perturbation $\delta = (\xi, \psi, \eta)$ will be added to the reflector position vector at some point $\mathbf{r}_o = (\rho_o, \theta_o, z_o)$ on the nominal orbit such that $\mathbf{r}_o \rightarrow \mathbf{r}_o + \delta$. The variational equation is then obtained from the non-linear equations of motion as

$$\frac{d^2\xi}{dt^2} - 2\omega\rho_o \frac{d\psi}{dt} + M_{11}\xi + M_{13}\eta = 0 \quad (25.13a)$$

$$\frac{d^2\psi}{dt^2} + \frac{2\omega}{\rho_o} \frac{d\xi}{dt} = 0 \quad (25.13b)$$

$$\frac{d^2\eta}{dt^2} + M_{31}\xi + M_{33}\eta = 0 \quad (25.13c)$$

where the linear expansion coefficients are found to be

$$M_{11} = 3\omega^2 + \tilde{\omega}^2 \left[1 - 3 \left(\frac{\rho}{r} \right)^2 \right] \quad (25.14a)$$

$$M_{13} = M_{31} = -3\tilde{\omega}^2 \left[\frac{\rho z}{r^2} \right] \quad (25.14b)$$

$$M_{33} = \tilde{\omega}^2 \left[1 - 3 \left(\frac{z}{r} \right)^2 \right] \quad (25.14c)$$

Owing to the azimuthal symmetry of the problem all derivatives with respect to θ in the linear expansion will vanish. Therefore, the six terms $M_{2,j}$ and $M_{j,2}$ ($j=1-3$) are zero. The set of three coupled, linear ordinary differential equations defined by Eq. (25.13) may be reduced to two by integrating Eq. (25.13b) to obtain

$$\frac{d\psi}{dt} = -\frac{2\omega}{\rho_o}(\xi - \xi_o) \quad (25.15)$$

This equation is in effect a linearised form of Kepler's third law, describing the orbital angular velocity of the reflector relative to the nominal orbit due to the radial displacement ξ . This equation may then be substituted into Eq. (25.13a) to eliminate the azimuthal term. However, this substitution then leads to a constant term $4\omega^2\xi_o$ in Eq. (25.13a) so that the variational equation is no longer homogeneous. It can be shown that the non-homogeneity can be easily removed by rescaling the co-ordinates through a change of variable

$$\xi' = \xi - \frac{4\omega^2 M_{33}}{M_{11}M_{33} - M_{13}M_{31}} \xi \quad (25.16a)$$

$$\eta' = \eta + \frac{4\omega^2 M_{13}}{M_{11}M_{33} - M_{13}M_{31}} \xi \quad (25.16b)$$

Using this transformation a reduced variational system with a set of two coupled equations is then obtained as

$$\frac{d^2}{dt^2} \begin{bmatrix} \xi' \\ \eta' \end{bmatrix} + \begin{bmatrix} M_{11} & M_{13} \\ M_{31} & M_{33} \end{bmatrix} \begin{bmatrix} \xi' \\ \eta' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (25.17)$$

with the reflector free to drift along the nominal orbit in the azimuthal direction.

The stability of the family of orbits defined in Sect. 25.4 may now be investigated by calculating the eigenvalues of the variational equation. The eigenvalues may be obtained in the usual manner by substituting an exponential solution of the form

$$\begin{bmatrix} \xi' \\ \eta' \end{bmatrix} = \begin{bmatrix} \xi_o \\ \eta_o \end{bmatrix} \exp(\lambda t) \quad (25.18)$$

Substituting this trial solution into Eq. (25.17) yields a matrix equation of the form

$$\begin{bmatrix} \lambda^2 + M_{11} & M_{13} \\ M_{31} & \lambda^2 + M_{33} \end{bmatrix} \begin{bmatrix} \xi_o \\ \eta_o \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (25.19)$$

For non-trivial solutions it is required that the determinant of this matrix equation vanishes. The characteristic polynomial of the variational equation is then found to be

$$\lambda^4 + \text{tr}(\mathbf{M})\lambda^2 + \det(\mathbf{M}) = 0 \quad (25.20)$$

The stability characteristics may now be investigated by substituting for the appropriate functional form of ω from Eq. (25.12). Then, regions in the ρ -z plane may be identified where the roots of the characteristic polynomial are purely

imaginary with $\lambda_j^2 < 0$ ($j=1-4$) indicating stable, bound orbits. Such purely imaginary eigenvalues are obtained if $\text{tr}(\mathbf{M}) > 0$ and $\det(\mathbf{M}) > 0$, again noting that only a linear analysis is presented which provides necessary condition for stability, but sufficient conditions for instability.

The family of displaced orbits is partitioned into stable and unstable groups by the conditions $\text{tr}(\mathbf{M}) > 0$ and $\det(\mathbf{M}) > 0$, denoted by the dashed line in Fig. 25.5. This partition between the stable and unstable orbits is approximated by the boundary $\rho = 4z$, with stable orbits in the region $\rho > 4z$. It can be seen then that only orbits relatively close to the planetary terminator are linearly stable, while orbits with a large displacement are unstable. Clearly, it is advantageous to use stable orbits to avoid the need for active control and to mitigate against the risk of failure. Any failure in such an active control system would excite the unstable modes of the orbit and lead to a catastrophic loss of the reflector system. However although unstable, it can be shown that the orbits are strictly controllable (McInnes, 1999).

25.6 Solar Reflector Fabrication

The mass required to fabricate a solar reflector utilizing the passively stable orbit discussed above will now be considered. For an orbit with a radius ρ of 15 Mars radii and a displacement z of 2 Mars radii in the anti-Sun direction, Eq. (25.10b) shows that an acceleration of 5.05 mms⁻² is required. If an orbit with a smaller displacement is considered then the required reflector mass will rise sharply, while a larger displacement will soon lead to an unstable orbit as the partition between the stable and unstable orbits is reached. At the mean distance of Mars from the Sun the solar radiation pressure P exerted on a perfectly absorbing surface is 1.964×10^{-6} Nm⁻². For an ideal specular reflector of mass per unit area σ , the resulting acceleration a is just $2P/\sigma$. Therefore, to generate an acceleration of 5.05 mms⁻², the required mass per unit area σ of the reflector will be 1.01 gm⁻². This is a value similar to the canonical 1 gm⁻² requirement for high performance deployable solar sails (McInnes 1999). The pitch of the solar reflector α can also be obtained from Eq. (25.11) and is found to be 41.2°. The required reflector area to obtain a given increase in solar insolation \dot{Q} can therefore be obtained from $\dot{Q}/S \cos\alpha$. For a 30% increase in total insolation (24.38×10^{15} W) the reflector area is found to be 1.44×10^{13} m². Again, a large number of reflectors are envisaged, distributed azimuthally about the displaced orbit to provide the total required surface area. The mass of the reflector can now be obtained from the required reflector mass per unit area as $\sigma \dot{Q}/S \cos\alpha$, where σ is obtained from the required reflector acceleration as $2P/a$. For the orbit of interest the total reflector mass is found to be 1.120×10^{10} kg. A comparison with large-scale terrestrial engineering ventures is shown in Table 25.1, while a comparison with space-based geo-engineering schemes is shown in Table 25.2.

Table 25.1 Mass comparison with terrestrial engineering ventures

Scale	Mass (kg)	Engineering venture
10^8	6.5×10^8	'Knock Nevis' oil tanker (fully laden)
10^8	6×10^9	Great pyramid of Giza
10^{10}	6×10^{10}	Concrete used for Three Gorges dam
10^{11}	2×10^{11}	Water stored in London's reservoirs
10^{12}	7×10^{12}	World annual CO ₂ emissions

Table 25.2 Mass comparison with terrestrial geo-engineering

Geoengineering concept	Mass (kg)	Area (m ²)	Density (g m ⁻²)
Struck (Lunar L _{4/5} dust cloud)	2.1×10^{14}	-	-
Pearson (Earth orbit dust ring)	2.3×10^{12}	1.10×10^{14}	-
McInnes A (Solar L ₁ reflecting disks)	2.6×10^{11}	6.57×10^{12}	40.2
McInnes B (Solar L ₁ absorbing disks)	5.2×10^{10}	6.57×10^{12}	7.9
Angel (Solar L ₁ refracting disks)	2.0×10^{10}	4.70×10^{12}	4.2

Clearly, the fabrication of a reflector with a total mass of order 10^{10} kg would require an advanced industrial space capacity with the ability to exploit extraterrestrial resources. For example, the mass requirements for the solar reflector can be satisfied by M-type asteroids, which are relatively abundant in Nickel-Iron materials (Gehrels 1979). Assuming the reflector is fabricated from thin metallic film processed from such a body, and that the asteroid has a bulk density similar to that of Iron (7860 kg m^{-3}), an asteroid with a radius of order 70 m will provide the required mass. Including a non-metallic content of as much as 50% will increase the required asteroid radius to order 90 m. The asteroid would require to be processed in-situ, probably using solar heating, and the metallic products extruded into thin film for fabrication of the solar reflector elements. Again, assuming a density of order 1 gm^{-2} , the film thickness would be of order $0.1 \mu\text{m}$. Processing thin film of this thickness is clearly challenging, but ultimately it is the film thickness which drives the total mass of the reflector system. Terrestrial experiments have been performed to produce thin metallic film of comparable thickness, although only on a laboratory scale (Drexler 1979).

Lastly, it is interesting to speculate as to the effect of such a displaced ring of orbiting solar reflectors on the sky, as seen from the surface of Mars. For an observer near the equator, the tip of the ring will slowly rise above the eastern horizon, and would likely be rather dim in the darkening dusk sky due to diffusely scattered light. Since the ring is composed of a vast number of individual reflectors, it would like appear as a dense ribbon of point sources, packed denser than the Pleiades. As local sunset is reached the ring would brighten relative to the darkening sky and gradually rise until it formed a thin bright arc stretching from the northern horizon to the south. Since the orbit of the ring is displaced, the peak brightness of the ring would occur sometime after sunset, providing a spectacularly bright structure in the night sky. After peak brightness, the ring would begin

to dim, drifting westwards and sinking again towards the horizon. This spectacular display would then be repeated prior to dawn with the ring providing a brilliant precursor to local sunrise.

25.7 Conclusions

Large-scale Mars terraforming using orbiting solar reflectors has been considered. In particular, a family of displaced planetary orbits has been presented which are suitable for use by solar reflectors as part of a large Mars terraforming effort. The orbits are configured to ensure that solar radiation is continually reflected onto the planetary surface to increase the total solar insolation. A particularly useful stable family of orbits has been identified which lie close to the planetary terminator. For comparison, static equilibria have been considered as potential locations for Mars solar reflectors. However, the reflector orientation required to ensure static equilibrium leads to significant losses of reflected solar radiation which is not intercepted by the planetary body. Static equilibria therefore have serious problems for use by Mars solar reflectors, while displaced orbits can be configured to be Sun-synchronous and passively stable leading to benefits for their use in large-scale terraforming.

References

- Angel, R.: Feasibility of cooling the Earth with a cloud of small spacecraft near the inner Lagrange point (L_1). *Proc. Nat. Acad. Sci.* 103, 17184–17189 (2006)
- Badescu, V.: Regional and seasonal limitations for Mars intrinsic ecopoiesis. *Acta Astronautica* 56, 670–680 (2005)
- Birch, P.: Terraforming Mars quickly. *J. British Interplanetary Soc.* 45, 331–340 (1992)
- Bookless, J., McInnes, C.R.: Dynamics, stability and control of displaced non-Keplerian orbits, IAC-04-A.7.09. In: 55th International Astronautical Congress, Vancouver (2004)
- Drexler, K.E.: High Performance Solar Sails and Related Reflecting Devices. In: 4th Princeton/AIAA Conference on Space Manufacturing Facilities, Princeton, Paper AIAA-79-1418 (1979)
- Early, J.T.: Space-based solar shield to offset greenhouse effect. *J. British Interplanetary Soc.* 42, 567–569 (1989)
- Fogg, M.: A synergistic approach to terraforming Mars. *J. British Interplanetary Soc.* 45, 315–329 (1992)
- Fogg, M.J.: Terraforming: Engineering Planetary Environments. SAE International, ISBN 1-56091-609-5 (1995)
- Forward, R.L.: Statite: A Spacecraft That Does Not Orbit. *J. Spacecraft and Rockets* 28, 606–611 (1991)
- Gehrels, T.: Asteroids. Univ. of Arizona Press, Tucson (1979)
- Gerstell, M.F., Francisco, J.S., Yung, Y.L., Boxe, C., Aaltonee, E.T.: Keeping Mars warm with new super greenhouse gases. *Proc. Natl. Acad. Sci. USA* 98, 2154–2157 (2001)
- Mautner, M., Parks, K.: Space-based control of the climate. In: Proc. Space 1990, Amer. Soc. Civ. Eng., pp. 1159–1169 (1990)

- McInnes, C.R.: Solar Sail Halo Trajectories: Dynamics and Applications. In: 42nd International Astronautical Congress, Montreal, Paper IAF-91-334 (1991)
- McInnes, C.R., Simmons, J.F.L.: Halo Orbits for Solar Sails II - Geocentric Case. *J. Spacecraft and Rockets* 29, 472–479 (1992)
- McInnes, C.R., McDonald, A.J.C., Simmons, J.F.L., MacDonald, E.W.: Solar sail parking in restricted three-body systems. *J. Guidance Dynamics and Control* 17, 399–406 (1994)
- McInnes, C.R.: Solar Sailing: Technology, Dynamics and Mission Applications. Springer-Verlag Series in Space Science and Technology. Springer, Heidelberg (1999)
- McInnes, C.R.: Non-Keplerian orbits for Mars solar reflectors. *J. British Interplanetary Soc.* 55, 74–84 (2002)
- McInnes, C.R.: Planetary Macro-engineering using Orbiting Solar Reflectors. In: Badescu, V., Cathcart, R., Schuiling, B.D. (eds.) *Macro-Engineering: A Challenge for the Future*, pp. 215–250. Springer, Heidelberg (2006)
- McKay, C.P., Toon, O.B., Kastling, J.F.: Making Mars habitable. *Nature* 352, 489–496 (1991)
- Oberg, J.E.: *New Earths: Restructuring Earth and Other Planets*. New American Library Inc., New York (1981)
- Sagan, C.: The planet Venus. *Science* 133, 849–858 (1961)
- Sagan, C.: Planetary engineering on Mars. *Icarus* 20, 513–514 (1973)
- Zubrin, R., McKay, C.: Technological requirements for terraforming Mars. *J. British Interplanetary Soc.* 50, 83–92 (1997)

Chapter 26

Ecopoiesis and Liquid Water Transportation on Mars

Viorel Badescu, Dragos Isvoranu, and Richard B. Cathcart

Polytechnic University of Bucharest, Romania and Geographos, Burbank, CA, USA

26.1 Introduction

As the means of safe transportation between Earth and Mars is technically developed, one of the outcomes will likely be assisted colonization of Mars. In this chapter we shall exemplify potential actions under a suite of possible future Mars settlement scenarios. The focus is on the means of providing human colonies with a vital resource: water.

The existing climate on Mars is too cold and dry for Earth-like life. The level of UV radiation at the planet's surface is very high and the atmosphere is thin and contains about 95 % of carbon dioxide. However, concepts and models for permanent settlement and eventual terraforming of Mars are already developed. Two steps are usually envisaged when terraforming is considered. The first step is *Mars ecopoiesis*, which consists in changes needed to introduce anaerobic life. The result is expected to be a free biosphere able to use solar energy and to run biogeochemical cycles on a planet – wide scale. The second step is the effective *Mars terraforming*, that finally leads to an environment habitable for man (Fogg 1995 p. 219, 302). Five main tasks should be solved for Mars terraforming: (i) surface temperature must be raised; (ii) atmospheric pressure must be increased; (iii) chemical composition of the atmosphere must be changed; (iv) the surface must be made wet and (v) the surface flux of UV radiation must be reduced. Various technological solutions were proposed to accomplish these five tasks and they are grouped into two categories. First, there are the *intrinsic techniques*, which include for example reducing the albedo of CO₂ ice by applying a thin dark material layer (Sagan 1973; Averner and MacElroy 1976) or inject trace greenhouse gases into the atmosphere (Lovelock and Allaby 1984; McKay et al. 1991). Second, there are the *extrinsic techniques*, which include augmentation of insolation with reflected sunlight (Zubrin and McKay 1997; see also Chap. 25) and optimize orientation of Martian rotation axis with its orbit's perihelion by applying a torque to Mars with orbiting masses (Burns and Harwit 1973).

A first action during the *ecopoiesis* of Mars is commonly considered to be thickening the atmosphere and using its increased greenhouse effect to warm the planet's surface. Various techniques proposed to facilitate the release of Martian volatiles are shown in Table 6.4 of Fogg (1995 p. 277). They include for example nuclear mining of carbon dioxide on Mars, regolith vaporization and impact devolatilization. Importation of volatiles is also considered. The next action consists in raising the surface temperature of Mars. To this aim one usually proposes the modification of three parameters, namely the insolation, the albedo and the greenhouse effect. In their study, Averner and MacElroy (1976) concluded that the extensive growth of terrestrial organisms requires an increase in the stability and abundance of liquid water. This requirement sets a limitation for ground surface temperature, which should normally exceed the water freezing point. The so-called "runaway greenhouse model" of Martian ecopoiesis represented for a long time the "standard paradigm" due to its promises. However, the success of the runaway greenhouse mechanism is far from guaranteed (Fogg 1995 p. 266). This prompted some researchers to develop new approaches. A synergic scenario combining processes which appear, individually, incapable of warming Mars adequately was proposed in Fogg (1995 p. 300). The synergic model accepts a more conservative assessment of the Mars volatile inventory than the standard ecopoiesis paradigm.

Section 26.2 gives details about the thermal environment on Mars after the ecopoiesis process is completed. In the absence of rainfall—even post-terraforming—we suggest that artesian basins and the two polar water-ice caps may serve to provide a Mars civilization with a basic water supply. Section 26.3 deals with large scale water transportation.

26.2 Ecopoiesis of Mars

Most researchers approached Mars ecopoiesis by using (i) steady-state and (ii) globally surface averaged models. In reality, Mars' surface has a low thermal inertia, and the alternation of day/night does not permit a thermal steady-state. In this chapter we shall relax in part the two assumptions above. Only intrinsic ecopoiesis techniques are considered here. Consequently, no device of augmenting the level of incident solar irradiance at global level on the planet's surface is considered. However, such devices may be envisaged for local or regional purposes, as described below. Also, the volatile problem is supposed to be already solved. We are interested to see whether additional *intrinsic actions* are necessary to diminish the influence of the day/night alternation on the amplitude of the daily excursion of ground surface temperature. The basis of our approach is the synergic model proposed in Fogg (1992).

26.2.1 Mars Ecopoiesis Models

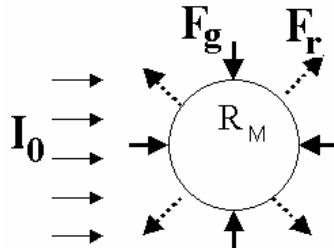
We denote by φ the latitude of a site on Mars' surface. We assume the Martian year begins at the vernal equinox, when the areocentric longitude L_s of the Sun, measured in the orbital plane of the planet, is $L_s=0^\circ$. Other details about the astronomical model

are given in Chap. 2. Note that due to the higher eccentricity of Mars as reported to Earth, there is an important asymmetry of seasons. Consequently, the southern summer is warmer and has more sunshine than the northern summer on Mars.

26.2.1.1 Steady-State Models

In this section the model in Fogg (1992) (called the “synergic model”) and the model by Zubrin and McKay (1997) are briefly presented. The first one has a globally averaged temperature for Mars’ surface while the second gives a rough latitudinal distribution of Mars surface temperature.

Fig. 26.1 Energy balance for the whole Mars surface. I_0 - solar energy irradiance; F_g – energy flux per unit surface area due to greenhouse gases; F_r – energy flux per unit surface area emitted by the planet surface.



The synergic steady-state model involves an energy balance at the level of the whole Mars ground surface. Two main energy flux densities are incident on that surface: the solar irradiance I_0 and the energy flux density F_g due to the IR radiation emitted by the atmospheric greenhouse gases, respectively (Fig. 26.1). The heated surface emits the energy flux density F_r . As a result, an average Mars surface temperature (say T_s) is obtained. Then, the steady state-energy balance equation is:

$$\pi R_M^2 (1 - a_M) I_0 + 4\pi R_M^2 F_g = 4\pi R_M^2 \sigma T_s^4 \quad (26.1)$$

Here R_M is Mars radius, a_M is the ground albedo averaged over the whole surface of Mars while σ is the Stefan-Boltzmann constant. The first term in the l.h.s. of Eq. (26.1) is the solar energy flux absorbed by Mars surface while the second term in the l.h.s. of Eq. (26.1) is the energy flux received by the ground from the greenhouse gases (here the absorptance of the ground for IR radiation was assumed to be 1). The r.h.s. of Eq. (26.1) contains the energy flux emitted by the ground through IR radiation. Equation (26.1) can be re-written in the form:

$$\frac{1-a}{4} I_0 + F_g \equiv \frac{1-a}{4} I_0 (1 + \tau_{st}) = \sigma T_s^4 \quad (26.2)$$

This defines the equivalent “standard” greenhouse gases opacity as:

$$\tau_{st} \equiv \frac{4}{1-a} \frac{F_g}{I_0} \quad (26.3)$$

Note that τ_{st} is computed as a sum of non-interacting contributions τ_i from all the η_{gas} greenhouse gases:

$$\tau_{st} = \sum_{i=1}^{n_{gas}} \tau_i \quad (26.4)$$

Use of Eq. (26.2) allows us to obtain the averaged ground temperature as:

$$T_s = \left[\frac{1 - a_M}{4\sigma} I_0 \left(1 + \sum_{i=1}^{n_{gas}} \tau_i \right) \right]^{1/4} \quad (26.5)$$

The greenhouse gases considered in the standard model are carbon dioxide, methane, ammonia, CFCs and water vapor. Their individual opacities are computed in the standard model by using the following semi-empirical formulas:

$$\tau_{CO_2} = 0.9 \cdot P_{tot}^{0.45} \cdot P_{CO_2}^{0.11}, \quad \tau_{CH_4} = 0.5 \cdot P_{CH_4}^{0.278}, \quad (26.6-9)$$

$$\tau_{NH_3} = 9.6 \cdot P_{NH_3}^{0.32}, \quad \tau_{CFC} = \frac{1.1 \cdot P_{CFC}}{P_{CFC} + 0.15}$$

where P_{tot} is the total atmospheric pressure (in bar) while P_{CO_2} (in bar), P_{CH_4} (in bar), P_{NH_3} (in bar) and P_{CFC} (in mbar) are the partial pressures of the respective component.

There are indices showing that Mars does have the capacity to sequester several hundreds meters of water, mostly as ground ice overlying liquid at greater depth (see Table 5.5 of Fogg (1995 p 217)). The synergic model assumes the water vapor is in thermal equilibrium with the ground surface. Consequently, the water vapor partial pressure (in bar) is computed with:

$$P_{H_2O} = R_h P_0 \exp \left(- \frac{L_{H_2O}}{RT_s} \right) \quad (26.10)$$

where R_h is the relative humidity, $P_0 (= 1.4 \cdot 10^6 \text{ bar})$ is a reference water vapor pressure, $L_{H_2O} (= 43655 \text{ J/mol})$ is water latent heat and $R (= 8.314 \text{ J/mol/K})$ is the universal gas constant. The water vapor opacity is computed by using:

$$\tau_{H_2O} = P_{H_2O}^{0.3} \quad (26.11)$$

where P_{H_2O} is in bar. When the globally averaged ground temperature T_s is computed by using the synergic model an iterative procedure should be used because of the dependence of water vapor pressure on T_s . This is a consequence of the hypothesis that water vapor is in thermal equilibrium with the ground surface. If this assumption is relaxed the water vapor partial pressure can be chosen as a free variable in Eq. (26.11) and computation of T_s can be performed directly (i.e. without iteration). Both procedures will be used in this section.

The total atmospheric pressure is given by:

$$P_{tot} = P_{CO_2} + P_{CH_4} + P_{NH_3} + P_{CFC} + P_{H_2O} \quad (26.12)$$

A simplification of the synergic model was adopted here as we neglected the influence of CO₂ deposits on the atmospheric pressure as a result of the increase in ground temperature. Similar to the synergic model, no dependence of the gas pressure on the altitude was considered here.

We follow Zubrin and McKay (1997) in assuming the polar temperature T_p is given by:

$$T_p = T_s - \frac{\Delta T}{1 + 5P_{tot}} \quad (26.13)$$

where $\Delta T = 75 K$, which is the difference between the mean global and polar temperatures in the absence of an atmosphere, and P_{tot} is the atmospheric pressure (in bar), included to take account of atmospheric transport of heat from tropics to polar regions. The relation for the maximum tropical temperature T_t is taken to be

$$T_t = 1.1 T_s \quad (26.14)$$

and the latitudinal global temperature distribution is given by:

$$T(\varphi) = T_t - (T_t - T_s) \sin^{1.5} \varphi \quad (26.15)$$

The assumption of the quoted authors was that when the tropical temperature $T_t > 0^\circ C$, which occurs when the mean global temperature T_s exceeds $-25^\circ C$, then some part of the planet will be habitable, on the assumption that this is for those parts of Mars above freezing.

The following values are adopted for the modified Mars atmosphere (Fogg 1995 p. 18): $P_{CO_2} = 0.35$ bar, $P_{PFC} = 10 \mu\text{bar}$, $P_{NH_3} = 100 \mu\text{bar}$, $P_{CH_4} = 1 \text{ mbar}$. This atmosphere is referred to as “Atmosphere 1” in the following. When water vapor in thermal equilibrium is considered, the relative humidity $R_h = 0.7$ was adopted. These values are considered to be reasonable upper limits of climate forcing methods. The albedo was assumed to be $a_M = 0.09$ on the whole surface of Mars. This value is the albedo of exposed bedrock in the Syrtis Major (dark region). A similar reflectance value can be obtained by depositing a 1-mm thick layer of material with albedo of carbon black as suggested by Sagan (1973). This thin layer could be prone to dispersal by the wind but technological solutions to fix this problem exist.

The present-day mean global surface temperature on Mars is $-56^\circ C$ (Fogg 1995 p. 202-203)). When Atmosphere 1 is used, Eq. (26.5) predicts an average Mars surface temperature T_s of $29.1^\circ C$ (Badescu 2005) and the water vapor partial

pressure at equilibrium 29.0 mbar. Another Martian modified atmosphere (Atmosphere 2) has been also studied in Badescu (2005). There, $P_{CO_2} = 0.35$ bar and $P_{PFC} = 1 \mu\text{bar}$ and water vapor in thermal equilibrium with the ground is considered. Then, the obtained average Mars surface temperature T_s is 9.6 °C and the water vapor partial pressure at equilibrium 8.9 mbar. These values are close to the Earth climate, where the average temperature is around 14 °C.

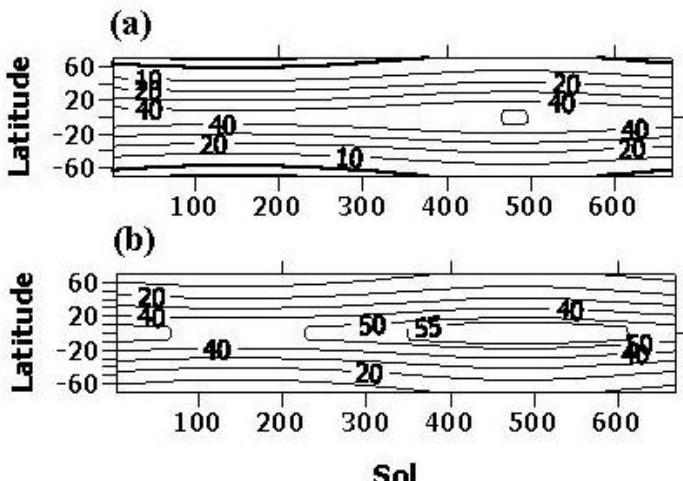


Fig. 26.2 The dependence on latitude and sol number of the surface ground temperature (degree Celsius) predicted by the synergic steady-state model. (a) Water vapor neglected; (b) Water vapor at thermal equilibrium with the ground is considered. Atmosphere 1 has been considered.

Figure 26.2 shows the dependence of the surface temperature predicted by Eq. (26.15) on latitude and sol number. Atmosphere 1 has been considered. Latitudes lower than 70° N and S were considered as the existence of ice and solid carbon dioxide caps is not a restriction for terraforming. When water vapor is neglected, the maximum temperature is about 50 degrees Celsius (Fig. 26.2a). For higher north and south latitude (more than 55 degrees) the temperature is below freezing during the first half of the Martian year. When water vapor at equilibrium are considered, the freezing water temperature is exceeded during the whole year for all the latitude range analyzed (Fig. 26.2b). The maximum temperature is about 60 degrees Celsius. Note that the water partial pressure varies between about 0.023 bar at sol 100 and about 0.038 bar at sol 480.

26.2.1.2 Time-Dependent Model

Now we turn our attention to an arbitrary moment and an arbitrary site on Mars' surface (Fig. 26.3). In this case the steady-state energy balance per unit surface area is given by:

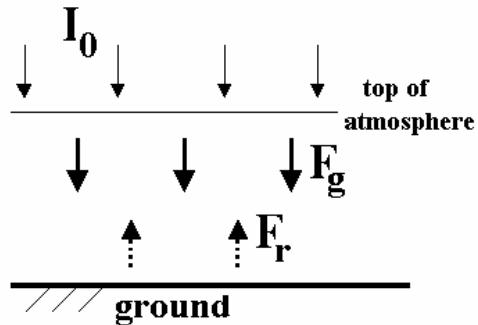
$$(1 - \alpha_{soil})\chi_{atm}I_0 + F_g = \varepsilon_{soil}\sigma T_{soil}^4 \quad (26.16)$$

where T_{soil} is the ground surface temperature, α_{soil} and ε_{soil} are Mars soil albedo and surface emissivity, while χ_{atm} is the transmittance of the atmosphere for solar radiation, defined as:

$$\chi_{atm} \equiv \frac{I_{0,ground}}{I_0} \quad (26.17)$$

where $I_{0,ground}$ is the global solar irradiance at ground level. $I_{0,ground}$ is the sum of direct and diffuse solar irradiance. In writing Eq. (26.16) one supposes the ground is a gray-body.

Fig. 26.3 Energy balance for a location on Mars surface. I_0 - solar energy irradiance; F_g – energy flux per unit surface area due to greenhouse gases; F_r – energy flux per unit surface area emitted by the planet surface.



One can define an equivalent *instantaneous* gray opacity of the atmospheric greenhouse gases τ_{inst} as follows:

$$\varepsilon_{soil}\sigma T_{soil}^4 = (1 - \alpha_{soil})\chi_{atm}I(1 + \tau_{inst}) \quad (26.18)$$

By comparing Eqs. (26.16) and (26.18) one sees that:

$$\tau_{inst} \equiv \frac{F_g}{(1 - \alpha_{soil})\chi_{atm}I} \quad (26.19)$$

It is easy to see from Eqs. (26.3) and (26.19) that, under the condition $T_s \equiv T_{soil}$:

$$\tau_{inst} = \tau_{st} \frac{1 - a}{1 - \alpha_{soil}} \frac{1}{4\chi_{atm} \cos \theta_z} \quad (26.20)$$

where θ_z is solar zenith angle. Equation (26.20) is useful as it allows using the semi-empirical relationships Eqs. (26.4) and (26.6-26.9) proposed to evaluate τ_{st} in the steady-state synergic model to compute the gray opacities τ_{inst} for particular places and arbitrary moments on Mars surface.

In heat transfer problems with time-periodic boundary conditions the so-called penetration length is important. For the solar heating of Mars surface the time pe-

riod is normally one day (say t_{day}) and the penetration length is called “diurnal thermal skin depth” (say δ). It is computed with (Edgett and Christensen 1991):

$$\delta = \sqrt{\frac{t_{day} k}{\rho c \pi}} \quad (26.21)$$

where k is thermal diffusivity, c is ground’s specific heat and ρ is ground mass density. For the medium sand on Mars the diurnal thermal skin is about 7.5 cm (Edgett and Christensen 1991). Note that the time period is different from a normal day for latitudes close to North and South poles during “arctic” days and nights. The energy balance for a slab of thickness δ is:

$$\delta \rho c \frac{dT}{dt} = (1 - a_{soil}) t_{atm} I (1 + \tau_{inst}) - \varepsilon_{soil} \sigma T^4 \quad (\text{daylight}) \quad (26.22a,b)$$

$$\delta \rho c \frac{dT}{dt} = -\varepsilon_{soil} \sigma T^4 \quad (\text{nighttime})$$

Total upward thermal radiance at the top of the Martian atmosphere is almost indistinguishable from that of a gray Lambert surface with emissivity $\varepsilon_{soil} = 0.9$ (Crisp 1990). Other details can be found in Badescu (2005). Note that the observed surface temperature near the North Pole during periods with perpetual darkness during winter remain nearly constant at about 160 K (Santee and Crisp 1993; Hourdin et al. 1995). This proves that other mechanisms to heat the ground except the radiative transfer exist, as for example carbon dioxide condensation. These mechanisms are neglected in the simple model Eqs. (26.22).

One can easily see from Eq. (26.21) that the quantity $\delta \rho c$ entering Eq. (26.22) can be expressed as:

$$\delta \rho c = \sqrt{\frac{t_{day}}{\pi}} i \quad i \equiv \sqrt{k \rho c} \quad (26.23,24)$$

where i is thermal inertia parameter, reported for the whole present Mars surface in Fig. 2 of Pollack et al. (1990), where this parameter varies with latitude between 100 and 425 $\text{J m}^{-2} \text{s}^{-1/2} \text{K}^{-1}$. Note that the bedrock of about two-third of the planet – the southern hemisphere and some extensions over the equator – is thought to be overlain by regolith (i.e. a layer of debris and dust made of meteoritic impact that forms the uppermost surface of planets, satellites and asteroids – for details see Chap. 19). Data about the thermal inertia parameter are reported in the file INERTIA.DAT in the databases with information collected by the Viking Landers (Lee 1995). The average value in this file is $262 \text{ J m}^{-2} \text{s}^{-1/2} \text{K}^{-1}$. This value is adopted here.

Most of the subsurface rocks on Mars are expected to be basalts with low silicon content. Indeed, the analyses conducted by Viking Landers suggest that rocks in the two sites studied are similar to terrestrial basalt lava (Fogg 1995 p. 205). The following standard properties of basalt are used in this chapter: $c = 0.79$

kJ/(kgK) , $k=1.67 \text{ W/(mK)}$, $\rho=2900 \text{ kg/m}^3$ (Rasznjevic 1978). The resulting thermal inertia parameter of the basalt calculated with Eq. (26.24) is $1956 \text{ Jm}^{-2}\text{s}^{1/2}\text{K}^{-1}$, much higher than in case of regolith.

In the Equatorial Zone of Mars, temperatures may peak as high as 20°C in the early afternoon but plunge as low as -90°C just before dawn. The temperature of the wintertime pole can fall to below -130°C (Fogg 1995 p. 202-203).

Let us examine the energy balance during the daytime. At thermal equilibrium the energy flux incident on Mars' surface (first term in the r.h.s. of Eq. (26.22a)) should be equal to the energy flux radiated by the surface (the second term in the r.h.s. of Eq. (26.22a)). A hypothetical instantaneous *diurnal* thermal equilibrium ground surface temperature can be defined in this case. However, due to day/night alternation there is not enough time for this thermal equilibrium to be reached and normally the ground surface temperature is below the thermal equilibrium temperature during the daylight period and higher than the *nocturnal* thermal equilibrium temperature during nighttime.

The ground surface temperature was evaluated during the whole Martian year on a hourly basis, solving Eqs. (26.22), for all latitudes between 70°S and 70°N with steps of 10 degrees. The longitudinal influence was neglected. Water vapor in thermal equilibrium with the soil is considered here. In this case a positive feedback relationship between water vapor and surface temperature exists. Water vapor freezes out of the atmosphere as the surface cools to adjust to radiative equilibrium, reducing the greenhouse effect and accentuating the cooling (Fogg 1995 p. 233). Figure 26.4 shows the minimum and the maximum daily ground surface temperature. Atmosphere 1 has been considered. When the ground consisting of regolith is considered, the minimum soil temperature is generally below -60 degrees Celsius during the whole year (Fig. 26.4a) while the maximum soil temperature exceeds 60 degrees Celsius for most of the latitudes considered (Fig. 26.4b). The partial water pressure is higher than 0.1 bar with a maximum of 1.06 bar around sol 480 (Fig. 26.4c).

The low temperature during the night is a serious obstacle for the ecopoiesis process. A consequence of the alternation of very low temperature during the night and high temperature during the daylight period would be the occurrence of massive "rains" during the evening followed by freezing during the night and ice melting in the morning. Our simple model does not take into account these processes that involve accumulation and removal of latent heat.

A possible solution is to remove or to compact the regolith layer. When a ground consisting of basalt is considered, the daily minimum ground temperature is above water freezing for some latitudes, for various times during the year (Fig. 26.4d). Thus, during the first half of the year the daily minimum temperature exceeds the freezing point in the Northern hemisphere while during the second half of the year this happens for the Southern hemisphere *and* part of the Northern hemisphere. The daily maximum soil temperature exceeds the freezing point during the whole year for the latitudes $55^\circ\text{S} - 55^\circ\text{N}$ (Fig. 26.4e). Note that the partial water pressure is a maximum of 0.45 bar around sol 480 (Fig. 26.4f).

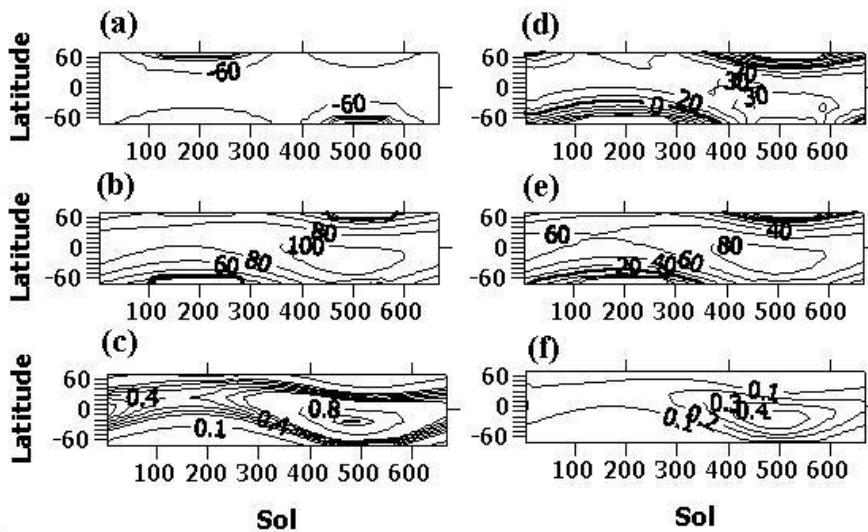


Fig. 26.4 The dependence on latitude and sol number of the surface ground temperature (degree Celsius) predicted by the time-dependent model (Eqs. (26.22)). Water vapor at thermal equilibrium with the ground is considered. (a-c) Ground consisting of regolith; (d-f) ground consisting of basalt. (a) and (d) daily minimum ground surface temperature; (b) and (e) daily maximum ground surface temperature; (c) and (f) partial water vapor pressure. Atmosphere 1 has been considered.

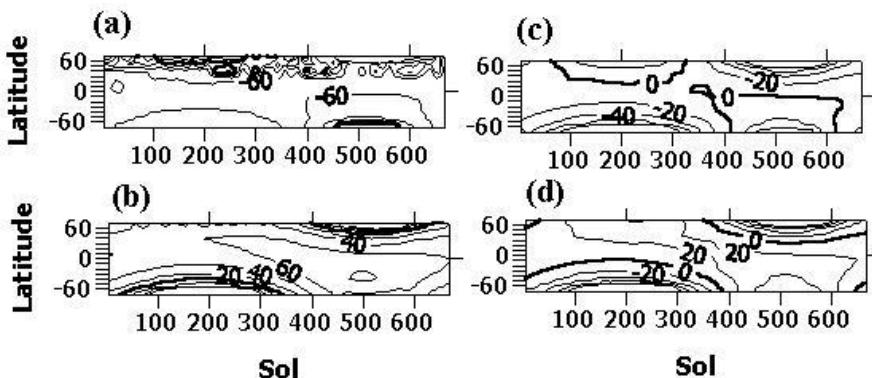


Fig. 26.5 The dependence on latitude and sol number of the surface ground temperature (degree Celsius) predicted by the time-dependent model (Eqs. (26.22)). A constant water vapor partial pressure was considered ($P_{H_2O} = 0.01$ bar). (a-b) Ground consisting of regolith; (c-d) ground consisting of basalt. (a) and (c) daily minimum ground surface temperature; (b) and (d) daily maximum ground surface temperature. Atmosphere 1 has been considered.

The rather high values of the water vapor pressure make questionable the acceptability of the assumption of thermal equilibrium between water vapor and soil. This would involve an atmosphere very rich in water, which is hard to accept. Another hypothesis is accepted now. We assume a constant water vapor partial pressure during the whole year and for all latitudes. As a numerical value we adopted $P_{H_2O} = 0.01$ bar, which is in the range of values found on Earth. Figure 26.5 shows the results. Atmosphere 1 has been considered. In case of a ground consisting of regolith, the daily minimum soil temperature is generally below – 60 degrees Celsius during the whole year (Fig. 26.5a) while the daily maximum soil temperature exceeds 20 degrees Celsius for most of the latitudes considered (Fig. 26.5b). From the point of view of Mars ecopoiesis this is quite similar to the case of Figs. 26.4a and 26.4b.

When a ground consisting of basalt is considered, the daily minimum soil temperature is positive for some latitudes, which change during the year (Fig. 26.5c). Thus, during part of the first half of the year (about 275 sols) the daily minimum temperature exceeds the freezing point in the Northern hemisphere while during about 200 sols in the second half of the year this happens for the Southern hemisphere. During the first half of the year the maximum daily temperature exceeds the freezing point in the Northern hemisphere while during the second half of the year this happens for the Southern hemisphere *and* part of the Northern hemisphere (Fig. 26.5d). This is a dramatic change as compared to the case of Figs. 26.4d and 26.4e).

Our results suggest that the list given in the Introduction with the five tasks to be accomplished for Mars terraforming should be added to with: (vi) the ground surface thermal inertia must be increased. Even in this case it seems that only terrestrial life adapted for hibernation and/or migration will survive on the terraformed Mars. Other results obtained under the assumption of Atmosphere 2 are reported in Badescu (2005).

26.3 Large Scale Water Transportation on Mars

26.3.1 Mars Settlements and Water Need

It is reasonable to accept that politically-defined settlement regions of a terraformed Mars conform exactly to Mars's anthropogenic periglacial catchments (Kauffman 2002). “Periglacial” applies to landscapes in which frost-related geomorphic event-processes and/or permafrost are dominant or characteristic (French 2003). There are two main forms of ground ice: ice that is structure-forming, bonding the enclosing sediments, or large bodies such as Mars's polar ice caps that are more or less pure water ice. Organization of a human society settling Mars catchment by catchment, successfully melds into domed and outdoor designer ecosystems two fundamental macro-engineering landscape concepts—“carrying capacity” and its tight conformance to explored and known “engineering geological structure” (Trofimov and Averkina 2007; Graham et al. 2008) as well as known regional differences in soil compositions.

Future human Mars settlers wielding high-technology will need to scrabble vigorously, artificially sublimating and receding polar ice, digging for frozen groundwater, and transporting mined useable liquid freshwater great distances, generating a unique Anthropic Hydrologic Cycle in doing so. The plan proposed in Birch (1992) may be modified, this time for the rapid ice melting (by focusing lens-concentrated sunlight on the planet's polar caps at various places). The existing technology could increase the amount of sunlight ("heat blasts") impinging Earth's surface by "...40 times that of the Sun on a [ground surface] disc 3 miles in diameter for a period of about a minute" (Buckingham 1971). Application to Mars seems feasible.

26.3.2 Previous Studies about Water Transportation on Mars

Giovanni Domenico Cassini (1625-1712) first observed how the white-colored non-identical poles of Mars waxed and waned seasonally during 1666. Widely reported naked eye and telescopic observations of Mars during the period 1892 to 1910 (prior to the revelatory imaging transmitted by Mariner 9 during 1971-72), forged the popular icon of a hyper-arid Mars landscape fully managed by some kind of intelligent animal, creatures possibly intellectually superior to humans, utilizing a desert network of canals globally distributing all available liquid freshwater (Lane 2005; Lane 2006). These canals were thought to be very wide and the cropland flood-irrigated in both hemispheres of Mars made them "visible" to Earthlings. Some of Mars's "canals" actually correspond to extant landscape features, imaged by Mariner 9, such as craters and valleys as well as some ground albedo features (Gerstbach 2003). Historically, it is interesting to note the deduction of the Royal Navy officer-electrician Herbert Francis Hunt (1878-1939) a century ago (Hunt 1909):

"I should like to suggest that these [Mars] canals may perhaps be used for power-storage purposes. In Mars, possibly, there are seasons of winds or monsoons during which the upper reaches of the canals would be pumped full by innumerable windmills, and the power thus stored utilised during calm seasons, and transmitted electrically for lighting, heating, and general power purposes. For a population which had exhausted all its mineral fuel, which possessed no extensive ocean, and whose soil and climate were unsuitable for the growth of fuel, this would indeed appear to be the only means of obtaining heat and power. The same canals could serve the triple purposes of communication, power, and irrigation."

Nowadays, planet observers are certain Mars is without true canals (Zahnle 2001; Jones 2008). Since 1905, thousands of areographic mapping photographs have been made at the Lowell Observatory in Arizona, USA of Mars's high-latitude regions and yet it was not until the mid-1960s, after Earth-orbiting satellites acquired them for the space-faring ecosystem-states, that mapping images of the Earth's North and South Poles became publicly available (Fleming 2007).

Percival Lowell (1855-1916) undoubtedly provided an earlier conceptualization of a global scale water circulation system on Mars and its basic control mechanisms.

However, the first engineering solution for large scale water transportation on Mars has been given by Charles Edward Housden and his brief biography follows.

Housden was born 19 July 1855; he was educated as a civil engineer at the Indian Civil Engineering Colleges. By October of 1876, he was appointed to the Public Works Department of the Government of India, having passed a vernacular examination at the departmental standard as well as a colloquial examination in Burmese. *Circa* June 1903, he was serving as a sanitary engineer to the Government of the City of Rangoon; 1908 saw him engaged as Superintending Engineer and Sanitary Engineer, East Bengal and Assam, stationed at Shillong. While resident there, he helped plan and execute the continuing regional infrastructure repairs necessary subsequent to the tremendously destructive 12 June 1897 Great Assam earthquake (Billham and England 2001). An organizational rule with the effect of a law then governing all UK civil servants in India imposed mandatory retirement at age 55, so Housden left the service during July 1910. On 27 November 1912, C.E. Housden became a member of the British Astronomical Association, while then residing in London. His landmark article, leading shortly to his pre-World War I book on Mars's optical puzzles was published in 1913. Between 1907 and 1914, Housden had published three major reference books on topics relevant to macro-engineered water supply and sanitary disposal systems as well as macroproject management. Housden died during 1921, while residing in Hove, Sussex, England¹.

In a succinct technical-level book text Housden worked up what amounts to a briefly outlined Mars macro-project of enormous planning, construction and social complexity; using his honed practical skills, he devised a scheme essential to successfully harvest the presumed hemisphere-alternating periglacial runoffs (Housden 1914). Although Housden's main-point was to support the hypothesis that living builders were at work on the 'canals', his scheme can also obviously be used to consider the colonisation of Mars by *Homo sapiens*. He calculated a complex freshwater canal global network that might enable the tapping and storage of the summertime Mars polar water-ice polar glaciers deriving from seasonal melts that were channeled and distributed by the biologically enfeebled, perhaps even hungry (Vernon 2007), but technically accomplished ancient Martians postulated by Percival Lowell. He also estimated the energy budget actually required to move freshwater from Mars's North and South Poles in order to facilitate precision irrigation pipeline agriculture (known as "site-specific management") and to keep city reservoirs filled within an Equatorial Zone belt-shaped farmed and inhabited region (Housden 1914). His statements and calculations pertaining to the "riddle"—the difference in the precipitation (snowfall and rainfall) of atmospheric water vapor in Earth and Mars—do have the tone of modernity that 21st Century expertise displays (Montanes 2005).

UK geologist Robert Lionel Sherlock (1875-1948) valued C.E. Housden's hydrological expertise and his findings regarding Mars. Sherlock made this insightful areomorphological deduction (Sherlock 1922):

¹ No published biography of C. E. Housden is available; this information was obtained with the on-line assistance of the staff in the British Library (London).

“Can we foresee the final results when all possible engineering works shall have been carried out? Is it possible that in the yet far distant end Earth will come to a state said to exist in Mars, and be covered by enormous canals from pole to equator? The canals of Mars, if they really exist, a question still under discussion, are far greater than any our engineers imagined. The battle of the Martians with Nature has been on a much more gigantic scale than Man’s conflict, and yet we hear that the Martian is on the point of extinction, and Mars of becoming totally lifeless. Even on Mars the mighty engineering works seem merely to scratch the skin of the planet, and the final result of Martian activity in the solar system seems likely to be infinitesimal.”

Sherlock’s is one of the few attempts made before 1966-1972 to gain an omniscient view of Earth humankind’s ever-progressing and increasingly powerful transformative mechanical powers. Presciently, Sherlock also opined in his 1922 book that “...there are indications that the doctrine of Uniformity has been carried too far”. Sherlock’s 1922 insight is supported by modern researchers (Smil 2008; Bostrom and Cirkovic 2008; Hanslmeier 2009). An interesting discussion of how the early Mars science inspired a reconsideration of the human-environment relationship may be found in Markley (2005).

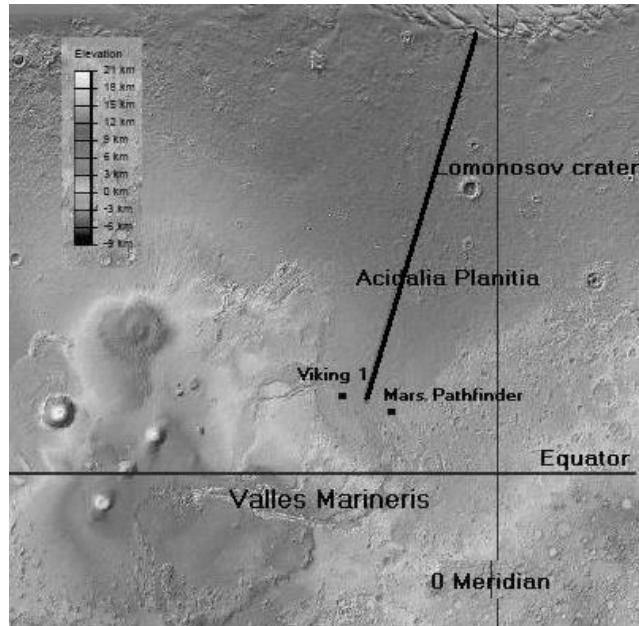
26.3.3 A Case Study

By the mid-19th Century Thomas Dick (1774-1857) had postulated that a natural Mars unaltered by humans might host a population of $15.5 \cdot 10^9$ persons whereas Paul Birch foresaw Mars—quickly terraformed during a ~50 year-long 21st Century period—ultimately becoming a homeland with potentially $3 \cdot 10^9$ settled and established pioneering human inhabitants (Birch 1992). Here a more conservative case will be considered, i.e. a Mars colony of 3 million habitants. Whenever and wherever the prime Mars base is established, that place will become the equivalent of a capital city on Earth. However, one expects that this settlement will be located near the Equator due to mild weather conditions. The total volume of freshwater to be delivered to the human settlement depends on net irrigation needs (evapotranspiration plus leaching minus extant soil water stores plus any other contributing sources) and on application efficiencies, the “site-specific management” aspect, such as drip-irrigation.

Noteworthy is the macro-engineering concept of a continuous line of connected pipe segments (installed on the ground surface or buried) supplying freshwater to consumers the purpose of which is to convey water from one region to another without causing land erosion and reducing the amount of evaporation into the atmosphere. Note that on Earth, the very first really long-distance freshwater pipelines were built in dry but fast-developing Australia and the USA’s Southwest: (1) the 530 km-long “Goldfields Water Supply Scheme” in Western Australia, started in 1896, was completed by 1903 and (2) the “First Los Angeles Aqueduct” was laid over a distance of 358 km between 1905 and 1913.

The polar water-ice deposits will be melted (by using a local source of heat or under an artificially dense Mars atmosphere) and the liquid water will be collected in large reservoirs. To be more specific, we envision a pipeline extending from Mars’s North pole region to a large human settlement near the Equator.

Fig. 26.6 Water pipeline route



The topography of the Mars surface has been taken into account and we have finally decided that the coordinates of the pipeline starting point are 8.5 W, 79.28 N while the coordinates of the end point are 43 W, 20 N (Fig. 26.6). The criteria that stood in making our choice were (a) an as small as possible water head and (2) an as short as possible route. However, the total length of the route is quite large: $L = 4023$ km.

26.3.3.1 Water Pumping Power

The pumping power p_{pump} required to move the freshwater in the pipe is given by:

$$P_{pump} = \dot{V} \Delta p_{pipe} \quad (26.25)$$

where \dot{V} is the volumetric flow rate, Δp_{pipe} is the linear pressure drop along the pipe (i.e. $\Delta p_{pipe} \equiv p_{inlet} - p_{outlet}$, where p_{inlet} and p_{outlet} are the water pressure at the inlet and outlet of the pipe).

In order to assess the pressure drop along the pipe, we have chosen to simulate the water flow in the duct based on the *incompressible* isothermal Reynolds averaged Navier-Stokes equations. Steady-state is accepted here. The scalable k-epsilon scheme has been used as closure turbulence model. A few details about the model are given next. The continuity equation is:

$$\frac{\partial}{\partial x_j}(\bar{v}_j) = 0 \quad (26.26)$$

where \bar{v}_j means Reynolds average, $x_j (j = 1, 2, 3)$ are spatial coordinates and $v_j (j = 1, 2, 3)$ are water velocity components. The Einstein summation rule is accepted here and in the following. The conservation of momentum reads:

$$\bar{\rho} \frac{\partial}{\partial x_j}(\bar{v}_j \bar{v}_i) = -\frac{\partial \bar{p}}{\partial x_i} + \frac{\partial}{\partial x_j}(\bar{t}_{ij} - \bar{\rho} \bar{v}'_i \bar{v}'_j) + \bar{\rho} g_i \quad (26.27)$$

where g_i are the components of the gravitational acceleration, $\bar{\rho}$ and \bar{p} are water density and pressure, respectively. The primes denotes fluctuating quantities. The viscous and turbulent stresses are defined as:

$$\begin{aligned} \bar{t}_{ij} &\equiv \mu_L \left(\frac{\partial \bar{v}_i}{\partial x_j} + \frac{\partial \bar{v}_j}{\partial x_i} \right), \\ \bar{t}_{ij} &\equiv -\bar{\rho} \bar{v}'_i \bar{v}'_j = \mu_t \left(\frac{\partial \bar{v}_i}{\partial x_j} + \frac{\partial \bar{v}_j}{\partial x_i} \right) - \frac{2}{3} \bar{\rho} \bar{k} \delta \\ \mu_t &\equiv C_\mu \bar{\rho} \frac{\bar{k}^2}{\bar{\epsilon}} \end{aligned} \quad (26.28-30)$$

where μ_L is laminar dynamic viscosity, μ_t is the turbulent dynamic viscosity and δ_{ij} is Kronecker's symbol. The transport equations for turbulent kinetic energy, k and turbulent dissipation, ϵ , are:

$$\bar{\rho} \frac{\partial}{\partial x_j}(\bar{v}_j \bar{k}) = P_k - \bar{\rho} \bar{\epsilon} + \frac{\partial}{\partial x_j} \left[\left(\mu_L + \frac{\mu_t}{\sigma_k} \right) \frac{\partial \bar{k}}{\partial x_j} \right], \quad (26.31)$$

$$\bar{\rho} \frac{\partial}{\partial x_j}(\bar{v}_j \bar{\epsilon}) = C_{\epsilon 1} \frac{\bar{\epsilon}}{\bar{k}} P_k - C_{\epsilon 2} \frac{\bar{\rho} \bar{\epsilon}^2}{\bar{k}} + \frac{\partial}{\partial x_j} \left[\left(\mu_L + \frac{\mu_t}{\sigma_\epsilon} \right) \frac{\partial \bar{\epsilon}}{\partial x_j} \right] \quad (26.32)$$

where σ_k and σ_ϵ are fitting coefficients which are determined by adjusting the results to experimental data and P_k is the turbulent kinetic energy production term, defined by:

$$P_k = -\bar{\rho} \bar{v}'_i \bar{v}'_j \frac{\partial \bar{v}_i}{\partial x_j} \quad (26.33)$$

A linear segment of pipe of length $L_{\text{sample}}=200$ m has been discretized in order to model the pressure drop and water averaged velocity. The pressure drop for the whole pipe was simply obtained by multiplication of the pressure drop for that segment by the number of such segments the pipe of length L would contain.

Boundary conditions comprise inlet turbulent intensity and mass flow rate and null area averaged relative pressure at outlet (i.e. $p_{\text{outlet}}=0$).

The standard model coefficients (Wilcox 1993) have been adopted in calculations

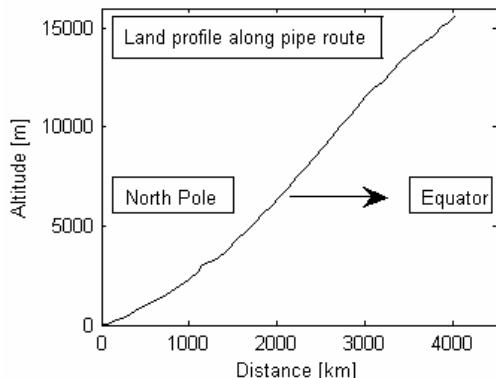
$$\sigma_k = 1, \sigma_\varepsilon = 1.3, C_\mu = 0.09, C_{\varepsilon 1} = 1.44, C_{\varepsilon 2} = 1.92 \quad (26.34)$$

These values are appropriate for a large number of fully developed turbulent flows including this one, where Reynolds number is over 10^6 . Simulation has been performed with a custom designed 3D finite volume Navier-Stokes code (Ferziger 1995).

Our discretization procedure took into account an average variation of the z -coordinate of 3.86 m/km according to the Mars altitude profile represented in Fig. 26.7 that was extracted from the topographical map illustrated in Fig. 26.6 (PDS 2009).

The life-sustaining requirements for an adult worker on Earth may total 8000 kg for 365 days, of which 7629 kg is drinking water, sanitary water and domestic water (Potter 2000). Mars has a year of 669 days, each Martian day with 24.65 terrestrial hours; every adult worker will need ~ 13380 kg of water per Mars year. For the envisaged 3 million people in the Mars colony this gives a water mass flow rate $\dot{m} = 680$ kg/s. A water density $\bar{\rho} = 995$ kg/m³ and an isothermal flow at an average of 5°C temperature have been considered during all calculations. The corresponding averaged velocity of the fresh water is approximately 1 m/s for a pipe diameter of 1m. The pressure loss for the sample pipe has been evaluated at $\Delta p = 18.942$ Pa/m. This means 76.2 MPa for the whole pipeline.

Fig. 26.7 Ground level altitude profile along pipeline route shown in Fig. 26.6



This result may be compared with results obtained for a pipeline of similar topology placed on Earth, where the gravitation acceleration is larger. On Earth the pressure loss for the sample pipe has been evaluated by the authors at $\Delta p = 42.53 \text{ Pa/m}$.

26.3.3.2 Maximum Distance between Pumping Stations

This section applies to a single pipe, operating alone or in a bundle of pipes. The meridian tension σ_1 and the circumferential tension σ_2 in a thin wall revolution shell are related to the pressure p_{int} inside by Buzdugan et al. (1972 p. 494):

$$\frac{\sigma_1}{R_1} + \frac{\sigma_2}{R_2} = \frac{p_{\text{int}}}{\delta} \quad (26.35)$$

where R_1, R_2 and δ are the main curvature radii and wall thickness, respectively. Equation (26.35) applies in case the pressure outside the shell is very low as compared to p_{int} . For a cylindrical pipe of inner diameter D_{int} one has $R_1 \rightarrow \infty$ and use of Eq. (26.35) yields:

$$p_{\text{int}} = 2\sigma_2 \frac{\delta}{D_{\text{int}}} \quad (26.36)$$

Fresh water motion in the pipeline is ensured by a number of pumping stations placed at equal distance L_{st} . The value of the pumping pressure at the exit of a pumping station is denoted p_{in} . This value must be large enough to cover the pressure losses due to friction during the water movement towards the next pumping station. Thus, along the pipe segment between two stations the pumping pressure p decreases due to friction from p_{int} to zero. Once a pipe material is chosen, the value of σ_2 may be estimated. Also, δ and D_{int} are design parameter. Considering $\sigma_2 \equiv \sigma_y$, where σ_y is the yield strength for steel, one derives:

$$p_{\text{int}} = 2\sigma_y \frac{\delta}{D_{\text{int}}} \quad (26.37)$$

The distance between two consecutive pumping stations comes from dividing the pressure head at pump exit to the pressure loss per unit length computed from the sample pipe flow simulation.

$$L_{st} = \frac{p_{\text{int}}}{\Delta p} \quad (26.38)$$

Usage of Eqs. (26.37) and (26.38) and the reasonable assumption that the altitude profile is almost linear (see Fig. 26.7) yield the number of pumping stations:

$$n_{PS} = \frac{L}{L_{st}} \quad (26.39)$$

For the present case, we considered $\sigma_y = 350 \text{ N/mm}^2$ and the steel pipe wall thickness $\delta = 8 \text{ mm}$ (Lui 1999 p. 225). The average pumping station pressure is $p_{\text{int}} = 57.6 \text{ bar}$ and this gives the distance between pumping station $L_{st} = 304 \text{ km}$ and the number of stations $n_{PS} = 14$. The pumping power per one station is $P_{st} = 3917 \text{ kW}$. This gives an input power of 54.8 MW needed to move the freshwater in the whole pipeline. An early calculation based on simple hydraulic formulae was performed in Houdsen (1914 p. 65-69). The results there were strongly affected by the diminished admissible water pressure at pipe inlet ($\sim 10 \text{ bar}$); this led to approximately 56 km between pumping stations, which is about 1/6 of the distance we have evaluated here, taking into account the performance of present-day materials.

Our results obtained for Mars are compared with results obtained for a pipeline of similar topology placed on Earth. The distance between pumping station is $L_{st} = 135.4 \text{ km}$ and the number of stations $n_{PS} = 30$. The pumping power per station is $P_{st} = 3916 \text{ kW}$ on Earth, which is very close to the value found for Mars. The explanation of this “coincidence” is simple: the pipeline segment between two adjacent pumping stations was designed to obey the same maximum inlet pressure (i.e. 57.6 bar), because the pipe wall material is the same in both cases (i.e. steel). However, the total input power needed to move the freshwater in the whole pipeline is 117.4 MW on Earth, because the number of pumping stations is larger on Earth than on Mars.

26.4 Conclusions

If humans are to settle Mars permanently, then they have to gather, route and consume freshwater derived from local hydrologic resources. Mars ecopoiesis seems to be a pre-requisite if large scale utilization of water resources is considered. This chapter provides basic information about Mars intrinsic ecopoiesis and water transportation by long pipelines under Martian conditions.

Increasing the ground surface thermal inertia seems to be necessary for Mars intrinsic ecopoiesis. This can be done either by removing the regolith layer covering the bedrock or by soil compression. Taking into account the thermal inertia of the Martian soil gives a more accurate idea about the perspectives of Earth-like life on Mars.

Due to the day/night alternation the ground surface temperature is normally below its steady-state temperature during daytime and above this temperature during nighttime. The amplitude of the daily excursion of ground temperature does not exceed 80-90 °C in case of a terraformed Mars covered by regolith, but decreases to less than 20 °C, when the basalt bedrock is considered. In the latter case the ground temperature on a large part of Mars surface has reasonable low and high temperatures for most of the year. After terraforming the Northern hemisphere seems to be more hospitable than the Southern hemisphere, because the amplitude of the daily temperature excursion there is lower and the freezing point appears at higher latitudes.

Our results suggest that a regional (and seasonal) ecopoiesis/terraforming of Mars is more realistic than a global scale project, in agreement with the early study of Averner and MacElroy (1976). These authors were rather conservative, particularly concerning the instability of liquid water *anywhere* on the Mars surface. They concluded ‘It is not a strong possibility that even a massive engineering project could trigger the conversion of *all* of Mars into a more hospitable environmental state’. Our results are also in concordance with the early paper by McKay (1982), who stated “a terraformed Mars need not duplicate the Earth in every respect”.

A case study has been presented in Sec. 26.3.2. It refers to a pipeline transporting water from the North polar cap to a 3 million person Martian colony in the equatorial region. The 4023-km long pipeline requires an input power of 58.4 MW to transport 680 kg/s fresh water. A total of 14 power stations are necessary to sustain the transport.

Hydraulics is cheaper on Mars than on Earth. There is about 2.3 times less effort needed to pump water on Mars. This decrease in the required power is due to the difference in the gravitational acceleration. The influence of friction of the flow mainly depends on fluid speed and for similar flows friction has the same influence on both Mars and Earth.

References

- Averner, M.M., MacElroy, R.D.: On the habitability of Mars: An approach to planetary ecosynthesis. NASA SP-414 (1976)
- Badescu, V.: Regional and seasonal limitations for Mars intrinsic ecopoiesis. *Acta Astronautica* 56, 670–680 (2005)
- Billham, R., England, P.: Plateau ‘pop-up’ in the great 1897 Assam earthquake. *Nature* 410, 806–809 (2001)
- Birch, P.: Terraforming Mars Quickly. *J. British Interplanetary Soc.* 45, 331–340 (1992)
- Bostrom, N., Cirkovic, M.M.: Global Catastrophic Risks, p. 554. Oxford University Press, New York (2008)
- Buckingham, A.G.: System and Method for Irradiation of Planet Surface Areas, US Patent 3564253 (1971)
- Burns, J.A., Harwit, M.: Towards a more habitable Mars – or – the coming Martian spring. *Icarus* 19, 126–130 (1973)
- Buzdugan, G., Voinea, R., Beles, A., Mitescu, C., Petre, A., Blumenfeld, M., Constantinescu, I.: Culegere de probleme din Rezistenta Materialelor. Didactica si Pedagogica, Bucuresti (1972)
- Crisp, D.: Infrared radiative transfer in the dust-free Martian atmosphere. *J. Geophys. Res.* 95(B9), 14577–14588 (1990)
- Edgett, K.S., Christensen, P.R.: The particle size of Martian aeolian dunes. *J. Geophys. Res.* 96(B5), 22765–22776 (1991)
- Ferziger, J.H., Peric, M.: Computational Methods for Fluid Dynamics. Springer, Dordrecht (1995)
- Fleming, J.: A 1954 color painting of weather systems as viewed from a future satellite. *Bull. Am. Met. Soc.* 88, 1525–1527 (2007)

- Fogg, M.J.: A synergic approach to terraforming Mars. *J. British Interplanetary Soc.* 4, 315–329 (1992)
- Fogg, M.J.: Terraforming: Engineering planetary environments. SAE International, Warrendale (1995)
- French, H.: The development of periglacial geomorphology: 1-up to 1965. *Permafrost and Periglacial Processes* 14, 29–60 (2003)
- Gerstbach, G.: Mars Channel Observations 1877–90, compared with Modern Orbiter Data. *Publications of the Astronomical Observatory of Belgrade* 75, 347–354 (2003)
- Graham, J., Simpson, A., Crall, A., Jarnevich, C., Newman, G., Stohlgren, T.J.: Vision of a Cyberinfrastructure for Nonnative, Invasive Species Management. *Bio-Science* 58, 263–268 (2008)
- Hanslmeier, A.: Habitability and Cosmic Catastrophes, 270 pages. Springer, Dordrecht (2009)
- Hourdin, F., Forget, F., Talagrand, O.: The sensitivity of the Martian surface pressure and atmospheric mass budget to various parameters: A comparison between numerical simulations and Viking observations. *J. Geophys. Res.* 100(E3), 5501–5523 (1995)
- Housden, C.E.: The Riddle of Mars: The Planet, pp. 65–69. Longmans, Green and Co., London (1914)
- Hunt, H.F.: The functions of the Martian Canals. *Nature* 82, 69 (1909)
- Jones, B.W.: Mars before the Space Age. *Int. J. Astrobiology* 7, 143–155 (2008)
- Kauffman, G.J.: What if...the United States of America were based on watersheds? *Water Policy* 4, 57–68 (2002)
- Lane, K.M.D.: Geographers of Mars. *Isis: Int. Review History of Science and Its Cultural Influences* 96, 477–506 (2005)
- Lane, K.M.D.: Mapping the Mars Canal Mania: Cartographic Projection and the Creation of a Popular Icon. *Imago. Mundi*. 58, 198–211 (2006)
- Lee, S.W.: Viking Lander Meteorology and Atmospheric Opacity Data Set Archive. In: vol. VL-1001, Laboratory for Atmospheric and Space Physics, Campus Box 392, University of Colorado, CO 80309-0392 (1995)
- Lovelock, J.E., Allaby, M.: The Greening of Mars. Warner Brothers Inc., New York (1984)
- Lui, E.M.: Structural Steel Design. In: Fah, C.W. (ed.) *Structural Engineering Handbook*. CRC Press, Boca Raton (1999)
- Markley, R.: Dying Planet: Mars in Science and the Imagination. Duke University Press, Durham (2005)
- McKay, C.P.: Terraforming Mars. *J. British Interplanetary Soc.* 35, 417–433 (1982)
- McKay, C.P., Toon, O.B., Kasting, J.F.: Making Mars habitable. *Nature* 352, 489–496 (1991)
- Montanes, J.L.: *Hydraulic Canals: Design, Construction, Regulation and Maintenance*, 389 pages. Routledge, London (2005)
- PDS, Planetary Data System. Geosciences Node (2009), <http://pds-geosciences.wustl.edu/missions/mgs/megdr.html> (Accesses on December 15, 2008)
- Pollack, J.B., Haberle, R.M., Schaeffer, J., Lee, H.: Simulations of the general circulation of the Martian atmosphere. 1 Polar processes. *J. Geophys. Res.* 95(B2), 1447–1473 (1990)
- Potter, J.F.: Seeking a new home: some thoughts on the longer term trends in planetary environmental engineering. *The Environmentalist* 20, 191–194 (2000)
- Raznjevic, K.: Tabele si diagrame termodinamice. Editura Tehnica, Bucuresti (1978) (in Romanian)

- Sagan, C.: Planetary Engineering of Mars. *Icarus* 20, 513–514 (1973)
- Santee, M., Crisp, D.: Thermal structure and dust loading of the Martian atmosphere during late southern summer: Mariner 9 revisited. *J. Geophys. Res.* 98(E2), 3261–3279 (1993)
- Sherlock, R.L.: Man as a Geological Agent: an Account of his Action on Inanimate Nature, p. 347. H.F. & G. Witherby, London (1922)
- Smil, V.: Global Catastrophes and Trends: the Next Fifty Years, 307 pages. MIT Press, London (2008)
- Trofimov, V.T., Averkina, T.I.: Engineering Geological Structures of the Earth. *Earth Science Frontiers* 14, 257–267 (2007)
- Vernon, J.: Hunger: A Modern History, 369 pages. Harvard University Press, Cambridge (2007)
- Wilcox, D.C.: Turbulence Models for CFD. DCW Industries Inc., La Canada, CA (1993)
- Zahnle, K.: Decline and fall of the Martian empire. *Nature* 412, 209–213 (2001)
- Zubrin, R.M., McKay, C.P.: Technological requirements for terraforming Mars. *J. British Interplanetary Soc.* 50, 83–92 (1997)

Author Index

- Antol, Jeffrey 401
Badescu, Viorel 25, 67, 99, 661
Bar-Cohen, Yoseph 431
Bearman, Greg 401
Behar, Alberto 401
Berggren, Mark 563
Bolonkin, Alexander A. 287, 599
Boston, Penelope 401
Buehler, Martin 401

Canessa, Enrique 587
Carrera, Stacy 563
Cathcart, Richard B. 661
Coleman, Max 401
Cooper, Chase 351
Cory, Joseph 83
Crawford, Ronald 401
Crossman, Frank 369

Fraser, Simon D. 1, 139
Gurfil, Pini 83

Hajos, Gregory 401
Hashimoto, Hirofumi 517
Herndon, J. Marvin 551
Hoffman, Jeffrey 351
Hofstetter, Wilfried 351
Homnick, Mark S. 369

Ishikawa, Yoji 543
Isvoranu, Dragos 661
Jones, Jack 401

Kelliher, Warren 401
Kuhlman, Kimberly R. 401

Maise, George 627
McInnes, Colin R. 645
Mikouchi, Takashi 483
Miley, George H. 213
Morgan, Paul 331

Palaia, Joseph E. 369
Paniagua, John 627
Popa-Simil, Liviu 175, 241
Powell, James 627
Powell, Jesse 627

Rapp, Donald 461
Rice, Eric 213
Rose, Heather 563
Rothschild, Lynn 401

Silver, Matthew 351
Stimpson, Alexander 369

Takeda, Hiroshi 483
Truett, John 369

Wada, Hidenori 517
Wilson, Cherie 563
Wilson, Daniel W. 401

Yamashita, Masamichi 517
Yang, Xiaoling 213
Yazawa, Yuuki 483

Zacny, Kris 431
Zubrin, Robert 563

Index

- 2001 Mars Odyssey 494, 496
 ^{210}Po 176
 ^{238}Pu 176
4Frontiers Corporation 374
- AB DOME 618
AB plasma beam reflector 310
AB plasma reflector 293
AB spherical reflector 294
abort to orbit 463
absorber 99
absorbed doses 206
AB-Space Propulsion 310, 311
accidents 265
agriculture 372, 483, 484, 491, 499, 500, 502, 503, 504, 505, 508, 509
agro-ecosystems 518
air 132
albedo 276
algae 530, 629
ALH84001 487, 490, 492
Alkali-Metal Thermoelectric Converter (AMTEC) 189
Alkaline Fuel Cells 355
allometric human growth 597
ALPH 628
ALSD 439
alumina 564
aluminum 380, 383, 388, 389, 545, 631
amenity 520
American West 388
amino acid score 521
ammonia 664
anaerobic condition 529
aneutronic reactions 297
anorthosite 506
- antimatter annihilation 280
Apollo fuel cell 151
Apollo Lunar Surface Drill 439, 440
Apollo 436, 439, 440
aquaculture 526
aquatic fern 527
aqueous alteration 486, 491, 494, 495, 497
arbuscular mycorrhizal fungi 528
Architecture Decisions Graph 353
areocentric longitude 27
artificial light 381
artificial gravity 249
artificial inflatable Dome 603
artificial lighting 383
artificial magnetic field 606
artificial seawater 531
ascent 462, 463, 464, 466, 472
asteroid 393, 394 646
astrobiology 499, 535
atmosphere of Mars 473
atmosphere 549
atmospheric pressure 34
atmospheric temperature 34
autoclavin 529
autocorrelation coefficients 58
autocorrelation function 56
autoregressive (AR) models 53
auto-regressive integrated moving average (ARIMA) models 53
auto-regressive integrated moving average 57
autumn 43
avian egg 525
azimuth angle 31
Azolla 534

- backward shift operator 60
 bacteria 364
 balloon 83
 bamboo 560
 barter economy 380, 381
 basalt 546, 669
 battery 273
 benefits of ISRU 462
 bentonite 546
 beta emitters 274
 BHA 448
 bi-directional trade 381
 binary power plant 341
 bio-electric 364
 biological combustion 527
 biological parameters 593
 biomass 518
 biometric formulae 592
 biometric relations 591
 bio-regenerative life support system 517
 bioremediation 531
 bio-safety 529
 biosphere 523
 blackbody receiver 101
 blackness 611
 bladder 531
 blueberry 497
 body blackness 610
 body mass index 588
 boil-off 473
 Boltzmann law 610
 Boltzmann's constant 298
 Bottom Hole Assembly 448
 braking radiation 293
 Brayton Cycle 188
 Brayton engine 355
 Brayton 244
break-even 290
 Bremsstrahlung radiation 293
 Bremsstrahlung 293
 brick 544
 building of a film dome 616
 buoyancy 84
 burnup 264

 cable space launch 304
 calcareous algae 533
 calcia 564
 campaign 470 471

 canals 672
 capillary force 534
 capital city 674
 carbon dioxide 132, 375, 506, 507, 635,
 664
 carbon nano-tubes 261, 364
 carbon, 390
 carbonate 490, 491, 492, 494, 495, 498,
 500, 507
 carbonyl, 560
 carbothermal process 468
 Carnot cycle 106
 cast basalt 545
 CAT 446, 447
 cation exchange capability, 534
 cell walls 523
 cellulose 525
 cement 389
 ceramics 563
 Ceres 390, 392
 cer-liq-mesh, 258
 CFC 664
 Chambadal-Novikov-Curzon-Ahlborn
 engine 102
 chassignite 485, 486, 487, 490
 chemical rockets 303
 chemical weathering 502, 503, 509
 cholesterol, 522
 circular orbit 464
 clay minerals 529, 545
 clay, 388
 Closed Ecology Experiment Facilities 536
 CO₂ acquisition 475
 CO₂ 473
 coating 243
 cogeneration 208
 coilgun 305
 cold press method 545
 collector tilt angle 107
 colonization 661
 color aberration 532
 column 564
 commodities 379
 commodity 380
complete_records 34
 composite 277
 Compost 527
 compressible fuel 261
 concentration systems 72
 concrete 380, 545, 563

- conductor-insulator, 273
- confinement 289
- construction material 543
- consumables 462, 464, 472
- Continuous Miner, 454
- controlled transparency 603
- convection heat losses coefficient 71
- cooling of dome 621
- cooling rate 488, 489
- Corer Abrader Tool 446
- cosmic radiation 543
- cosmic spherules 553
- cost 379
- Coulomb barrier 288
- cover damage, 613, 624
- cover film 617
- critical length 117
- critical mass 278
- criticality 276
- crop efficiency 521
- crustal thermal gradient 332
- cryogenic 270
- crystallization 563
- currency 380, 381, 388, 389, 399
- cyanobacteria 527

- daily direct irradiation 31
- dairy products 525
- DARPA 307
- DBT 433
- DEC 217, 219, 221, 222
- delta v, 390
- denitrification 528
- density of electric currency 610
- desalination 533
- descent 466
- Deuterium (D) 288, 296
- dietary fiber 521
- differencing procedure 55
- diffuse solar irradiance 31
- diffusion membranes 206
- diluted radiation 69, 103
- dilution factors, 104
- direct conversion 257
- direct drive 289
- dissolution rate 501, 502, 504
- diurnal thermal skin depth 668
- double film 603
- DPF 214

- Drill and Blast 453
- dry washers 553
- dry-steam power plant 341
- D-T 288
- ductile to brittle transition temperature 433
- dust storm periods 72
- dust storm season 49, 77, 100
- dwarf planet 390
- dynamic component 53
- dynamic pressure 612

- Earth microbes 182
- eccentric anomaly 28
- eccentricity of Mars orbit 28
- ECLSS 464, 465
- economic models 393
- ecopoiesis 661
- EDL 434, 435
- effective temperature 104
- effective transmittance - absorptance products 107, 114
- efficiency of AB-space propulsion 322
- efficiency of heater 611
- efficiency 241
- electric field intensity 291
- electricigens 364
- electrolyser 145
- electrolysis 392
- electrons 292
- electrostatic artificial gravity 607
- elliptical orbit 464
- energy intake 522
- energy
 - consumption per capita 370, 371, 372
 - demand, 374 375
 - dependence 399
 - of human endeavor 387
 - thermal 376
 - significance of 370
- energy independence 378, 379, 399
- energy systems
 - scale of 374
- Engineering Trade Analysis 381
- equivalent energy 381
- equivalent energy cost 380
- energy storage 304, 353, 354
- energy 483, 500, 506, 508

- engineered geothermal systems 331
entry, descent and landing 434
EPRI 216
ESAS 462
eutrophic coastal waters. 530
EVA 442
Evergreen dome 603
excess pressure 53
exergy of solar radiation 101
exergy of thermal radiation 101
exploration
 energy requirements 374
exploration vehicles 177
extra vehicular activity 442
extraction of oxygen from regolith 468
extraction 563
extrinsic techniques 661
- farming area 521
fast reactor 244
feces 527
feldspar 489, 494, 501, 502, 533
fermentation 527
fertilizer 528
filling station 466
fire safety 520
fish farming 629
fish 522
fissile material 180
fission product 262
flat plate solar collectors 32, 100
flowering plants 524
flyers 182
fog 32
food web 525
FRC 214
free flying 606
free market 381
freezing point 669
fuel cell 139
 alkaline Fuel Cell (AFC) 143
 applications 161
 configurations 149
 fundamentals 140
 history 151
 molten Carbonate Fuel Cell (MCFC)
 143
 phosphoric Acid Fuel Cell (PAFC) 143
 polymer Electrolyte Membrane Fuel Cell
 (PEMFC) 143
 solid Oxide Fuel Cell (SOFC) 143
- fuel dispersion 259
fully concentrated scattered radiation 105
fulvic acid 502, 503, 504, 505, 508, 509
funeral service 525
fusion barrier energy 288
fusion power plant 287, 290
fusion reactions 288
fusion 278
- galactic cosmic radiation 549
gallium arsenide 69
gas production 376
Gaussian white noise 58
GDP 394, 396, 398
gear ratio 463
GEKKO XII laser 289
Gemini fuel cell 151
general circulation models 52
general purpose heat sources 355
general radiation heat 611
generation 258
GEO 392
Geobacter 536
geometric factor 102
geothermal heat pumps 333
geothermal reservoir 331
geothermal resource 333
geothermal 379
germination 520
glass composite 545
glass fiber 545
glass 545, 563
global dust storms 31
global solar irradiance 31
global solar radiation 25
global temperature distribution 665
Gopher 448, 449
graphene 364
gravitropism 587
gravity of planet 610
gray thermal radiation 103
green manure 527
greenhouse gases opacity 663
greenhouse gases 661
greenhouse 381, 384, 385
ground ice 470, 475
growth charts 588, 593
growth velocity 588
Gusev Crater 496, 497, 498, 500

- habitat 543
 habitation on Mars 518
 halophytes 530, 531
 hawkmoth 524
 heat balance 611
 heat pipe 180, 193, 251
 heat production rate 334
 heat regeneration factor 130
 heat storage system 106
 heat transfer coefficient 610
 heat transferring 611
 heating of the dome 621
 Heat-pipe Power System (HPS) 254
 helium 84, 99, 128, 132, 243
 Hellas Planitia 601
 hematite 388, 494, 496, 497, 508, 533, 553
 hemispherical blackbody radiation source 101
 hibernation 671
 high-energy particles 606
 higher hydrocarbons 479
 HiPER 289
 Hohmann transfer 390
 HOMER 178, 252
 hornblende 533
 hornworm 524
 hour angle 29
 human development index 372
 human diet 519
 human growth 588, 593
 human mission 248
 human productivity 520
 HVAC system 376
 hydrated minerals 375
 hydrogen bomb 289
 hydrogen reduction 467
 hydrogen 390, 466, 467, 468, 470, 471, 473, 474, 478, 479, 507, 629
 hydroponics 614
 hyperbolic velocity 390
 hyperthermophilic aerobic bacteria 527
 hypobaric conditions 523
 hypogravity 523
 ice plant 531
 ice 388
 ICF 213, 288
 Icy Soil Acquisition Device 437, 439
 iddingsite 486, 491, 500
 ideal solar cells 73
 IEC 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 228, 229, 233, 234, 235, 236
 IMLEO 461, 464
 impulse 389
 incineration 529
 indirect drive 289
 industrial processes 372
 inertial confinement fusion 288
 inflatable space hotel 607
 inflated domes 602
 inflated film dome 616
 inflated textile shell 603
 inner solar system 393, 399
 insects 519
 in-situ propellant production 359
 in-situ resource utilization 154, 431, 456
 In-Situ Resource Utilization 544
 internal combustion engine 388
 internal heat budget 333
 International Space Station 431, 441, 442
 interplanetary transfer vehicle 186
 intrinsic techniques 661
 ion engine 390
 ionization 273
 iron carbide 560
 iron oxide 546
 iron pentacarbonyl 557
 iron 389, 545, 551, 563, 631
 irrigation 613
 ISAD 437
 ISRU 431, 456, 544
 ISS 431, 441
 ITER 288
 jarosite 492
 JET 287
 KE Reactor 242
 Kevlar 354
 kinetic thrust 326
 KIWI 177
 Kolmogorov-Smirnov test 59
 La Grange Point 395
 labor 379
 Lagrange function 113
 Laird-Gompertz empirical formula 596
 Laird-Gompertz formula 588

- landing sites 46
 laser beams 289
 latent heat storage 100
 launch vehicles 463
 launch 249
 lava 338
 Lawson criterion 290
 leaching 533
 LEDs 384, 385
 LENR 213, 229, 230, 231, 232, 233, 235
 LEO 392
 life support 464, 471, 472
 light energy 381
 Li-ion batteries 354
 limestone 556
 linear electrostatic engine 304
 lipid 525
 liquid droplet 194
 liquid metal 187
 lithium 197
 LLNL 288
 loach fish 526
 loamy sand 534
 local dust storm 25
Luna. See Moon *See Moon*
 Lunar Reconnaissance Orbiter 470, 471
 lunar resources 467

 macroprojects for Mars 612
 magma 485, 486, 488, 490, 506
 magnesia 564
 magnesium sulfate 563
 magnetic confinement fusion 288
 magnetic constant 609
 magnetic control mirror 605
 magnetic fusion 186
 magnetic intensity 312, 609
 magnetic levitation 605
 magnetic pressure 312
 magnetic thrust of cables 326
 magnetite 553
 major dust storm 79
 mantle convection 335
 Mars Ascent Vehicle (MAV) 463, 464
 Mars atmosphere 601
 Mars colonization 213, 236, 225, 587
 Mars Colony 636, 674
 Mars Direct 374, 464
 Mars Exploration Rovers 176, 434, 446, 494, 496
 Mars exploration 84
 Mars Express 494, 500, 601
 Mars Global Surveyor 494, 496, 601
 Mars' gravity 601
 Mars Hopper 630
 Mars Ice Cap Explorer 630
 Mars orbit 390
 Mars Pathfinder 496
 Mars Reconnaissance Orbiter 495
 Mars Science Laboratory (MSL) 182
 Mars settlement 374, 379, 387, 388, 389, 393, 399
 Mars settlers 375, 387, 395
 Mars surface energy pathways 18
 Mars surface power system options 17
 Mars-bound vehicle 465
 Mars-Earth Rapid Interplanetary Transport 639
 Mars exploration 1
 1960-1990 1
 1990-2000 4
 2000-2010 8
 2010+ 12
 marsquakes 338
 Martian day 27
 martian geothermal resources 331
 Martian ice 183
 Martian ionosphere 601
 Martian meteorites 484, 487, 490, 496, 498, 499, 500, 502
 Martian seasons 27, 602
 Martian surface temperatures 602
 Martian surface 600
 maskelynite 487, 488, 490
 mass density of reflector 610
 mass independence 379
 material independence 399
 materials 196
 MATLAB 385
 maturation of seeds 520
 Max Planck 619
 maximum concentration ratio 105
 maximum likelihood method 58
 Maxwellian distribution 292, 298
 MCF 288
 medical requirements 520
 medium of exchange 380
 MER 434, 446, 447
 mercury 99, 111
 Meridiani 494, 496, 497, 500, 508

- meteorites 553
 methane hydrate 506
 methane 388, 395, 463, 479, 483, 500,
 503, 505, 506, 629, 664
 methanol 629
 MHD conversion 191
 MHD 229
 microbes 528
 microeconomics 389
 micro-hetero structured 257
 micrometeorite 543
 microorganism 499, 500, 501, 505
 migration 671
 mineralogy 483, 484, 487, 488, 489, 490,
 494, 495, 496, 497, 498, 501, 502, 508
 mining 388, 450, 552
 MIT study 464
 mixtures of noble gases 111
 model diet 525
 modular 252
 molybdenum 196
 monmorillonite 500, 548
 Moon 392, 393, 394, *See*
 moons of Mars 602
 more than 100% closure 519
 moving-average coefficients 58
 mulberry 522, 524
 multi-junction solar cells 364
 multi-layer insulation (MLI) 473
 Multi-Mission RTG 179
 multiply scattered solar radiation 104
 mushroom 523
- nakhlite 485, 486, 487, 489, 490, 491
 nano-cluster-sinter 257
 nano-materials 364
 nano-structured radiation guides 258
 nanotube 293
 NASA Design Reference Mission 374
 NASA Human Exploration Initiative 462
 NASA plans 461
 NASA Science Mission Directorate 354
 NASA 182
 National Ignition Facility 288
 Navier-Stokes equations 675
 needed thickness 608
 nematodes 528
 NERVA 177
 net transparency 299
 neutron spectrometer 470, 471, 474
- neutron 290
 New Zealand spinach 531
 Newton relationship 129
 Ni-Cd batteries 100
 nickel carbonyl 558
 NIF 288
 nitrogen fixation 525, 535
 nitrogen 384, 390
 nontronite 495, 500
 North Polar Cap 628
 NOVA 288
 Nuclear Electric Propulsion (NEP) 177, 185
 nuclear particles 270
 nuclear power 176
 nuclear thermal propulsion 177 627
 nuclear waste 264
 nutrient 499, 502, 505, 508, 509
 nutritional content 521
- Object-Process Methodology 352
 obscuration rate 79
 olivine 487, 488, 489, 490, 491, 494, 495,
 500, 533
 Olympus Mons 600
 OMEGA 289, 494, 495, 500
 open loop life support 517
Opportunity rover 602
 optical depth 25, 30, 78, 118
 optimum azimuth angle 32
 optimum tilt and orientation 72
 optimum tilt angle 26, 35, 72
 orbit 250, 646
 orbital mechanics 390
 organic acid 501, 502, 503, 504, 508, 509
 organic chemicals 390
 organic fluids 111
 organic life 587
 organic matter 535
 organic waste 535
 organo-metallic compound 557
 orographic component 53
 outer space hotel 607
 outpost 178, 466
 overall heat losses coefficient 108
 oxygen 384, 388, 389, 463, 464, 466, 467,
 468, 469, 470, 472, 479, 563, 629
- parabolic concentrator 99
 parasites 527
 parsimony principle 58
 partial heat regeneration 128

- Pathfinder 114
 pathogens 527
 pearlite 560
 peat moss 533
 Pellet Bed Reactor System 243
 pellet 289
 penetration length 668
 perfect burning 265
 performance metrics 356
 perihelion 49
 periodogram 56
 permanent settlement 374
 permanently shadowed craters 467
 P-F 230
 pH 563
 Phoenix Lander 600
 Phoenix Mars Lander 601
 Phoenix 498
 phosphate 528, 534
 photon fuel cell 153
 photosynthetic bacteria 533
 photosynthetic plants 518
 photovoltaic energy conversion 68
 photovoltaic 83, 181, 354
 phyllosilicate 494, 495, 499, 509, 545
 phylogenetic tree 526
 physiological activities 504
 pipeline 677
 plagioclase 486, 488, 489, 490, 492, 502,
 503, 504, 506, 507
 planetary drills 439
 planetary protection 535
 planetary scoops 436
 plant fertilization 502
 plant roots 528
 plant's defenses 524
 plasma beam 307
 plasma cable 307
 plasma cord 307
 plasma reflector 291
 plasma temperature 291
 plastic 384
 plate tectonics 340
 poisoning 260
 polar cap recessions 54
 polar caps 601
 polar ice deposits 469
 polar outpost 463
 polar temperature 665
 political independance 379, 399
 political issues 389
 pollination 524
 polycarbonate 381, 383, 384
 polymers 380, 390
 polyvinyl alcohol 534
 potassium 111, 529
 power density 271
 power generation 353
 power management 248, 353
 power requirements 467, 470, 477, 468,
 472, 480
 power systems
 nuclear 374, 375
 power usage pattern 356
 precipitation 563
 pressure 431, 432, 433, 434, 435, 441,
 450, 451
 Prometheus 355
 propellant requirements 464
 propellants 462, 463, 464, 465, 466, 472,
 473, 479
 propulsion 462, 463, 479
 prospecting 470
 proton exchange membrane 355
 pulsed thermal-to-electricity converter 267
 pumping stations 678
 pupa 524
 PV array 68
p-value 58
 PWR 221, 222
 pyroxene 487, 488, 490, 494, 501, 509, 533

 Q 216, 219
 QUID 381, 388, 392

 radiation activity 606
 radiation belts 293
 radiation heat flow 610
 radiation shielding 462, 466
 radiation 279
 radiogenic heat producing elements 334
 radioisotope heat sources 358
 radioisotope heater units 432
 radioisotope heating units 178
 radioisotope thermal generators 470
 Radioisotope Thermoelectric Generators
 (RTGs) 176, 178, 434
 radioisotopes 273, 295
 radiolysis 206

- railgun 305
- rainfall 673
- rains 669
- Rankine cycle 186, 187
- RAT 446
- reaction rates 298
- reactor concept 257
- reactor coolant 177
- recombination efficiency 71
- Red Planet 291
- re-entry shield 181
- reflected solar irradiance 31
- reflector 606
- reflectors 276
- refreshment 523
- regenerative fuel cells 149, 352, 355
- regolith simulant 546
- regolith 206, 389, 435, 439, 440, 441, 446, 491, 495, 499, 502, 505, 507, 533, 543, 668
- restart capability 248
- reverse water gas shift 476, 477, 478
- reversible engine 101
- RF 219
- RHU 432
- rice 521, 526
- road curve 177
- robot 388
- robustness 252
- rock abrasion tool 446
- rocket engine 304
- rocket 303
- rovers 177
- RTG 434
- rubidium 99, 111
- runaway greenhouse model 662
- Russia 183
- Sabatier/Electrolysis (S/E) process 479
- SAFE 255
- safety tensile stress 608
- safety 202
- salinity 529
- saltwort 531
- sample spectrum 57
- sandy loam 534
- saponite 495, 500
- science
 - energy requirements 374
- seasonal ARIMA models 63
- secondary power generation 354
- Seebeck effect 189
- selective flat-plate solar collector 106
- sheet silicate 494
- shelter 532
- shergottite 486, 487, 488, 489, 491
- shield 250
- shielding 206
- shock metamorphism 486
- silica 388
- silicon 69
- silkworm larvae 523
- silkworm 524
- sintered material 544
- small thermonuclear engines 291
- smectite 491, 495, 496, 500, 536, 545
- smelting 552
- SNC meteorites 485, 486, 487
- snowfall 673
- social issues 389
- sodium management 529
- sodium 519
- soil albedo 667
- soil bacteria 528
- soil 483, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 508, 509, 563
- Sojourner 114
- solar array 178
- solar concentrators 32
- solar constant 33
- solar declination angle 29
- solar dynamic power system 99
- solar energy 83, 379
- solar insolation 647
- solar irradiation 375
- solar panels 532
- solar power plants 106
- solar power 376
- solar radiation 605
- solar reflector 646
- solar wind radiation 606
- solar wind 467, 606
- solid oxide electrolysis 476
- solid state-compact nuclear 269
- sorption compressor 476
- Soybean 521
- SP-100 177
- space agriculture 531
- space economy 395, 396, 398
- Space Shuttle Orbiter fuel cell 152
- Space Station Freedom 100
- spectral density 53
- spherical source of radiation 102
- spheromak 186

- Spirulina* 533
 Spitzer conductivity 314
 spring 43
 SQUID 231 232
 standard of living 370, 372, 373
 stationary collector 73
 stationary process 54
 statistical equilibrium 53
 steel 380, 381, 383, 389, 563
 Stefan-Boltzmann constant 610
 stellarators 288
 Stirling cycle 189
 Stirling engine 99, 128, 180, 355
 Stirling 255
 stoichiometric relation 518
 storm wind 624
 strategies to collect solar radiation 43
 subsurface ice 474, 475
 sulfate 491, 492, 494, 495, 496, 498, 500, 508
 sulfuric acid 563
 summer 43
 sunlight 381, 382, 385
 sunset hour angle 29
 super-colliders 278
 superconductivity ground ring 607
 supply and demand 379
 support cable 608
 surface albedo 30
 surface emissivity 667
 surface gravity 587
 surface heat flow 333
 surface miner 454, 455
 survivability 521
 sustainable agriculture 529, 535
 sweet potato 521
 symbiotic protozoa 525
 synergic model 662
 synthetic data generation 27
 Syrtis Major 665
- tanker 466
 tantalum 197
 tectonism 333
 temperature electrons 325
 temperature 431, 432, 433, 434, 435, 441, 450
 tent-shaped array 68
 termites 523
- termocouple 99
 terraforming 645, 648, 661
 thermal conductivity 262
 thermal control 353
 thermal engine 99
 thermal fracturing 445
 thermal inertia parameter 668
 thermal photovoltaics 250
 thermal radiation 290
 thermal resistance method 107
 thermionic reactor 183
 thermionic 190
 thermionic-emitter 99
 thermodynamic cycles 99
 thermo-electric nuclear power 241
 thermoelectric 266
 thermonuclear AB-reactor 293
 thermonuclear ignition 288
 thin film rollout solar arrays 359
 thin film solar architectures 362
 thin-film amorphous silicon solar cells 354
 thorium 275
 tilapia 533
 timeline 472
 titanium 389
 tokamaks 288
 TOPAZ 177
 total economic activity 394
 tracking arrays 354
 tracking PV array 68
 trade 379, 399
 - inner solar system 389
 - Mars surface 388
 trade network 389
 transmutation 264, 274
 transparency 617
 transparent film 603
 transpiration ratio 518
 transportation 379, 388, 389, 392, 393
 tri-culture 526
 triple electrostatic mirror 307
 triple point 433, 435, 450, 451
 Tritium 288, 296
 TRL 213
 tropical temperature 665
 true anomaly angle 28
 Tsiolkovsky 390
t-statistics 58
 turbation 536
 turbo-generator 99
 TWDay/year 241

- ultra-light thin-film arrays 354
- Ultrasonic/Sonic Driller/Corer 448
- Ulva* 530
- unitised regenerative fuel cell 149
- Urey ratio 334
- urine 527
- USDC 448, 449
- vacuum solar collectors 118
- Valles Marineris 600
- value 461
- value 379
 - of data 394
- value of ISRU 461
- Van Allen Belt 292
- variance 53
- vault 548
- vertebrate meat 525
- vertisols 536
- Viking 437
- Viking Landers 26, 495
- virtual specific impulse 321
- vitamins 521
- volatile 469
- volatiles 469
- volcanism 333
- volume of living 543
- waste processing 376
- water content 474
- water deposits 473
- water electrolysis 642
- water excavated on the Moon 465
- water ice hazes 48
- water ice 469
- water in atmosphere 620
- water on Mars 474
- water vapor opacity 664
- water vapor 664
- water 377, 380, 388, 390, 545
- watering 528
- weather forecasting 63
- weight on bit 444, 446
 - weightlessness 587
- wet-steam power plant 341
- wheat 521
- wind power 379
- wind speed 33
- winds 25
- winter 43
- WOB 444
- Xanadu 630
- xenon 99
- zenith angle 29
- Zirconia 476, 477