# Predicting Thermal Power Consumption of the Mars Express Satellite with Machine Learning

**Conference Paper** · September 2017

DOI: 10.1109/SMC-IT.2017.22



**9 authors**, including:

Dragi Kocev
Jožef Stefan Institute

Jurica Levatić
IRB Barcelona Institute for Research in Biomedicine

Matej Petković
Jožef Stefan Institute

Nikola Simidjievski
Jožef Stefan Institute

# Predicting thermal power consumption of the Mars Express satellite with machine learning

Martin Breskvar[1,2], Dragi Kocev[1,2], Jurica Levatić[1,2],
Aljaž Osojnik[1,2], Matej Petković[1,2],
Nikola Simidjievski[1], Bernard Ženko[1]
e-mail: name.surname@ijs.si
Jožef Stefan Institute[1]
Jožef Stefan International Postgraduate School[2]
Jamova cesta 39, SI-1000, Ljubljana, Slovenia

Redouane Boumghar[3], Luke Lucas[4]
e-mail: redouane.boumghar@esa.int
e-mail: luke.lucas@lsespace.com
European Space Agency, ESOC[3]
Data Analytics Team for Operations[3]
Robert Bosch 5, 64293 Darmstadt, Germany[3]
Mars Express, Mission Planning & Spacecraft Operations[4]
LSE Space GmbH, Robert Bosch 7, 64293 Darmstadt, Germany[4]

***Abstract*—The thermal subsystem of the Mars Express (MEX) orbiter keeps the on-board equipment within its pre-defined operating temperatures range. To plan and optimize the scientific operations of MEX, its operators need to estimate in advance, as accurately as possible, the power consumption of the thermal subsystem. The residual power can then be allocated for scientific purposes. We present a machine learning-based pipeline for the prediction of MEX's thermal power consumption. We show that the proposed pipeline is superior in accuracy to the models currently used by MEX's operators. We also demonstrate that machine learning can provide the operators with insight about the orbiter's thermal behavior. Better understanding of the thermal subsystem and improved predictive accuracy of the thermal power consumption could help operators to improve science return and to prolong the operating life of MEX.**

## I. Introduction

Mars Express (MEX), a spacecraft operated by European Space Agency (ESA), has been orbiting Mars since the end of 2003. Its scientific payload consists of seven instruments that provide global coverage of Mars' surface, subsurface and atmosphere [1]. The instruments and on-board equipment of MEX have to be kept within their operating temperature ranges, which range from –180°C for some equipment to room temperature for others. To maintain operating temperatures, the orbiter is equipped with an autonomous thermal system composed of heaters and passive coolers.

MEX is powered by electricity provided either by its solar arrays or batteries, when the arrays are in shadow. The thermal system, together with the platform units, consumes a significant amount of the available power, while the remaining power is used for science operations.

The thermal power consumption (TPC) changes through time, depending on various external and internal factors, such as exposure of the orbiter to Sun or heat generated by the on-board equipment units. Predicting the power consumption of the thermal system is a non-trivial but crucial task, which allows the optimization of science operations of MEX.

To predict the TPC, the MEX's operators currently use a manually constructed model based on simplified physical models, expert knowledge and experience. However, due to aging of the spacecraft and decaying capacity of its batteries, power is a precious resource and every little bit saved in the thermal subsystem can be used for science acquisitions. This prompts the need for a more accurate predictive model of the TPC, which would also prolong the operating life of MEX.

Machine learning studies algorithms that can learn from and make predictions based on the data, and improve with experience [2]. Such algorithms can capture and describe patterns in complex data, making them a valuable asset for studying a variety of domains ranging from life and earth sciences to social and behavioral sciences. In the context of MEX's TPC prediction, given the inherent complexity of the task and large quantity of the available telemetry data, models derived with machine learning algorithms can offer a better solution than the models constructed manually by human experts.

This motivated ESA to organize its first data mining competition, the *Mars Express Power Challenge*[1] [3], with the goal of acquiring more accurate modeling approaches and models to predict MEX's TPC.

In this paper, we present the winning solution of the Mars Express Power Challenge submitted by the team from the Jožef Stefan Institute. In particular, we developed a pipeline for predicting the TPC of MEX, which includes data preprocessing and machine learning. Besides yielding more accurate predictions than the rest of the competing teams, the proposed pipeline is substantially more accurate than the predictive model currently in use at ESA. This suggests that machine learning algorithms can be used to support operators to gain more accurate knowledge on the thermal behavior of their satellite. This approach could also be extended to the some other subsystems of a spacecraft to further improve the operation of the orbiter.

The paper is organized as follows. Section II briefly describes the MEX Power Challenge. Next, Section III presents the proposed methodology, i.e., the data preprocessing and machine learning methods. Section IV presents the results and discusses possible implications for spacecraft operators. Finally, Section V concludes the paper.

[1]https://kelvins.esa.int/mars-express-power-challenge [accessed:2017-03]

## II. Mars Express Power Challenge

For the purpose of the challenge, ESA provided data consisting of raw telemetry data (context data) and measurements of the electric current of 33 thermal power lines (observation data), for the three Martian years of MEX operations. We refer to these data as the training set. For the fourth Martian year of the operation, only context data were provided. We refer to these data as the test set.

As mentioned, the raw data comprise two parts: context data and observation data. The *context data* consisted of five components:

**SAA** (Solar Aspect Angles) data contain the angles between the Sun–MEX line and the axes of the MEX's coordinate system.

**DMOP** (Detailed Mission Operations Plans) data contain the information about the execution of different subsystems' commands at a specific time.

**FTL** (Flight dynamics TimeLine events) data contain the pointing and action commands that impact the position of MEX, such as pointing the spacecraft towards Earth or Mars.

**EVTF** (Miscellaneous Events) data contain time intervals during which MEX was in Mars's shadow or records of the time points when the MEX is in apsis of its elliptical orbit.

**LTDATA** (Long Term Data) contain the Sun–Mars distances and the solar constant.

The *observation data* consisted of the 33 electric current measurements, recorded once or twice per minute. The observation data was given only for the duration of the three Martian years of MEX's operation, whereas the fourth year was not provided and was used for assessing the quality of the solutions submitted during the challenge.

All raw data entries are timestamped (expressed in milliseconds), indicating when the entry was logged. The time spans between the two consecutive entries varies from less than a minute (SAA) to several hours (LTDATA).

The task was to predict the values of the observation data for the fourth Martian year, using a machine learning model, learned from the training set. The predictions are to be given as the hourly average electric current for each of the power lines. A detailed description of the task and the data can be found on the challenge website.

The predictions were evaluated against the fourth Martian year's actual average electric current using the root mean square error (RMSE) measure, defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(c_{ij}-r_{ij})^2},$$

where $c_{ij}$ is the predicted value for the $i$-th time interval in the fourth Martian year of the $j$-th power line and $r_{ij}$ is the corresponding recorded value; $N$ is the total number of evaluated measurements ($N = 16488$) and $M$ is the number of power lines ($M = 33$). The goal was to minimize the RMSE score, i.e., lower RMSE indicates better results.

## III. Methodology

In the typical machine learning setting, the input to a learning algorithm is given in the form of a table, where each row of the table is an example and each column is a numerical (i.e., continuous) or nominal (i.e., discrete) feature describing the example. In the context of the MEX challenge, an example corresponds to a time interval, while features are derived from context and observation data.

We propose a pipeline for the prediction of the TPC of MEX. It consists of three main steps. First, the raw telemetry data is processed to construct a dataset as described above. Second, machine learning methods are applied to this dataset to learn a predictive model. Third, the model is used to predict the TPC on the test set. In the reminder of this section, each step of the proposed pipeline is described.

### A. Data preprocessing

The raw data could not be used directly for learning the model due to two main reasons: (i) incompatible time resolutions of the different components of raw data, and (ii) unstructured format of some of the entries, such as text, that machine learning algorithms do not understand. Therefore, a suitable preprocessing schema had to be designed. First, we selected the time resolution of the dataset, and then aggregated entries from the raw data to this resolution. Second, we constructed features informative of the TPC from the raw data. By combining information from different components of the context data, we aimed to construct more informative features to achieve better predictive performance.

As mentioned in Sec. II, the values of observation data and some of the raw context data (e.g., angles in SAA data) were given in 30 to 60 seconds intervals, while others were given at coarser resolutions (e.g., solar constant in LTDATA). In the latter case, the values in the data changed only gradually. Hence, we assumed that these values can be interpolated without a considerable loss of predictive performance.

We chose to use a one minute time interval as a resolution of our dataset. To align the timestamps, we divided the time span $[t_{\mathrm{FIRST}}, t_{\mathrm{LAST}}]$ into 1 minute subintervals. Here, $t_{\mathrm{FIRST}}$ and $t_{\mathrm{LAST}}$ are the first timestamp in the training part of the observation data, and the last timestamp in the test part of the observation data, respectively.

To compute the values of the measurements at the new timestamps, we used linear interpolation. Due to issues with spacecraft communication, some values were missing during some time periods both in the descriptive space (i.e., the constructed features from the raw data) as well as in the target space (the heater lines thermal consumption) of the data. In principle, the machine learning method used in this study can handle data with missing values. However, longer periods with contiguous missing values can substantially hurt the accuracy of the method as well as add an additional computational overhead. For this reason, we first removed the longer gaps i.e. the time periods longer than 10 minutes with missing values in the target space. Next, in the descriptive space, for time periods longer than 10 minutes, we interpolated the missing values, and left them unchanged otherwise.

In the following subsections, we give descriptions of the groups of features that we constructed in the preprocessing step of the pipeline.

*1) Energy influx features:* There are seven features in this group: one for the solar panels and one for each of the six sides of the cuboid of MEX. The features describe the amount of solar energy that is collected through a given surface in a given time interval $[t_i, t_{i+1})$. The solar energy collected by a side of the cuboid directly influences the temperature of the satellite, thus affecting the power consumption of the corresponding heaters. The solar energy collected by the solar panels influences the amount of energy that can be generated and used.

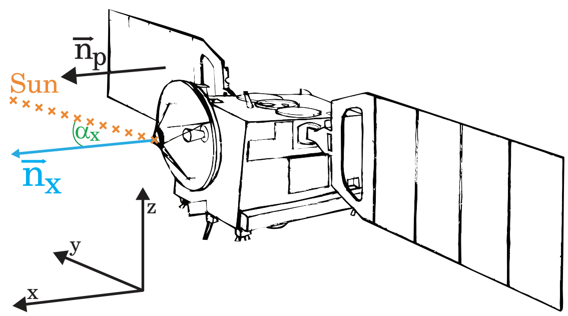


Fig. 1. Illustration of the MEX spacecraft and its coordinate axes $x$, $y$ and $z$, that correspond to *front*, *left* and *up* sides of MEX, respectively. $\alpha_x$ denotes the solar aspect angle of the front side, i.e., the angle between the normal $\vec{n}_x$ and the Sun-MEX line. $\vec{n}_p$ denotes the normal of the panels.

The amount of energy collected by a given surface is proportional to the product of the effective area $A_{\text{eff}}$ of the surface and the solar constant $c$. If the area $A$ is given, we compute $A_{\text{eff}}$ as $A_{\text{eff}} = A \max\{\cos\alpha, 0\}$, where $\alpha$ is the angle between the Sun–MEX line and the outer normal $\vec{n}$ to the surface (see Fig. 1). Without any loss of generality, we assume $A = 1$ for all surfaces, as the machine learning method that we use is invariant to monotonic transformations of features. The values of $\alpha$ for each of the seven surfaces were computed directly from the SAA data, while $c$ was given in LTDATA. Additionally, (pen)umbras have a considerable impact on energy influx, thus we also consider this information (extracted from EVTF data).

We define the energy influx $E_S^i$ for the surface $S$ at the time interval $[t_i, t_{i+1})$ as:

$$E_S^i = \int_{t_i}^{t_{i+1}} A_{\text{eff}}(t)c(t)U(t)\mathrm{d}t,$$

where $U$ is the *umbra coefficient*, an approximation of the proportion of Sun visible from the orbiter. $U$ takes the value $U(t) = 0$ if the orbiter is in an umbra, $U(t) = 0.5$ if the orbiter is in a penumbra, and $U(t) = 1$ otherwise. Instead of calculating exact integrals for $E_S^i$, we approximated the values using the trapezoid-rule.

*2) Historical energy influx features:* The thermal state of the satellite depends not only on the current energy influx, but also on the energy influx in the past. To capture this, we constructed historical energy influx features for each of the seven surfaces. A given historical feature for surface $S$ at time $t_i$ was computed as a sum of energy influx during the given historical time-frame, i.e.,

$$H_S^i = \sum_{j=1}^{N} E_S^{i-(j-1)},$$

where $N$ is the length of historical time-span considered. To account for different impacts of the historical energy influx we constructed historical features with different time-frames, namely, 4, 16, 32, 64 and 128 time intervals (minutes). This yields 35 features, 5 for each surface.

*3) DMOP features:* The DMOP files contain a log of commands issued to MEX's subsystems. The names of commands have been obfuscated, but the available documentation revealed two variants: (i) events that contain information about the subsystem and command that has been executed (e.g., ASXX383C), and (ii) events that represent flight dynamics events (e.g., MAPO.000005). The first 4 characters of variant (i) represent the subsystem, while the rest represent the command and its parameters. In variant (ii), the first 4 characters represent the name of the event, followed by an occurrence number. These events likely have different impacts on the thermal subsystem of the orbiter.

For subsystems' command that has been executed, we assumed that there were some delays before the thermal state of the orbiter was affected. Therefore, we constructed features that encode such a delayed effects in terms of "time since last activation" of a specific subsystem command from the raw DMOP data.

TABLE I
ILLUSTRATION OF THE DMOP FEATURES THAT ENCODE THE TIME SINCE LAST ACTIVATION OF A GIVEN SUBSYSTEM COMMAND.

| Raw data | | DMOP features | | | |
|---|---|---|---|---|---|
| t | Command | t | ASXX383C | ATTTF030A | ASXX303A |
| 1 | *none* | 1 | 0 | 0 | 0 |
| 2 | ASXX383C | 2 | 0 | 1 | 1 |
| 3 | ATTTF030A | 3 | 1 | 0 | 2 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 56 | ASXX303A | 56 | 54 | 53 | 0 |

The values of the DMOP features are calculated as follows:

$$f_k^i = \begin{cases} 0 & \text{if } k \text{ is activated at } t_i \\ \min(f_k^{i-1} + 1, \theta) & \text{otherwise} \end{cases},$$

where $f_k^i$ denotes the value of feature corresponding to event $k$ at time $t_i$. The threshold $\theta$ was introduced under the assumption that the effect of a given event diminishes with time, eventually (after $\theta$ time intervals) rendering its influence unimportant. We chose this interval to be 1 day ($\theta = 1440$, the number of minutes in a day) since last activation. We assumed that none of the subsystems were active at the first time point, i.e., $f_k^0 = 0$. Table I illustrates three calculated DMOP features.

We construct such features for each subsystem-command pair (331 features) and also for each subsystem (34 features). Additionally, we construct binary indicators (34 features) for subsystems, where a feature $f_k^i$ has value of 1 if the subsystem was triggered within the time-step $t_i$, and 0 otherwise.

*4) FTL features:* The FTL data contain time ranges of pointing events with their starting and an ending time-points, where simultaneous events are also possible. We created a feature for each pointing event, where a value of 1 in a specific time period indicates the specific event is in progress, or 0 otherwise. This approach rendered 20 FTL features.

Namely, for events that are shorter than the time resolution of our dataset (i.e., 1 minute), we checked the time difference to the nearest existing time-stamp in our dataset. If the difference is smaller than 30s, we also insert the value 1 to the nearest existing time point.

*5) The final dataset:* The final datasets consists of approximately $2.7 \cdot 10^6$ examples and $464$ features: 42 energy influx and historical energy influx, 402 DMOP and 20 FTL features. The dataset is available at: http://kt.ijs.si/jurica_levatic/MEX/MEXdata.zip.

### B. Machine learning algorithms

To analyze the data and learn a predictive model, we used predictive clustering trees (PCTs) [4]. PCTs are a type of regression trees where the definitions of the heuristic and prototype functions are flexible and can be adapted for a variety of tasks. Regression trees are tree-like structures that have internal nodes and leaves. The internal nodes contain tests on the descriptive variables (i.e., the different features extracted with preprocessing), while leaves give predictions for the target variable (i.e., a heater line thermal consumption).

PCTs are built with a greedy recursive top-down induction algorithm. The learning algorithm starts by selecting a test for the root node by using a heuristic function (variance reduction) computed on the training examples. The goal of the heuristic is to guide the algorithm towards small trees with good predictive performance. Based on the selected test, the training set is partitioned into subsets according to the test outcome. This is recursively repeated to construct the subtrees. The partitioning process stops when a stopping criterion is satisfied (e.g., the minimal number of examples per leaf is reached, the heuristic score no longer changes, etc.).

An ensemble is a set of predictive models (called base models). The prediction of an ensemble for a new example is obtained by combining (e.g., averaging) the predictions of all base models from the ensemble. Here, we consider random forest of PCTs [5], [6]: an ensemble method that constructs the base models by making bootstrap samples of the training set and using each of these replicates to construct a PCT. The PCTs in this case are randomized in a way that at each tree node a random feature subset (with a user defined size) is considered for selecting the best split.

The reasons for using this ensemble method are (i) its state-of the-art predictive performance, and (ii) ability to calculate feature importance scores, i.e., ranking of the features w.r.t. their importance for the target variables. Namely, random forests can measure how much each feature contributed to the quality of the predictive model. By exploiting this we can, for instance, provide the operators of the spacecraft with valuable new knowledge about the factors influencing the thermal subsystem and the TPC. For this purpose, we used the GENIE3 algorithm [7] that calculates the importance of a feature as a sum of variance reductions over the splits containing that feature in all trees in the random forest. The rationale is that if a feature reduces the variance more, then it is more important.

In our experiments we learned one random forest for each power line separately. In all cases, we instantiated the algorithms as follows. First, as a stopping criterion for the PCTs, we set the minimal number of examples in the leaves to 500. Next, we construct random forest with 200 PCTs and the feature subset size in each node was set to $\frac{1}{4}$ of the number of all features. Finally, we use the same settings for obtaining the feature importance scores.

## IV. Results and Discussion

Table II presents the final scores achieved by the top 5 teams at the MEX Power Challenge, compared to the performance of the model currently in use by the operators of MEX. We can observe that all machine learning solutions achieve much lower error when compared to ESA's current model. This suggests that machine learning approaches could substantially advance the task of predicting the satellite's TPC.

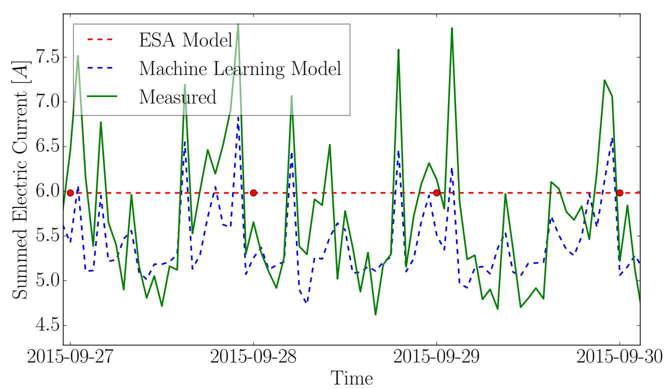


Fig. 2. Comparison of the measured TPC (green line) with the predicted power consumption with the machine learning model proposed in this study (blue dashed line) and model currently used by ESA (red dashed line). The power consumption is presented as a sum of the electric current (in Amperes) of 33 power lines.

TABLE II
Comparison of the final scores of the top 5 teams in the MEX competition (our approach is marked in bold) to the score of the ESA's model.

| Team | **MMMe8** | redrock | fornax | Alex | luis | ESA |
|---|---|---|---|---|---|---|
| RMSE | **0.0792** | 0.0803 | 0.0819 | 0.0838 | 0.0884 | 0.4903 |

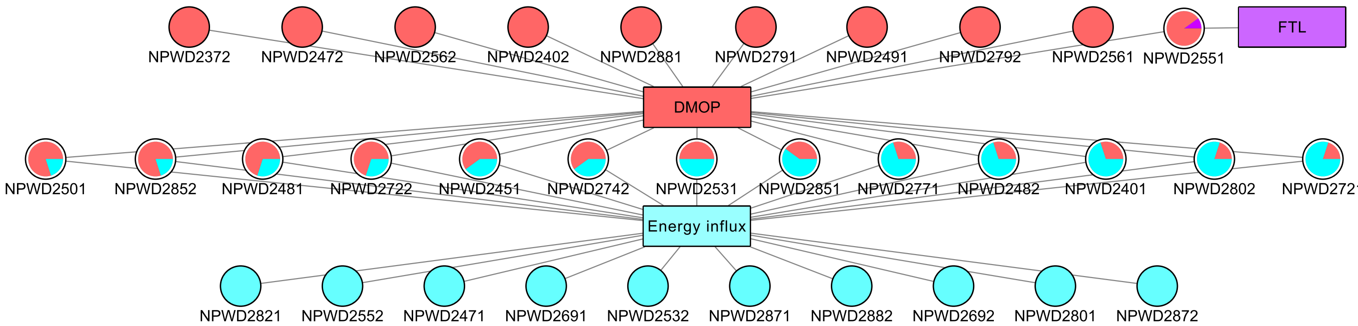


Fig. 3. Graph of feature influences. Each circle represents one of the power lines, while its color shows the proportion of the top 10 most important features (for that power line) that belongs to each of the groups of features, i.e., energy influx (cyan), DMOP (red), and FTL (purple).

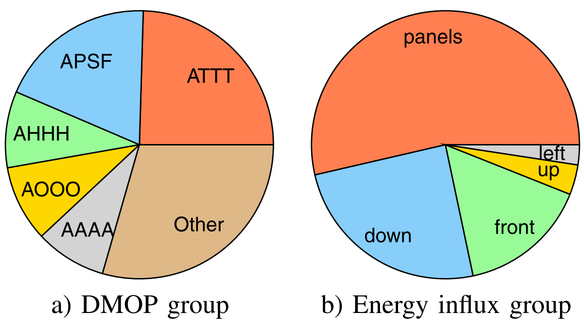


Fig. 4. a) The proportions of the features from the DMOP group (see Fig. 3) related to different subsystems of MEX. b) The proportions of the features from the Energy influx group (see Fig. 3) related to panels or sides of MEX.

We next highlight the difference between the measured and the predicted TPC of our machine learning model and ESA's model on Fig. 2. Notably, the predictions of our model closely follow the dynamics of the measured electric current, while ESA's model only predicts a general trend (i.e., it estimates average daily TPC). The model presented in this study makes an order of magnitude smaller error than the ESA's model, i.e., 0.0792 vs. 0.4903. The lower accuracy of the currently used MEX TPC model means also lower available power for science observations, because safety margin must be higher. A machine learning prediction model, with greater accuracy, would therefore increase the available power for science observations. This is the topic of an on-going study at ESA, assessing the performance of machine learning prediction models.

The feature importance analysis (Fig. 3) reveals that there are two groups of 10 power lines each, which are influenced either by DMOP features either by energy influx features (top and bottom row in Fig. 3, respectively). Recall that DMOP features represent activities of various subsystems of the spacecraft, and can thus be considered as "internal" influences on the TPC. On the other hand, energy influx features can be considered as "external" influences on the TPC, since they represent the amount of solar energy reaching the sides of the spacecraft. In other words, our analysis suggests that some power lines are predominantly influenced by the internal influences, some predominantly by the external influences, while the remaining power lines are influenced by both.

This analysis also suggests that the FTL features have low utility in predicting the TPC of MEX. These features capture the flight dynamics of MEX, e.g., pointing the spacecraft towards Earth. While such events may impact the thermal behaviour, their effect is also likely reflected in the solar aspect angles (i.e., energy influx features) and in the subsystem commands (i.e., DMOP features).

Furthermore, we analyzed the individual features that appear in the DMOP and the energy influx groups in Fig. 3. The analysis of the DMOP group (Fig. 4a) suggests that the influence of the MEX's subsystems to the TPC is heterogeneous. Namely, 5 out of the 34 MEX's subsystems are responsible for most of the influence, while the other subsystems only have a minor influence. Specifically, the radio transmission subsystem (i.e., ATTT) has the biggest impact on the TPC, followed by the APSF, AHHH, AOOO and AAAA subsystems. This finding was also confirmed by the MEX operators.

The analysis of the energy influx group (Fig. 4a), revealed that the most important external factors are the features related to (solar array) *panels* of the spacecraft, followed by the *down* and *front* sides of the spacecraft, while features related to *up*, *left*, *back* and *right* sides have low or even negligible importance. These findings are supported by the following reasoning: the features related to solar array panels indicate if the arrays are illuminated or not. This, in turn, reveals if MEX is receiving solar energy or is in eclipse during which temperatures drops dramatically. The fuel lines are located at the bottom (down) side of the spacecraft. These need to be constantly heated to prevent the fuel from freezing. When this side of the spacecraft is illuminated and receiving solar energy, indirectly the fuel lines are heated. They reach temperatures in ranges where the spacecraft electrical heaters do not activate, hence greatly decreasing the TPC.

Note that, some aspects of the proposed methodology are based on assumptions without expert knowledge in spacecraft construction and operation. The umbra coefficient of energy influx features, time-spans of historical energy influx features and the $\theta$ threshold of DMOP features could likely still be optimized.

Next, we discuss the computational complexity of the methodology for TPC presented in this paper. Two key factors influence the computational efficiency of the proposed machine learning pipeline: the size of the training dataset and the underlying learning method. We generated a training dataset with a time resolution of 1 minute, which results in nearly 3 millions examples, each described with 464 features. This yields a massive amount of data, which may be challenging to handle on a standard desktop computer. On average, it took about 15 hours per power line to learn a random forest model and calculate predictions for one Martian year. However, the models for different power lines can be built in parallel, since they are independent from each other. We used a computer cluster, and were able to obtain the complete solution in about 15 hours total. Notably, the models, once built, can be used to quickly predict the TPC beyond the time periods included in the MEX Power Challenge (given that the context data is available), without re-learning them.

The presented pipeline for prediction of the TPC was developed specifically for the purpose of the MEX Power Challenge and it was not optimized for computational efficiency, but for maximal predictive accuracy of the models. There are, however, several ways for reducing the computational demand of the proposed pipeline, without a considerable impact on its predictive power.

One step towards reducing the computational demand is to exclude the features with negligible importance from the dataset. This will reduce learning time and memory consumption of training the models, presumably without a noticeable impact on the predictive performance. Moreover, the size of the dataset can be reduced by using a time resolution coarser than 1 minute. However, the impact of this approach on the predictive performance remains to be investigated. The underlying learning method can also be refined to yield a more time efficient solution. Recall that, the we used random forests of PCTs, which generalize ordinary classification and regression trees towards predicting structured outputs, such as multi-target prediction. Multi-target models predict multiple numerical values simultaneously. Such models can predict the TPC of all 33 power line with one model, while the single-target approach we used in this study learns a separate model for each power line. Due to this, the multi-target approach is approximately 33 times faster, while achieving comparable predictive performance: Our proof-of-principle experiment shows that a random forest of multi-target PCTs [5] (with the same parameters as reported in Section III-B) achieves a RMSE of $0.083$.

## V. Conclusion

In order to optimize future scientific operations of the Mars Express orbiter, it is essential to estimate in advance the amount of power consumed by its thermal subsystem. However, this is a challenging task because of the sheer amount of telemetry data gathered from the orbiter and the complex relationship between the TPC and various external and internal factors. This challenge led to the organization of ESA's first data mining competition. This paper presents a machine learning-based pipeline for TPC prediction that won the challenge. The proposed methodology transforms raw telemetry data into features usable by machine learning methods, learns predictive models, and predicts future TPC with high accuracy. The two main contributions of the proposed methodology are as follows.

First, the results presented in this study show that machine learning models achieve predictive performance superior to the performance of the model that is currently used by the operators of MEX. Machine learning models could thus assist the operators to estimate the TPC of MEX more accurately and in turn allocate more electricity for the scientific operations. Consequently, this could prolong the operating life of MEX.

Second, we demonstrate that machine learning can also be used to provide the operators with insight about the thermal behaviour of MEX. Our analysis suggests that some of the MEX's thermal power lines are predominantly affected by external factors (i.e., solar radiation), while some are predominantly affected by the internal factors (i.e., activities of MEX's subsystems). Finally, the analysis also suggests that the majority of the influence to the thermal subsystem can be contributed to the solar radiation hitting specific parts of MEX (i.e., its solar panels and its bottom and front sides), and the activities of 5 out of 34 of the satellite's subsystems.

There are several direction to continue the work presented in this paper. First, the predictive performance of the proposed solution can be improved by further optimizing the constructed features, also by constructing new (informative) features from context data that was not given during the competition (e.g., space weather data). Next, we plan to optimize the computational complexity of the proposed pipeline, for instance by reducing the feature set to include only the highly important features. Finally, the proposed methodology focuses on the thermal subsystem of the MEX orbiter, however, it could also be extended to other subsystems of the spacecraft.

## Acknowledgements

We thank Sašo Džeroski for support and helpful scientific discussions. We also thank the Slovenian grid initiative and the Academic and Research Network of Slovenia for support with the computer infrastructure. We acknowledge the financial support of the Slovenian Research Agency, via the young researcher grants to MB, JL, AO and MP, as well as the European Commission, via the grants ICT-2013-612944 MAESTRA and ICT-2016-720270 HBP-SGA1.

## References

[1] Augustin Chicarro, Patrick Martin, et al. The Mars Express mission: An overview. In *Mars Express: The Scientific Payload*, volume 1240, pages 3–13, 2004.

[2] Tom Mitchell. *Machine learning*. McGraw Hill, 1997.

[3] Luke Lucas and Redouane Boumghar. Machine learning for spacecraft operations support - The Mars Express Power Challenge. In *Sixth International Conference on Space Mission Challenges for Information Technology, SMC-IT 2017*, 2017.

[4] Hendrik Blockeel, Luc De Raedt, et al. Top-down induction of clustering trees. *Proc. of the 15th Int'l Conf. on Machine learning*, pages 55–63, 1998.

[5] Dragi Kocev, Celine Vens, et al. Tree ensembles for predicting structured outputs. *Pattern Recognition*, 46(3):817–833, 2013.

[6] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[7] Vân Anh Huynh-Thu, Alexandre Irrthum, et al. Inferring regulatory networks from expression data using tree-based methods. *PLOS ONE*, 5(9):1–10, 09 2010.