

# International Society for Photogrammetry and Remote Sensing (ISPRS) Book Series

*Book Series Editor*

**Paul Aplin**

*School of Geography*

*The University of Nottingham*

*Nottingham, UK*



**Information from imagery**

Environmental tracking and modeling studies have shown – whether at the local, regional, or national level – that a significant amount of environmental information can be collected through automated systems and effectively integrated to monitor, report and analyze broad areas of interest associated with disease. In addition, these types of data provide useful information for decision making and planning.

# Environmental Tracking for Public Health Surveillance

Editors

Stanley A. Morain & Amelia M. Budge

*Earth Data Analysis Center, University of New Mexico, Albuquerque,  
New Mexico, USA*



**CRC Press**

Taylor & Francis Group

Boca Raton London New York Leiden

CRC Press is an imprint of the  
Taylor & Francis Group, an **informa** business  
A BALKEMA BOOK

*Cover image:*

Environmental tracking involves a host of international space agency satellites and their on-board sensors. The upper portion of the image illustrates one of several satellite constellations (the *A-Train*) carrying land, ocean, and atmospheric sensors. The bottom three elements of the image show a pollen burst resulting in a respiratory reaction (left), typical disease carriers like mosquitoes, hookworms and ticks (center), and public health interventions like vaccinations (right). The challenge is to improve current environmental tracking capabilities with next generation sensor systems to predict disease threats and mitigate their outcomes.

*Image credits:*

*A-Train:* Courtesy US National Aeronautics and Space Administration (NASA)

*Girl sneezing:* Courtesy frogblog

*Mosquito:* Photo by James Gathany courtesy Centers for Disease Control and Prevention (CDC)

*Tick:* Photo by James Gathany courtesy CDC

*Hook worm:* Courtesy CDC

*Boy being vaccinated:* Photo by James Gathany courtesy CDC

*CRC Press/Balkema is an imprint of the Taylor & Francis Group, an informa business*

© 2013 Taylor & Francis Group, London, UK

Typeset by MPS Limited, Chennai, India

Printed and Bound in the United States of America by Edwards Brothers Malloy

All rights reserved. No part of this publication or the information contained herein may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, by photocopying, recording or otherwise, without written prior permission from the publisher.

Although all care is taken to ensure integrity and the quality of this publication and the information herein, no responsibility is assumed by the publishers nor the author for any damage to the property or persons as a result of operation or use of this publication and/or the information contained herein.

Library of Congress Cataloging-in-Publication Data

Environmental tracking for public health surveillance / editors, Stanley A. Morain & Amelia M. Budge.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-415-58471-5 (hbk.) – ISBN 978-0-203-09327-6 (eBook)

I. Morain, Stanley A. II. Budge, Amelia M.

[DNLM: 1. Disease Outbreaks—prevention & control. 2. Environment. 3. Information Systems. 4. Population Surveillance. 5. Satellite Communications. WA 105]  
362.1—dc23

2012025102

Published by: CRC Press/Balkema

P.O. Box 447, 2300 AK Leiden, The Netherlands

e-mail: Pub.NL@taylorandfrancis.com

[www.crcpress.com](http://www.crcpress.com) – [www.taylorandfrancis.com](http://www.taylorandfrancis.com)

ISBN: 978-0-415-58471-5 (Hbk)

ISBN: 978-0-203-09327-6 (eBook)

## Table of contents

Preface	vii
Foreword	ix
List of contributors	xi
Acronyms	xix
<i>Part I Introduction</i>	
Chapter 1 Earth observing data for health applications S.A. Morain & A.M. Budge (Author/editors)	3
<i>Part II Infectious and contagious diseases in the environment</i>	
Chapter 2 Vector-borne infectious diseases and influenza R.K. Kiang (Author/editor)	21
Chapter 3 Water, water quality and health S.I. Zeeman & P. Weinstein (Author/editors)	87
Chapter 4 Air quality and human health D.W. Griffin & E.N. Naumova (Author/editors)	129
Chapter 5 Emerging and re-emerging diseases C.J. Witt (Author/editor)	187
<i>Part III Data, modelling, and information systems</i>	
Chapter 6 Data discovery, access and retrieval S. Kempler (Author/editor)	229
Chapter 7 Environmental modelling for health S.A. Morain, S. Kumar & T.J. Stohlgren (Author/editors)	293
Chapter 8 Early warning systems P. Ceccato & S.J. Connor (Author/editors)	333
Chapter 9 Towards operational forecasts of algal blooms and pathogens C.W. Brown (Author/editor)	345
Chapter 10 Information and decision support systems W. Hudspeth (Author/editor)	369
Author index	411
Subject index	413
ISPRS Book Series	419
Colour plates	421



## Chapter 6

### Data discovery, access and retrieval

S. Kempler (Auth./ed.)<sup>1</sup> with: G.G. Leptoukh<sup>1</sup>, R.K. Kiang<sup>1</sup>, R.P. Soebiyanto<sup>1</sup>, D.Q. Tong<sup>2</sup>, P. Ceccato<sup>3</sup>, S. Maxwell<sup>4</sup>, R.G. Rommel<sup>4</sup>, G.M. Jacquez<sup>4</sup>, K.K. Benedict<sup>5</sup>, S.A. Morain<sup>5</sup>, P. Yang<sup>6</sup>, Q. Huang<sup>6</sup>, M.L. Golden<sup>7</sup>, R.S. Chen<sup>7</sup>, J.E. Pinzon<sup>8</sup>, B. Zaitchik<sup>9</sup>, D. Irwin<sup>10</sup>, S. Estes<sup>10</sup>, J. Luvall<sup>10</sup>, M. Wimberly<sup>11</sup>, X. Xiao<sup>12</sup>, K.M. Charland<sup>13</sup>, R.P. Stumpf<sup>14</sup>, Z. Deng<sup>15</sup>, C.E. Tilburg<sup>16</sup>, Y. Liu<sup>17</sup>, L. McClure<sup>18</sup>, & A. Huff<sup>19</sup>

<sup>1</sup>NASA GSFC, Greenbelt, MD, US

<sup>2</sup>SISS, George Mason University, Fairfax, VA, US

<sup>3</sup>IRI, Columbia University, Palisades, NY, US

<sup>4</sup>BioMedware, Ann Arbor, MI, US

<sup>5</sup>EDAC, University of New Mexico, Albuquerque, NM, US

<sup>6</sup>George Mason University, Fairfax, VA, US

<sup>7</sup>SEDAC, Columbia University, Palisades, NY, US

<sup>8</sup>SSAI, Lanham, MD, US

<sup>9</sup>Johns Hopkins University, Baltimore, MD, US

<sup>10</sup>Marshall Space Flight Center, Huntsville, AL, US

<sup>11</sup>South Dakota State University, Brookings, South Dakota, US

<sup>12</sup>University of Oklahoma, Norman, OK, US

<sup>13</sup>Children's Hospital, Boston, MA, US

<sup>14</sup>NOAA National Ocean Service, Silver Spring MD, US

<sup>15</sup>Civil & Environmental Engineering, Louisiana State University, Baton Rouge, LA, US

<sup>16</sup>University of New England, Biddeford, ME, US

<sup>17</sup>School of Public Health, Emory University, Atlanta, GA, US

<sup>18</sup>University of Alabama, Birmingham, AL, US

<sup>19</sup>Battelle Memorial Institute, Columbus, OH, US

**ABSTRACT:** This chapter explores the complex, and sometimes frustrating, world of data discovery, access, delivery and use by reference to the US National Aeronautics & Space Administration's (NASA's) public health applications portfolio in 2011. It also provides examples of global information system applications in health.

#### 1 INTRODUCTION

NASA's Applied Sciences Division (ASD) promotes innovation by public and private sector organizations applying satellite data, models, products, and scientific findings for air quality management and policy activities benefitting human health and safety. The program focuses on themes of air quality planning, forecasting and compliance and the two crosscutting themes of climate and emissions inventories. It also focuses on areas of infectious disease, emergency preparedness and response, and environmental impacts.

Although satellites obviously cannot monitor swarms of malaria-carrying mosquitoes or other zoonotics, they carry sensors that record environmental attributes that control mosquito populations and distributions, and are excellent tools for gathering and transmitting data from remote regions. This makes it possible to record and monitor geographic or meteorological factors favouring mosquito habitats and the possible onset of diseases before they occur (CNES 2011).

In general, international space assets, infrastructure and expertise aim to improve human understanding of the solar system and beyond. Earth observations form a subset of this understanding

Table 1. URLs referenced in section 2.0.

Product/service	URL (all accessed 18th January 2012)
A. Resilience human populations	<a href="http://www.enotes.com/public-health-encyclopedia/vector-borne-diseases">http://www.enotes.com/public-health-encyclopedia/vector-borne diseases</a>
B. SERVIR	<a href="http://www.servirglobal.net/en/AboutSERVIR.aspx">http://www.servirglobal.net/en/AboutSERVIR.aspx</a>
C. EASTWeb	<a href="http://globalmonitoring.sdsstate.edu/eastweb">http://globalmonitoring.sdsstate.edu/eastweb</a>
D. Hong Kong observatory	<a href="http://www.hko.gov.hk/wxinfo/pastwx/extract.htm">http://www.hko.gov.hk/wxinfo/pastwx/extract.htm</a>
E. Climate and health effects	<a href="http://www.cdc.gov/climatechange/effects/default.htm">http://www.cdc.gov/climatechange/effects/default.htm</a>
F. Healthy water	<a href="http://www.cdc.gov/healthywater/disease">http://www.cdc.gov/healthywater/disease</a>
G. Tracking algal blooms	<a href="http://ccma.nos.noaa.gov/stressors/extremeevents/hab/RSFieldOps.aspx">http://ccma.nos.noaa.gov/stressors/extremeevents/hab/RSFieldOps.aspx</a>

as scientists probe for Earth-like planets. The knowledge gained by looking back to Earth helps ground-positioning, communication satellites and associated applications, becoming key tools towards strengthening preparedness, improving surveillance, and providing effective early-warning (ESA 2011).

The Japanese Space Agency (JAXA) also has a concept on space initiatives for health (Igarashi 2010). It is founded on a similar principal that space based remote sensing data can be used to monitor and predict air quality conditions that enable decision support tools and early warning systems for specific diseases. In addition, the international GEO recognizes that collaboration is essential for exploiting the growing potential of Earth observations to support decision making in an increasingly complex and environmentally stressed world (GEO 2011a). In regards to health, GEO defined two tasks in its 2012–2015 Work Plan: HE-01 *Tools and Information for Health Decision Making*, and HE-09-02 *Tracking Pollutants* (GEO 2011b).

These are all testaments to the awareness and willingness of the international community of space agencies and Earth observing organizations to invest resources that bridge scientific discoveries and practical applications that benefit society. These initiatives form the backdrop stimulating remote sensing aided public health research and applications projects. The keystones for facilitating the success of these projects is both the ability to transform long-term environmental data into relevant information useful for public health studies, and the research community's understanding of how emerging and evolving sources of information can be utilized to enhance their work. Using satellite remote sensing data provides the following benefits to environmental data collection: 1) enables continuous data acquisition; 2) helps to refresh environmental data sets with each satellite overpass, conditions permitting; 3) offers synoptic coverage and good spectral resolution to augment point source ground measurements 4) offers accurate data for information and analysis; and 5) serves as a large archive of historical data (ClearLead 2011). The purpose of this Chapter is to provide insights on the availability of EO data and information services applicable to public health interests, and how this information is currently being applied in public health research and development.

Sections 2, 3, and 4 address different aspects and perspectives for how convergence of EO data and health can occur. Section 2 provides synopses of ASD projects to give readers a quick introduction to the variety of topics being developed. They have been contributed by their respective project Principal Investigators as indicated in Tables 2, 3, and 4, noting that project titles have been edited for length. Section 3 is an in-depth look at three initiatives showing how sensor data often are used for public health, and how remote sensing data enhance research. Section 4 focuses on three case studies on how to access and use appropriate data sets.

## 2 EARTH SCIENCE SENSOR DATA FOR PUBLIC HEALTH APPLICATIONS

Section 2 is a sample of project synopses from ASD's health and air quality program 2011 portfolio of health projects (NASA 2012). The purpose for profiling them is to show how they link environment and health communities; how they advance the roles of spectral and spatial science; and how they lead to social and economic benefits. It is clear that both communities need to form collaborations to synergize their distinct skill sets, and to fuel ideas for further data and methods

Table 2. Selected NASA-funded research applications on vector-borne diseases.

Principal investigator	Brief title of project
J.E. Pinzon, GSFC/SSAI	Predicting zoonotic haemorrhagic fever in sub-Saharan Africa
B. Zaitchik, Johns Hopkins Univ.	Detection and early warning for malaria risk in the Amazon
D. Irwin & J. Kessler, MSFC	SERVIR Africa
S. Estes, MSFC/USRA	Potential range expansion of <i>Aedes aegypti</i> in Mexico
M. Wimberly, SDSU	Forecasting mosquito-borne disease outbreaks using AMSR-E
R. Kiang, GSFC	Avian influenza risk prediction in Southeast Asia
X. Xiao, Univ. of OK	Sensor imagery and satellite telemetry for avian influenza
K. Charland, Children's Hospital, Boston	Data for an influenza forecasting system

sharing amongst new and experienced users. Since 2003, ASD has funded some thirty projects related to health. They describe uses of EO data sets, collateral data sets; methods for using data; techniques for modelling and transforming data into public health information; and, anticipated project outputs and benefits to appropriate user communities. Projects are presented in three broad disease categories: vector-borne, water-borne, and air-borne. For ease of presentation, all URLs for this section are listed in Table 1. Other citations for URLs are given in the reference section to this Chapter in standard reference format. Table 1 is a list of URLs referenced in Section 2, and Table 2 lists a selection of NASA-funded projects focused on for vector-borne diseases.

### 2.1 Data sets for vector-borne disease studies

Vectors are the transmitters of disease-causing organisms that carry pathogens from one host to another. By common usage, vectors are considered to be invertebrate animals, usually arthropods. Furthermore, key components that trigger vector-borne diseases include: 1) abundance of vectors and reservoir hosts; 2) prevalence of disease-causing pathogens suitably adapted to the vectors and the human or animal host; 3) local environmental conditions, especially temperature and humidity; and 4) resilience behaviour and immune status of the human population (Table 1, A). It is the third component, environmental conditions that most projects address. The significance of these project synopses is realized when considering the fragile balance of our physical and biological environment. Abrupt or long-term changes in abiotic factors may lead to an alteration in ecosystem equilibrium, resulting in more or less favourable vector habitats. Anticipated changes in temperature and precipitation resulting from global warming will affect vector reproduction and longevity, rate of development of the parasites, and pathogens in the vector, as well as the geographic extent of invasions (WHO 1990). Abiotic environmental factors also affect vegetation patterns that determine distributions of disease-causing vectors. Some of these are temperature, precipitation, relative humidity, wind, solar radiation, topography, and fresh water rivers, ponds, and lakes (Kay *et al.*, 1989). Research continues in understanding and modelling environmental factors used to predict the magnitude and direction of an infected vector, translate these predictions to early warning systems, and aid policy decision makers.

#### 2.1.1 Predicting zoonotic haemorrhagic fever in sub-Saharan Africa

Knowledge generated about infectious diseases through basic research on associations, causes and effects of climate has been an important source of information for public health decision makers. The processes involved and their dynamics show a need for synergistic interpretations with field observation that include monitoring and analysis of remotely sensed data. This project involves a multidisciplinary team of earth scientists, epidemiologists and public health experts using satellite data to map the risk of outbreaks of deadly viruses like EBOV, MARV and RVFV.

Emerging infectious diseases are a global and regional security issue with the capacity to have serious human health and economic impacts, worldwide. This project is integrating EO data into a global emerging infectious surveillance and response system (GEIS) to complement it with a systematic method for monitoring and forecasting environmental and climatic risk factors associated

with emerging infectious diseases. NDVI temperature data from the MODIS instruments on Terra and Aqua; TRMM/GPCP (precipitation) data, SRTM (elevation data) and simulated products from both the future NPOESS preparatory project (NPP) and the Global Precipitation Mission (GPM) are projected. These data sets will aid in migrating a reactive surveillance system to one that is proactive. The goal of this synergy is to facilitate decision making with early warning tools.

The end-of-project goal is an eco-climate monitoring algorithm that assesses environmental and climatic risk factors quantitatively, and that could lead to outbreaks of vector-borne diseases. It will provide risk maps that highlight areas where targeted surveillance should be implemented. Monthly environmental risk maps for zoonoses focusing on EBOV, MARV filoviruses and RVF in Africa will be provided. Environmental risk maps support GEIS efforts to improve surveillance systems crucial for preventing, detecting and containing emerging infections that threaten military personnel, their families, and national security, and enhance GEIS overseas military research units with their service to host country counterparts, WHO, and FAO to improve local epidemiological capabilities.

#### *2.1.2 Detection & early warning system for malaria risk in the Amazon*

Malaria is a leading cause of morbidity in the Amazon basin. Major challenges remain in targeting, intervention, and control strategies; in particular the distribution of health resources (treatments, diagnostics, and long-lasting impregnated nets) due to eco-social dynamics that result in differences between where people are infected and where they are diagnosed. An effective malaria risk monitoring system begins with reliable estimates of the eco-physiological factors that drive *Anopheline* mosquito populations. On the Amazon frontier, as in many other malaria-prone regions, *Anopheline* density is known to be associated with meteorological factors, including precipitation and air temperature, with land surface characteristics, including land cover and local topography, and with surface hydrological conditions, including soil moisture and surface water ponding. Thus, this project's first objective is to merge multiple EO data sets through a land data assimilation system (LDAS) that can be used to drive spatially explicit ecological models of anopheline mosquito distribution. Assimilations will be driven by observational data. LDAS simulations will make use of satellite-derived meteorological forcing data, parameter data sets, and assimilation observations, including: 1) precipitation from TRMM and GPM; 2) land cover from MODIS, Landsat, ASTER, and LDCM; 3) soil moisture from AMSR-E (where applicable) and Soil Moisture Active-Passive (SMAP); 4) terrestrial water storage from the Gravity Recovery and Climate Experiment (GRACE and GRACE-II); 5) surface temperature from MODIS, Landsat, ASTER, LDCM; 6) Vegetation Fraction/Leaf Area Index from MODIS, Landsat, ASTER, LDCM; and 7) topography from SRTM. Earth science data will be important drivers of models for mosquito population dynamics. The LDAS will drive ecologically based models of mosquito dynamics, developed using data from larva and adult mosquito collection sites maintained in the study region.

The second objective is to develop a human activity and settlements map that uses a spatially explicit model of human settlement locations derived from census and regional studies, areas of forest concessions, and indicators of forest disturbance to identify permanent and temporary sites of human activity. Sensor data will help generate permanent and temporary areas of human activity and augment maps of human settlements.

The third objective integrates ecological and human population models resulting from the companion objectives to create spatial risk maps of human malaria risk. Overall, the project builds on collaborations between investigators expert in land use, climate and ecological modelling, and epidemiologists who are expert in vector-borne diseases, biostatistics, and demography to inform health interventions.

#### *2.1.3 SERVIR Africa*

Efforts to implement public health early warning and decision support systems are not limited to small teams using new technologies and information sources to improve state-of-the-art health surveillance. For example, SERVIR (a Spanish acronym for *Regional Visualization and Monitoring System*) (Table 1, B) enables EO data sets and predictive models to be used for timely decisions

through regional platforms in Mesoamerica, East Africa, and the Hindu-Kush Himalayas. The system was developed initially in 2004 through joint sponsorship of NASA, the US Agency for International Development (US/AID), the World Bank, and the Central American Commission for Environment and Development (CCAD). It is an initiative that applies earth observations and predictive models to support decision making by government officials, managers, scientists, researchers, students, and the public. In 2005, the Water Center for the Humid Tropics of Latin America and the Caribbean (CATHALAC) in Panama became the first regional SERVIR facility, serving Central America and the Dominican Republic. In late 2008, a SERVIR facility at the Regional Center for Mapping of Resources for Development (RCMRD) in Nairobi, Kenya, was dedicated to serve East Africa. A SERVIR facility for the Hindu-Kush Himalaya region in Asia was inaugurated in Kathmandu, Nepal in October, 2010. US/AID and NASA provide primary support for SERVIR, with the long term goal of transferring its capabilities to other host countries.

The objective of this project is to initiate a SERVIR-like system using NASA science research results to improve decision support in Africa. The initial focus is on flooding and RVF because they are both related to several of the environmental parameters necessary to monitor them: precipitation, topography, soil moisture and land cover. Data from multiple missions, sensors, and models are being used. These include SRTM, AMSR-E, TRMM, MODIS, the global hazard model-flood (GHMF), and the infectious disease eco-climatic link (IDEL) algorithm. The goal is to create a functioning SERVIR-Africa node at RCMRD and to improve current decision making processes of the Kenya flood response system. Additional benefits will be improved flood forecasting and monitoring capabilities at the operational SERVIR facility in Central America.

In addition, SERVIR has become a link to several relevant articles (Kalnay *et al.*, 1996; Huffman *et al.*, 2001; Uppala *et al.*, 2008; Caminade *et al.*, 2011). Caminade (*et al.*, 2011), in particular, discusses using climatic indicators to map RVF and malaria over West Africa. In order to model and map both diseases, different climate data sets are used. Daily rainfall is estimated using mixed satellite and rain gauge observations from the global precipitation climatology project (GPCP) data set, and rainfall and temperature from both the National Centers for Environmental Prediction (NCEP) and the European Centre for Medium Range Weather Forecasts (ECMWF) reanalysis projects, as described in ECMWF Newsletter No 110 (ECMWF 2007). Dynamic models driven by daily rainfall and temperature simulate malaria incidence in the human population, and a methodology for tracking weather and climate events that enhance RVF risk over West Africa are being employed.

#### 2.1.4 Range expansion of *Aedes aegypti* in Mexico

This project is an international collaboration between the Universities Space Research Association (USRA), NASA, The National Center for Atmospheric Research (NCAR), CDC, and the University of Veracruz in Mexico. By modelling the social, economic, environmental and epidemiological factors that influence the survival and abundance of *Ae. Aegypti*, it will be easier to control the primary transmitters of dengue viruses. The ultimate goal is to employ this integrated modelling approach to understand the potential of *Ae Aegypti* to expand its range into heavily populated, high elevation areas like Mexico City under various climate and socioeconomic scenarios.

The first objective is to employ EO data to augment the environmental monitoring and modelling component of the National Science Foundation project. Several remote sensing products have been proposed and their usefulness is being determined through ground verification. The data and their uses include: 1) MODIS NDVI to monitor vegetation and seasonality in a transect area from Veracruz to Mexico City. This 250 m resolution product is a sixteen-day average ratio of the reflectance of red and near infrared spectral bands. It is a gridded, product, masked for water, clouds, and cloud shadows; 2) MODIS LST is used as an indicator of Earth's surface energy balance often used to relate climate, hydrological, ecological and other environmental variables; 3) MODIS LCLU data are used to represent land cover derived from a year's input of Terra observations. The scheme identifies seventeen land cover classes defined by the International Geosphere Biosphere Programme (IGBP); 4) soil moisture data from the AMSR-E are used in regions of low to moderate vegetation and processed to provide soil moisture conditions to a depth of one to two centimetres (Njoku

*et al.*, 2003); 5) digital elevation from SRTM are used to provide surface topography (Farr *et al.*, 2007); and 6) climate prediction centre morphing technique (CMORPH) will provide precipitation data based on microwave observations from multiple sensors. Products are available at several time intervals, the most appropriate for this purpose being thirty minute and three hour intervals. These data are critical for understanding the habitat requirements necessary for mosquito survival and spread.

#### 2.1.5 Forecasting mosquito-borne disease outbreaks using AMSR-E

This research uses data from AMSR-E to develop improved environmental models of mosquito-borne disease risk. AMSR-E products provide several critical environmental variables that are directly relevant to mosquito ecology, including near-surface air temperature, soil moisture, and fractional water cover. These variables will be used to model and forecast inter-annual fluctuations in WNV cases and mosquito populations. A limitation of vegetation indices, land surface temperature, and many other products derived from existing remote sensing platforms is that they provide indirect measurements of the proximal environmental factors influencing mosquito populations and disease risk. The recent development of a new set of daily global land surface parameters derived from AMSR-E offers novel environmental metrics and expanded opportunities for mosquito-borne disease risk forecasting. The specific objectives of this research are to: 1) develop statistical models of WNV risk and mosquito population dynamics; 2) compare their performance with models based on MODIS and TRMM products to quantify potential improvements in forecasting WNV outbreaks; and 3) generate and disseminate early warning predictions from these models to gain qualitative feedback from vector control experts and public health practitioners.

These objectives are to use the archive of AMSR-E land surface observations commencing in 2002 and to summarize the data temporally to match eight-day MODIS composites. These will be spatially matched to human case data at the county level and mosquito data at the city boundary level. Appropriate statistical methods are being applied to account for spatial and temporal autocorrelation, seasonality, and the effects of other environmental variables such as land cover and land use. Data from the AMSR-2 sensor on the JAXA GCOM-W satellite will be used in future to produce seasonal forecasts of WNV risk for the South Dakota Department of Health. The expected benefits of this project are: 1) its prospect for incorporating novel information from a region of the electromagnetic spectrum currently underutilized in mosquito-borne disease research; and 2) its prospect for producing a more accurate and effective early-warning system.

The larger dimension of the effort is to develop early warning systems that can forecast areas of future disease risk based on environmental variability in space and time. The epidemiological applications of spatial technologies project (Table 1, C) is focused on WNV in the northern Great Plains of the US, and malaria in the Amhara region of Ethiopia (EASTWeb 2012). So far progress has been made applying geospatial technologies for mapping, risk analysis, and ecological forecasting of infectious diseases. An earlier project applied merged sensor data with a GIS to manage and process spatial statistics for analysing disease patterns and developing predictive models to explain how climate and land use patterns influence outbreaks of vector and host populations.

#### 2.1.6 Avian influenza risk prediction and early warning of pandemic influenza

This research utilizes sensor data, models and analysis techniques to enhance decision support capabilities for avian influenza (AI) and pandemic influenza (PI) risks. The first objective focused on assessing AI risks for poultry farms and humans and on the potential for early detection of pandemic influenza. In particular, spatial and-temporal risks of H5N1 outbreaks for selected districts in Indonesia were generated, along with short- and medium-term influenza-like illness forecasts for selected regions. The second objective was to advance the established capabilities for modelling seasonal influenza created in the former project, by refining the models and extending their capabilities to a global scale. Toward this end, climate-based models to predict influenza risks in major cities around the world are being developed.

Another research focus uses weekly counts of laboratory-confirmed influenza viruses, and climatic and meteorological parameters collected from two primary sources: ground-based and satellite-derived measurements. Daily meteorological observations were retrieved from the Hong Kong Observatory (Table 1, D) including maximum, mean, minimum temperature, mean dew point temperature, mean relative humidity, global solar radiation and total evaporation. For the US study site in Maricopa County, Arizona, daily climatic observations were acquired from the local flood control district. Data were aggregated from thirty-two stations. These data include daily mean air temperature, minimum, mean and maximum dew point, minimum and maximum relative humidity, maximum wind speed, minimum and maximum air pressure, and maximum solar radiation. EO measurements of daily rainfall also were obtained for both Hong Kong and Maricopa County from instruments on TRMM. Daily LST data from MODIS also were extracted. Both Terra and Aqua missions carry this instrument. Temperature and precipitation are very important environmental determinants of infectious disease transmission (Kiang *et al.*, 2006; Xiao *et al.*, 2007; Soebiyanto *et al.*, 2010). Consequently, the multinational GPM is an important successor to TRMM. It is scheduled for launch in 2013. If successful, it should provide a key data set for continued influenza monitoring.

Understanding and controlling AI outbreaks brings substantially more benefits to the society than just the farms where the outbreaks occur. It spares extensive culling, preserves the livelihood of small farmers, and protects food security and biodiversity. The expansion of developed capabilities to the global scale requires collaboration with CDC's international epidemiology and response team, influenza division, to acquire data from targeted countries. Predictive capabilities for seasonal influenza in major population centres are being developed. The predictive and early warning capabilities developed by this project will promote influenza surveillance, prediction and control at CDC and key public health agencies around the world (Table 1, E).

Research on malaria surveillance modelling uses NASA data, model outputs, and analytical and modelling expertise to enhance decision support capabilities for malaria risk assessment and control. The capabilities that are developed concern detection, prediction and reduction of malaria risk. Since rainfall provides vector breeding sites and prolongs vector life span by increasing humidity, anomalies in precipitation are the attribute most frequently used for predicting malaria outbreaks. However, it has also been shown that rainfall, or the lack of it, has a complex effect on malaria transmission for various parts of the world (Kovats *et al.*, 2003). For example, although moderate rainfall may promote malaria transmission, intense and prolonged rainfall may flush away larval habitats and thus reduce transmissions. Similarly, lack of rainfall does not always reduce larval populations. On the contrary, lack of rainfall may create new habitats, such as pools and puddles, in some regions and therefore increase larval population. In addition, droughts may be deleterious to predator populations or may cause human populations with no immunity to move to areas endemic with malaria (Kovats *et al.*, 2003). These factors may indirectly increase overall malaria transmissions.

Another meteorological variable that is often used for predicting malaria transmission is temperature. Warmer temperature hastens larval and vector development and therefore increases the rate of vector production (Craig *et al.*, 1999). It also shortens the time for spores to reproduce via multiple fission, which thereby produce more vectors, and more time for them to transmit malaria. In addition, warmer air holds more moisture and therefore enhances mosquito survivorship.

Relative humidity is important for the survivorship of malaria vectors. While it is not a standard remote sensing data product, relative humidity can be computed from dew point and air temperatures, which are usually provided for from satellite instruments that measure atmosphere properties.

Vegetation is often associated with vector breeding, feeding, and resting locations. A number of vegetation indices have been used in remote sensing and Earth science disciplines. The most widely used is NDVI (Tucker 1979). It is simply defined as the difference between the red and the near infrared bands normalized by twice the mean of these two bands. For green vegetation, the reflectance in the red band is low because of chlorophyll absorption, and the reflectance in the near infra-red band is high because of the spongy mesophyll leave structure. The more vigorous and denser the vegetation is, therefore, the higher the NDVI becomes.

Remote sensing instruments and geophysical parameters used in these studies include: ASTER (ground cover, vegetation index, rainfall), AVHRR (ground cover, vegetation index, rainfall), MODIS (ground cover, vegetation index, surface temperature, humidity), TRMM (rainfall), IKONOS (ground cover, vegetation index, rainfall), and SRTM (DEM).

**2.1.7 Sensor data and satellite telemetry of wild birds for decision support of avian influenza**  
This research also addresses avian influenza (H5N1 subtypes), but from a slightly different perspective. It examines the interplay between the local persistence of HPAI (an avian-adapted strain of H5N1) in poultry, and episodic long-distance dispersal by migratory wild birds, which pose serious threats to poultry production, wild birds, and human health. Geospatial technologies are needed to understand migration patterns among wild birds that spread diseases over long distances. Specifically, MODIS time series data are used to map cropping intensity (double and triple crop cultivation per year), crop calendar (planting dates and harvesting date), and paddy rice fields. PALSAR data are used to map water bodies and natural wetlands. In addition, satellite telemetry data are used to track wild waterfowl migrating from Southern China, to Siberia and Russia. This project provides updated and improved geospatial data sets for cropping intensity, paddy rice, wetlands and LST; and uses the emergency prevention system (EMPRES) model for priority animal, plant pests, and diseases to assess, forecast, and communicate risks of avian influenza.

Maps of avian influenza *hot-spots* and *hot times* provide timely support for disease surveillance, pandemic preparedness planning, and disease management and response. Another objective is to train young scientists in the era of one world/one health, including agro-ecology, wild bird biology, veterinary, epidemiology, public health, eco-informatics and geo-informatics.

#### **2.1.8 Develop an influenza forecasting system**

This research combines data on influenza activity from CDC with key environmental data sets on climate and phenology obtained by TOPS to understand environmental drivers affecting influenza transmission patterns, and the timing of peaks in influenza activity. Influenza is a contagious respiratory illness caused by viruses that affect between five and twenty per cent of the US population each year, resulting in more than 200,000 hospitalizations and 36,000 deaths on average. Influenza in temperate regions exhibits a pronounced seasonality that has been well described. A number of hypotheses have been proposed to explain this seasonality, and recent studies have demonstrated both a temperature threshold effect (in animal studies) and a significant correlation between latitude and the timing of annual influenza epidemics. These results suggest that environmental drivers play a role in influenza transmission, innate immunity, and/or virus-host interaction, and could be used to forecast the date of peak influenza activity. In particular, the average difference in peak week for southern *vs.* northern cities is approximately sixty days, which matches phenological patterns closely in the US and suggests a potential link between the two.

In recent years, absolute humidity, solar radiation, temperature and relative humidity have emerged as important determinants of the timing of seasonal influenza epidemics. To date, research results have shown that drops in absolute humidity tend to be followed by increases in influenza mortality (Shaman *et al.*, 2010). Other studies observed a relationship between latitude, solar radiation and the timing of peak influenza activity (Charland *et al.*, 2009; Finkelman *et al.*, 2007). Precipitation is also an important predictor of increased influenza incidence in tropical and sub-tropical regions. This research utilizes various statistical analyses to relate climatological information to health data. Daily estimates of average saturation vapour pressure deficit, solar radiation, maximum temperature and precipitation were obtained from TOPS. Interestingly, results show limited evidence that absolute humidity, solar radiation, precipitation and maximum temperature could add to successful forecasting of influenza. Time series models that accounted for seasonality performed as well as the models with meteorological inputs. Only peak timing appeared to be meaningfully associated with vapour pressure deficit, solar radiation, maximum temperature, but the forecast error suggested that forecasts of the peak timing could be miss-forecast by between two to three weeks on average.

Table 3. Selected NASA-funded research applications on water-borne diseases.

Principal investigator	Brief title of project
R. Stumpf & T. Wynne, NOAA	Monitoring and forecasting cyanobacterial blooms
Z. Deng, LSU	Satellite detection and forecasting of oyster norovirus outbreak
C. Tilburg, Univ. of New England	Land-use and precipitation for hydrology & public health



Figure 1. MODIS true colour image showing a bloom in the western basin of Lake Erie, August 2009.  
(see colour plate 39)

## 2.2 Data sets for water-borne disease studies

Many illnesses, contaminants, and injuries can be water-, sanitation-, or hygiene-related. Water-borne diseases are caused by organisms that are spread through water directly. Water-related illnesses can be acquired through an absence of water for good hygiene, poor sanitation, or increasing insect populations that breed in water and then spread disease (Table 1, F). The projects discussed in Table 3 describe how the addition of remote sensing data has aided in detecting and forecasting water-borne diseases.

### 2.2.1 Monitoring and forecasting cyanobacterial blooms

This project addresses cyanobacterial harmful algal blooms (CyanoHABs), a global problem affecting public and environmental health (see also Chapter 10). The project is developing a modelling and forecasting system (MFS) to identify, document, and forecast CyanoHABs in major water bodies to aid environmental and public health managers in planning responses. Satellite sensors are used to collect data on ocean colour to characterize the amount of live phytoplankton biomass and turbidity data, identify algal blooms, and track plumes (Table 1, G). This project combines environmental, meteorological, and disease surveillance data with colour and temperature satellite data from MODIS and VIIRS. Ocean colour sensors provide large synoptic sampling otherwise impossible to collect with standard *in situ* techniques that may be useful for identifying harmful CyanoHABs (Kahru 1997; Kutser 2004; Wynne *et al.*, 2008). Once an algal bloom has been identified, a method must be in place to track the bloom's position and to compensate for missing surface data caused by cloud cover (Figure 1). The nowcast is created using imagery that is generally older than one day. Wynne (*et al.*, 2011) discusses this further and provides detailed methodologies for sensor data, hydrodynamic models and particle tracking. The MFS capability includes: 1) routine bloom identification; 2) forecasting CyanoHAB events; and 3) transferring results to health and

environmental management agencies. Successful implementation should reduce monitoring costs, improve water management practices, and reduce public health impacts.

#### 2.2.2 Satellite-assisted detection and forecasting of oyster norovirus outbreak

This project utilises Terra and Aqua MODIS data to detect and forecast oyster norovirus outbreaks in coastal Louisiana. The strategy is to combine environmental data with *in-situ* bacteriological data from field samplers and laboratory analysis. The system consists of a series of retrieval algorithms or water quality models that link MODIS data to water quality indicators controlling norovirus disease outbreaks, an artificial neural network model for detecting and forecasting faecal coliform (norovirus indicator organism), and a hierarchical Bayesian model for detecting and forecasting norovirus disease outbreak risks.

The system enables shellfish managers to make two types of decisions: detection management decisions (open/close) and forecasting management decisions (classification and reclassification). The decision management capability makes it possible to reduce decision making time from its current time lag of two to four months, to one day. The forecast management function predicts oyster norovirus outbreaks in a probabilistic fashion. The system is essential to classify/reclassify oyster growing areas and for long-term planning and sustainable oyster management.

#### 2.2.3 Influence of land-use and precipitation on regional hydrology & public health

The central objective of this project is to determine if a relatively simple regression model can be used to predict water quality along the coast of Maine. In the north eastern US, climate change scenarios typically indicate an increase in overall precipitation, which would lead to larger river discharge. Increased discharge would then lead to greater pathogen loading in rivers and coastal waters from anthropogenic sources due to runoff from contaminated sites. To achieve the project's objective, EO and *in-situ* data of the Gulf of Maine watersheds, estuaries, and coastal ocean are being used in simple regression models. Observations for the project were collected over a twelve year time period from TRMM's TMI sensor designed to measure rainfall over the swath under the TRMM satellite. TRMM is able to distinguish precipitation variations at a resolution of 0.25°C. In addition, imagery from Landsat is being used to quantify changes in LCLU in the modelled region. River data from USGS and other sources allow hydrology and nutrient discharges to be quantified. Water quality data obtained from state agencies allow models to be calibrated and enable predictions of water quality for these same state agencies. Information from the models is used to determine the feasibility of this approach to climate scenarios.

The most common method for testing water for pathogens is to measure the concentration of *E. coli* and total coliforms, bacteria that are normally present in the intestinal tract of humans and other animals and that are used as indicators for recent faecal contamination. Since the State of Maine does not currently have the ability to test for *E. coli* and total coliforms quickly, they use river discharge as a simple proxy for water quality. Maine's current method relates low water quality events to high measured discharge. High discharge events would trigger increased run-off and low water quality. Water resource managers are particularly interested in the ability of a method to predict low water quality events but also to minimize false alarms or times when a method predicts a low water quality event that does not materialize.

Bulk measurements of discharge are not able to differentiate between run-off from high pollution areas and run-off from less-developed, pristine areas. A method that identifies where run-off occurs would likely predict reduced water quality events more accurately. A precipitation-based method that uses simple linear regression of observed precipitation and observed water quality shows promise for calibrating the model. To compare the effectiveness of the precipitation-based method and Maine's current method, the percentage of actual reduced water quality events that were predicted by each method and the percentage of predicted reduced water quality events that preceded actual events have been examined. The precipitation-based method predicted low water quality events more accurately, and with fewer false alarms. Models developed by this project are determining that accurate measurements of precipitation, land-use, and river discharge can be

Table 4. Selected NASA-funded research applications on air-borne diseases.

Principal investigator	Brief title of project
Y. Liu, Emory Univ.	Tracking and modelling particle exposure for health/epidemiology
J. Luval, MSFC	Air-borne pollen prediction for asthma alerts and health decisions
L. McClure, Univ. of Alabama	Environmental data linked to cohort data for health decisions
S. Morain, Univ. of NM	Adding EO data to EPHTN via the NM/EPHT system
A. Huff, Battelle Mem. Inst.	Using AOD data to create PM <sub>2.5</sub> fields for health and epidemiology

used to develop regression models to achieve accurate predictions of water quality and that can be expanded to larger temporal and spatial scales.

### 2.3 Data sets useful for air-borne diseases, air quality, and health

Air pollution affects health in many ways with both short-term and long-term effects. Examples of short-term effects include irritation to the eyes, nose, throat and upper respiratory infections such as bronchitis and pneumonia. Short-term air pollution can aggravate individuals suffering with asthma and emphysema. Long-term health effects can include COPD, lung cancer, heart disease, and even damage to the brain, nerves, liver, or kidneys. Chronic exposure to persistent organic pollutants (POPs) affects the lungs of growing children and may aggravate or complicate medical conditions in the elderly (LBL 2011). Earth observation data are advancing scientific understanding of the effects of elevated concentrations of POPs, and in tracking their movements for more informed policy decisions. Table 4 lists five projects in NASA's 2011 portfolio for air-borne diseases.

#### 2.3.1 Environmental public health tracking using particle exposure models and epidemiology

As one of six ambient air pollutants, fine particulate matter (i.e. air-borne particles less than or equal to 2.5 micrometres in aerodynamic diameter-PM<sub>2.5</sub>) has been linked to various acute and chronic health problems. Because satellite aerosol remote sensing expands the coverage of PM<sub>2.5</sub> monitoring to rural and suburban areas, it is important to examine how satellite aerosol remote sensing might extend CDC's coverage of fine particulates in the national environmental public health tracking network (EPHTN), and specifically, the utility of satellite-derived air quality estimates.

This project is developing spatial statistical models that integrate aerosol information from: 1) MODIS level-2 aerosol product; 2) the GOES aerosol and smoke product (GASP); 3) the MISR level-2 aerosol product; 4) the OMI aerosol index; 5) meteorological observations; 6) land use information; and, 7) EPA PM<sub>2.5</sub> measurements. The aim is to produce a spatially resolved daily PM<sub>2.5</sub> concentration surface. These data sets with different spatial resolutions are all reprojected to the twelve kilometre resolution community multi-scale air quality (CMAQ) model grid and compared to EPHTN's tracking network of existing CMAQ-simulated PM<sub>2.5</sub> concentrations. In addition, satellite predictions are validated prospectively with independent field sampling. Satellite predictions are also evaluated in epidemiological models that link PM<sub>2.5</sub> levels with cardiorespiratory morbidity outcomes collected in the Atlanta, Georgia metropolitan area. Outputs include detailed analyses of the spatial and temporal patterns of PM<sub>2.5</sub> pollution in the domain, a statistical evaluation of the advantages of estimating PM<sub>2.5</sub> concentrations, and an assessment of the potential benefit of ingesting satellite observations into EPHTN.

#### 2.3.2 Air-borne dust prediction and vegetation phenology for tracking pollen episodes

This health application explores details of phenology and meteorology, and their dependencies to produce a first-generation deterministic model for predicting and simulating pollen emission and downwind pollen concentration. Pollen can be transported great distances. In fact, pollen from *Juniperus spp.* can be transported 200–600 km downwind (Van de Water *et al.*, 2003). Hence,

local observations of plant phenology may not be consistent with the timing and sources of pollen collected by air sampling instruments. Based on the NCEP non-hydrostatic meteorological model, satellite sensor data, and an *in-situ* network of phenology cameras, a real-time and rapid response pollen release and transport system is being prototyped. This prototype is based on the rapid response MODIS direct broadcast system to acquire daily data with one hour lag time, similar to the MODIS rapid response system for fire detection. Research outputs will be used to support EPHTN, which includes the State of New Mexico's environmental public health tracking system, and the syndrome reporting information system (SYRIS®) for asthma and allergy alerts and decision support.

**2.3.3 Satellite data augment a national public health cohort study for better health decisions**  
This project links three national data sets: EPA's *in-situ* network of PM<sub>2.5</sub> observations; solar insolation data from the North American regional reanalysis (NARR) network; and LST from satellite sensor observations. These data sets are merged with public health data from the reasons for geographic and racial differences in stroke (REGARDS) national cohort study, to determine whether these environmental risk factors are related to cognitive decline. Environmental data sets and public health linkage analyses will be disseminated to end-users for decision making, priority setting, program evaluation, public health research, and resource allocation through CDC's wide-ranging online data for epidemiological research (WONDER) system. For further information on this project, see Section 4.3 below.

**2.3.4 Adding air quality observations to state and national tracking systems**

This application addresses respiratory public health at both the individual and community levels by forecasting atmospheric ozone, dust, and other aerosols that trigger asthmatic responses or myocardial infarction; and, by enhancing the State of New Mexico's ability to prepare and provide early warning forecasts and alerts to populations at risk through its EPHTS. Between 2003 and 2008 a dust entrainment model based on DREAM (Nickovic 2001) was re-configured to assimilate EO data, and nested within the NCEP/eta national weather forecast model, to develop a numerical system to simulate dust entrainment, transport and deposition on hourly meteorological conditions. Several static DREAM model parameters for land surface conditions were replaced by assimilating more frequently updated measurements of land cover from MODIS, elevation from SRTM, and surface roughness length from a look-up table based also on MODIS. Replacing these parameters showed marked improvement in DREAM's performance for forecasting dust entrainment in the American southwest. The model simulates timely forecasts of hourly mineral dust-patterns, but is less successful forecasting particle concentrations (Morain & Budge 2010). Rolling hourly simulations can be converted into animated gifs to produce forty-eight to seventy-two hour forecasts of dust episodes. These animations represent an important step in verifying that satellite sensor data not only improve model performance, but also add value to daily weather forecasts. The NCEP/eta model alone does not forecast atmospheric dust; and therefore, cannot be used for intervening possible asthma interventions at the local, state, or regional levels.

After 2008, capabilities from the NCEP/DREAM model system were extended to include hourly ozone and aerosol data collected by US/EPA. The new model system was improved by sharing hourly dust loadings from NCEP/DREAM with CMAQ to enhance its ability to distinguish between atmospheric dust entrainment and fugitive dust, and to add simulations for 107 species of POPs. Two other improvements were made: 1) by replacing the MOD12Q1 product generated in 2002 with the MCD12 product generated in 2008 (Friedl 2010); and 2) by refreshing the NCEP/DREAM model with sixteen-day NDVI rolling updates of changing dust source geographies across the model domain. NDVI up-dates of dust sources are attributed overwhelmingly to agricultural practices and seasonality. These changes were assimilated into the model regularly to better quantify barren land distributions in the Southwest. To improve CMAQ outputs, the model domain for ozone and aerosol observations was reduced to accommodate a finer resolution grid cell spacing of 7.5 km instead of the original 17 km spacing. This is important in context of health applications because epidemiologists need exposure data across large areas that are then aggregated to zip code or census

tract levels, most of which have complex shapes and attendant edge-pixel problems, best managed statistically at high resolution (Morain & Budge 2010).

Creating the nested model and combining it with CMAQ ozone and aerosol data represents an engineering approach to provide data that support health tracking and decision making. The aim throughout the effort has been to verify and validate outputs, and to build archives of hourly model runs for PM<sub>2.5</sub>, PM<sub>10</sub>, and 107 anthropogenic EPA air quality contaminants. As the archives continue to grow, they support short-term episodic alerts for daily health and long term data sets for epidemiological studies of exposures and outcomes.

### 2.3.5 Using aerosol optical depth to represent PM<sub>2.5</sub> fields

This project supports environmental public health tracking systems by establishing that AOD data from satellite observations can generate PM<sub>2.5</sub> data sets. Such data sets reflect spatial and temporal variations in ambient concentrations occurring at local scales, and are representative of the true PM<sub>2.5</sub> field. Fine particulates are critical air pollutants, and their adverse impacts on human health are well established. Traditionally, studies that analyse health effects of exposure to PM<sub>2.5</sub> use data from broadly scattered ground-based monitoring stations. However, due to large spatial and temporal gaps between these point source monitors, daily synoptic satellite AOD data provide information about particulate concentrations in areas where monitors do not exist.

Using a hierarchical Bayesian model (HBM) to combine monitored data with estimates of PM<sub>2.5</sub> derived from AOD and CMAQ, this approach could represent a significant step toward creating accurate and representative PM<sub>2.5</sub> data sets that can be used to make informed public health decisions. The accuracy of the combined monitor/satellite/air quality model data sets are being determined in relation to monitor values, and their performance over data sets analogous to the current best estimates of PM<sub>2.5</sub> fields are being quantified. Environmental public health tracking programs associated with Maryland, the CDC, and EPA have expressed interest in using the results of the feasibility study to enhance their existing decision making activities.

## 3 CONVERTING ESR DATA INTO HEALTH RESEARCH AND APPLICATIONS

This Section provides data and information services for EO data that: 1) facilitate their use in health research and applications; 2) invoke new ways of thinking about environmental health tracking; and 3) stimulate new ideas about how these services advance health surveillance. Section 3.1 is an overview by Doctors Tong and Soebiyanto of atmospheric data used for health applications. This is followed in Sections 3.2 through 3.6 by contributions addressing specific uses of sensor data, and the tools necessary to facilitate their use. The first of these by Doctor Ceccato describes EO data for malaria surveillance. Doctor Maxwell then presents a tool for evaluating environmental measurements of, and responses to, extreme heat events. The third by Doctor Benedict describes an IT system for delivering dust forecasts, and the fourth by Doctors Golden and Chen describes data, tools, and services provided by the Socioeconomic Data and Applications Center (SEDAC). The last sub-Section by doctors Kempler, Lynnes, Vollmer, and Leptoukh describes health data and services from GES-DISC.

Key information requirements for epidemiology are the spatial and temporal distributions of environmental agents, their levels of concentration, and their proximity (exposure potential) to effected cohorts. Due to lack of exposure data, many epidemiological studies use measurements from stationary ground monitors as surrogates of exposures for individuals proximal to those sites (Ito *et al.*, 2005; Jerrett *et al.*, 2009). This central monitoring approach is usually oversimplified (Tong *et al.*, 2009). However, sensor data supply area-wide, synoptic values for monitoring time, space, and concentrations of environmental agents across all land, air, and water surfaces. Nevertheless, satellite sensor data have only recently entered the realm of either epidemiological studies that establish the linkage between exposure and risks, or health assessments that quantify health effects based on epidemiological knowledge. The underlying obstacles that inhibit the marriage of

these two promising partners are discussed. If these obstacles can be overcome, EO data and their delivery systems will be of great service to health surveillance practitioners.

### 3.1 Atmospheric data for health applications

Since the launch of the first satellites, efforts have been made to measure atmospheric constituents and Earth's surface characteristics from space. The Nimbus satellites, which operated from 1963 to 1993, began the age of space-based Earth observations. The Nimbus program expanded human knowledge of the upper atmosphere significantly, including the discovery of the ozone hole. Like Nimbus, most of the early satellites focused on the upper part of the atmosphere. The TOMS ultraviolet instrument is the first to observe the total ozone column as well as the tropospheric ozone column (Krueger & Jaross 1999). The GOME observes not only ozone, but also ozone precursors, such as nitrogen dioxide ( $\text{NO}_2$ ), formaldehyde ( $\text{HCHO}$ ), and bromine oxide ( $\text{BrO}$ ) that catalyse polar ozone destruction (Burrows *et al.*, 1999). SCIAMACHY sensors extended these efforts by providing observations of additional atmospheric constituents [methane ( $\text{CH}_4$ ), carbon monoxide ( $\text{CO}$ ), and carbon dioxide ( $\text{CO}_2$ )] at even finer resolution (Bovensmann *et al.*, 1999). Continuous, routine measurements of trace gas distributions are being provided by several satellite instruments, including the MOPITT/Terra (Drummond & Mand 1996), AIRS/Aqua (Aumann & Pagano 1994), TES/Aura (Beer *et al.*, 2001) and the Infrared Atmospheric Sounding Interferometer (IASI on MetOP) (Schlussel *et al.*, 2005). Meanwhile, utilizing higher spectral resolution in the limb geometry, vertical profiles of atmospheric constituents are now being observed with the Michelson interferometer for passive atmospheric sounding (MIPAS) on board ENVISAT and Atmospheric Chemistry Experiment (ACE) instruments (Bernath *et al.*, 2005; Clerbaux *et al.*, 2008; Fischer *et al.*, 2008). EO data sets expanded in both spatial and temporal resolution and the capacity for retrieving atmospheric chemistry have provided unprecedented opportunities for science data users to study Earth systems and the societal impacts of environmental changes from a space perspective.

This article reviews the current status, challenges and opportunities for applying satellite data to environmental monitoring and health surveillance. Environmental tracking is central to understanding levels of hazards imposed on society. Epidemiological studies around the world have associated pollutant exposure with adverse health effects, including both morbidity and mortality (Bell *et al.*, 2004; Jerrett *et al.*, 2009). The WHO estimated that 800,000 annual premature deaths, or 1.2 per cent of all deaths, are caused by exposure to urban fine particles (Cohen *et al.*, 2004). It has been estimated that the global health burden of air pollution is much larger than the WHO study, perhaps by as much as five-fold or six per cent of all deaths, which excluded ozone impacts and included only part of the urban population (Annenberg *et al.*, 2010). Given the profound impacts of air pollution on human health, it is critical to understand the sources, formation, and chemical characterization of air pollution, so that linkages and mechanisms from pollution sources to health endpoints can be revealed and understood before effective mitigation strategies can be designed to protect public health.

Several approaches have been developed to extend point measurements to broader areas so that larger populations can be included. Kriging, for example has been applied to the relationship between fine particle exposures and mortality risks within-city that result in considerably higher risk levels than those found in central monitoring approaches, and three dimensional air quality models such as the CMAQ are able to estimate population exposure to air pollutants at places where monitoring data are not available (Jerrett *et al.*, 2005; Bell 2006). Both methods can provide exposure estimates at additional locations, thus alleviating some limitations of the central monitoring method. These methods, however, are subject to inherent limitations; in particular, the need to validate the estimates to ensure appropriate representations of ambient concentrations (Bell 2006). The distribution of monitoring sites for surface ozone ( $\text{O}_3$ ) from the two major networks, AQS and CASTNET, reveals that there are insufficient numbers of observations to verify gross national patterns in the air quality system (Tong *et al.*, 2007; Avnery *et al.*, 2011). In most cases, across the field of environmental health, one of the most pressing challenges is obtaining accurate exposure assessment data (Patz 2005). Data for statistical and sensor studies at monitoring sites are obtained

from US/EPA. Human population data are county-level, all-age sums provided by the US Bureau of Census, and corn production data are compiled by the USDA, National Agricultural Statistics Service (NASS) (see section 3.1.1.3).

### 3.1.1 *The range of air quality applications for health*

Because of its broad spatial coverage, satellite sensor data have considerable promise for expanding air quality data from ground monitoring networks into broader urban, suburban and rural air quality contexts (Patz 2005; Liu *et al.*, 2009). A series of vignettes is provided to reprise these opportunities. While each makes a case for successful satellite data use, such data have not been used widely in either epidemiological studies that might establish linkages between exposure and risks, or in health assessments that quantify the levels of health damage based on epidemiological linkages. What are the obstacles that inhibit the marriage of two seemingly promising partners? How can science move forward to bridge the gap between satellite data and their health application?

#### 3.1.1.1 Surface levels of air pollution

In general, satellite data provide column loading of atmospheric components, while health studies mostly focus on surface concentrations at the breathing level. To bridge the gap, several algorithms have been developed to derive surface concentration from existing satellite products. A simple approach was presented in which ground-level PM<sub>2.5</sub> concentrations were estimated by applying a localized vertical profile from a global chemical transport model to aerosol optical thickness retrieved from MISR (Liu *et al.*, 2004). They found the derived MISR PM<sub>2.5</sub> concentrations to be in good agreement with the ground measurements, with a correlation coefficient  $r = 0.81$ , an estimated slope of 1.0 and an insignificant intercept when three outliers are excluded. A similar approach was applied to another region using vertical profiles from the same global model to retrieve surface PM<sub>2.5</sub> concentrations from TOMS, MODIS and MISR measurements (Hu *et al.*, 2009). Regardless of its successful application, this approach is compromised by inherent uncertainties with several key inputs, including the cloud screening for satellite products, modelled vertical profile, and model and satellite spatial resolution. A two-stage generalized additive model (GAM) was subsequently used to estimate ground-level PM<sub>2.5</sub> concentrations in the north eastern US based on land use data, meteorological data and the GASP AOD product (Liu *et al.*, 2009). This approach, while immune to several weaknesses in the atmospheric model-based approach is able to achieve a higher model performance (model correlation coefficient = 0.89), but faces its own limitation in that the two-stage GAM cannot be applied in regions where ground PM<sub>2.5</sub> measurements are too sparse in space.

#### 3.1.1.2 Long-range transport of air pollutants

The large areal coverage of satellite sensors provides a reliable means for monitoring long-distance transport of air pollutants and other exposure agents. The magnitude of transboundary air pollutant transport has long been at centre-stage in a global debate on how to effectively reduce air pollution; through domestic emission control, or background mitigation (i.e. blaming one's upwind neighbours?). For instance, a global chemical transport model was used to estimate the impact of transPacific transport of mineral dust on aerosol concentrations in North America. The authors found that transPacific sources are responsible for forty-one per cent of the worst dust days in the western US (Fairlie *et al.*, 2007). Within the borders of the US it has been estimated that for over eighty per cent of the conterminous states, interstate transport was more important than local emissions for summertime peak ozone concentrations (Tong & Mauzerall 2008). These results are all based on model calculations that cannot be verified easily or independently through traditional ground-based measurements.

Several studies have used EO data to quantify air quality and health effects of long-range pollutant transport. The data can detect transboundary transport because of their large temporal and spatial scales. Quantitative or semi-quantitative assessments of intercontinental and hemispherical transport of aerosols, for instance, have been conducted since 1970s (Fraser 1976; Lyons *et al.*,

1978; Fraser *et al.*, 1984; Herman *et al.*, 1997). However, these studies are limited by poor accuracy of earlier satellite measurements. Recent improvements in data quality and enhanced new capabilities of satellite sensors are enabling new opportunities to investigate this issue with a more robust data set (Yu *et al.*, 2008). In particular, multi-wavelength, multi-angle, and polarization measurements have provided additional information on the physical and chemical characteristics of air pollutants, such as particle size (fine vs. coarse), shape (spherical vs. non-spherical), and absorptiveness (absorptive and scattering) (Higurashi & Nakajima 2002; Tanre *et al.*, 2001; Holzer-Popp *et al.*, 2008). The new capabilities associated with passive sensors are further complemented by an increased data pool of active LiDAR data that shed light on the vertical structure of pollution distribution (Spinhirne *et al.*, 2005). Built on these recent advances, scientists at GSFC assessed the pollution flux from Asia to North America using a four-year (2002–2005) climatological archive of MODIS AOD, relative humidity and vertical distribution data from field campaigns and satellite data retrievals (Yu *et al.*, 2008). With uncertainties within a factor of two, they estimated an influx of aerosols into the western coasts of North America to be approximately 4.4Tg/yr.

Dust and aerosols not only impose adverse health effects directly, but act also as carriers for long-range transport of diseases (Prospero *et al.*, 2002). A recent study in Japan attempted to apply a satellite-based monitoring approach to measure the influx of Chinese pollutant loads and their association with bronchitis mortality (Goto *et al.*, 2010). They found no significant association by examining the relationship between the annual average amount of incoming pollutant load and annual average mortality from asthma. However, this is perhaps the first attempt to apply synoptic sensor data to epidemiological research investigating chronic mortality effects.

### 3.1.1.3 Crop exposure and ecosystem health

Environmental health and human health are inextricably intertwined. Ecosystem health is an emerging field in satellite remote sensing that aims to assess impacts of air pollution and climate change. Previous field experiments using the open top chamber (OTC) approach have shown that ambient ozone, alone or with acid rain gases, accounts for up to ninety per cent of crop losses from air pollution in the US. Crop yield loss from ozone exposure is of particular concern because surface ozone levels remain high in many major agricultural regions (Tong *et al.*, 2007; Avnery *et al.*, 2011). Quantitative assessment of ozone damage to crops is challenging because of difficulties in measuring reductions in crop yield that result from exposure to surface ozone. Ozone monitors are sparse or non-existent in major corn production areas in the US.

A number of approaches have been proposed to estimate crop exposure to ambient ozone. Most of these fall into two categories: those that use ground monitoring to determine ozone exposure through spatially extrapolated measurements at discrete locations (Adams & Croker 1989; Hertlein *et al.*, 1995; Felzer *et al.*, 2004); and those that use atmospheric chemical transport models (CTM) to calculate ozone exposure (Tong *et al.*, 2007; Avnery *et al.*, 2010). The major limitation of the monitoring approach is that all extrapolation techniques, such as the widely used kriging, cannot capture the actual spatial variability in ozone concentrations perturbed by chemical and physical processes. The modelling approach alleviates this limitation by accounting for all underlying processes that control ozone variations. Using this approach, annual approximate losses of ten per cent for soybean yields in the US due to ozone exposure were reported (Tong *et al.*, 2007). Globally, ozone exposure may be reducing soybean, wheat, and maize yields by as much as 8.5–14 per cent, 3.9–15 per cent, and 2.2–5.5 per cent, respectively (Avnery *et al.*, 2010). It should be pointed out, however, that modelling approaches have their own limitations, such as the uncertainties in parameterization and the need of independent verification of the implementations of these complex processes.

A recent study by Fishman and colleagues is the first to utilize both ground and satellite ozone measurements to verify crop yield losses from ozone exposure (Fishman *et al.*, 2010). By using a multiple linear regression model, they found that soybean crop yields during a five year period across the mid-section of the US were on the order of ten per cent less, consistent with earlier modelling by Tong (*et al.*, 2009). Interestingly, their research demonstrates that space-based

measurements provide a means for quantifying crop losses that could be employed on a global basis. They also argue that satellite data may even be a better measure of ozone amounts outside of urban areas because of the general paucity of surface sites in predominately farmland regions (Fishman *et al.*, 2010).

#### 3.1.1.4 Linking human ecosystems to infectious diseases

Human-induced land use changes have had complex impacts on human health, mainly by altering food supply, shelter, and sanitation that in-turn have unintended health consequences (Patz *et al.*, 2004; Patz, 2005). Rural and urban developments, including roads, dams, and dwellings, all modify transmission of infectious diseases. In tropical regions, irrigation increases schistosomiasis and malaria by supplying habitat and breeding sites. Similarly, construction of hydroelectric dams can lead to mosquito proliferation and subsequent infectious diseases like filariasis and elephantiasis (Thompson *et al.*, 1996). An important feature of human ecosystems is that they be designed to withstand severe natural hazards and mitigate injuries and causalities (Glantz & Jamieson 2000). Hurricane Katrina demonstrated the combined effects of urban infrastructure failures and extreme weather by claiming 1,836 lives and causing widespread property destruction (Knabb *et al.*, 2006).

In the 1990s NASA's Ames Research Center (ARC) executed several projects in collaboration with New York Medical College and Yale School of Medicine to develop sensor-based models for Lyme disease transmission risk in New England (Beck *et al.*, 2000). One project used canine seroprevalence rate (CSR) as a surrogate of human exposure risk and compared municipality level CSR with Landsat TM data to examine relationships between Lyme disease transmission rate and tick bites on dogs near their owners' property (Dister *et al.*, 1993). A similar approach was employed to map relative adult tick abundance on residential properties based on TM-derived indices of vegetation greenness and wetness, assuming a relationship between forest patch size and white-tailed deer populations that are the major host of the adult tick and its primary mode of transportation (Dister *et al.*, 1997; Beck *et al.*, 2000).

#### 3.1.1.5 Linking weather and climate to diseases

The impact of weather on disease progression and spread has long been recognized since the beginning of medical science by Hippocrates (NRC 2001). However, the underlying mechanism in which weather and climate variability change the spatiotemporal dynamics of disease transmission remains poorly understood. Advances in technology and modelling techniques in the past few decades have enabled satellite meteorological data to be used to understand such phenomena in a more quantitative manner. In fact, a number of successful uses of EO data in disease spread have led to developing early warning systems that allow public health agencies to monitor, prevent and control diseases (Grover-Kopec *et al.*, 2005; Anyamba *et al.*, 2009; Witt *et al.*, 2011). In addition, with the surge of evidence on the impacts of climate variability and climate cycles on health, EO data have become even more popular in recent years. Elevated average temperature and frequent extreme heat events increase the risk for heat-related mortality and complicate chronic illnesses such as cardiovascular diseases. They also will expand the disease's geographic range, especially for vector-borne and zoonotic diseases such as dengue and WNV, where the vectors thrive under warm temperature and adequate rainfall. Satellite sensor observations can provide health organizations with decision tools for mitigating future risks and possibly preventing or controlling diseases.

In vector-borne and zoonotic diseases, environmental conditions play critical roles in propagating the vector populations. Vector abundance determines the number of new disease cases, and hence is important for controlling vector populations as a means for preventing disease (Hay *et al.*, 1997). Studies have indicated that the rate of mosquito egg development depends on temperature, and to a lesser extent relative humidity (Hoshen & Morse 2004). Rainfall also impacts mosquito breeding sites. Pools and puddles from rain can create breeding sites, but heavy, intense and prolonged rainfall may flush away larval habitats (Kovats 2003; Kiang *et al.*, 2011). Vegetation is another factor that often is correlated with mosquito breeding, feeding and resting sites. Remote sensing measurements of temperature, rainfall and vegetation are used frequently to study the spread of

mosquito-borne diseases and to identify areas at risks, especially for malaria (Craig *et al.*, 1999; Brooker *et al.*, 2006; Rahman *et al.*, 2006; Gomez-Elipe *et al.*, 2007; Kelly-Hope *et al.*, 2009; Adimi *et al.*, 2010; Haque *et al.*, 2010). For RVF, Linthicum (*et al.*, 1999) coupled satellite NDVI data with Pacific and Indian Ocean SST anomalies to predict outbreaks in East Africa up to five months in advance. Surface temperature has been used successfully at one kilometre<sup>2</sup> resolution from TOPS, along with mosquito field data to assess WNV risk in California (Nemani *et al.*, 2009; Barker *et al.*, 2010). Currently there are fewer applications of EO data for studying dengue vector transmission (Harrington *et al.*, 2001). This is due to the intrinsic nature of the dengue vector mosquito that is capable of breeding within a house or in an urban area where environmental conditions are difficult to measure remotely. However, Hales (*et al.*, 2002) have shown that the current geographical limits of dengue can be modelled using vapour pressure data. Global circulation model projections from the IPCC were further used to estimate the world population at risk under climate change scenarios.

In addition to vector-borne diseases, EO data have been used to model the spread of seasonal and pandemic influenza. Influenza's spatiotemporal dynamics have been observed to vary historically with latitude, which further suggests the role of weather and environmental factors. Several experimental studies have in fact indicated that influenza virus survival depends on temperature, and that transmission effectiveness is subjected to variability in temperature and humidity. Soebiyanto (*et al.*, 2010) used satellite derived LST and rainfall, in combination with ground station data, to predict influenza cases in Hong Kong, New York City and Maricopa County in Arizona with reasonable accuracy (see Chapter 2). It has also been shown that satellite solar radiation estimates are related significantly to the timing of influenza epidemic in thirty-five cities in the US (Charland *et al.*, 2009).

Another disease application that has shown promise for EO applications is meningitis, which is epidemic in Africa. Since epidemics seem to start during the dry season and end at the onset of rainfall, factors such as low absolute humidity and dusty atmospheric conditions have been associated with it. NDVI, cold cloud duration and rainfall estimates were used along with aerosol index, soil and land cover types to predict the outbreaks (Thomson *et al.*, 2006).

### 3.1.2 Remaining challenges

Although there are several successful applications, the mainstream epidemiological community is just beginning to use EO data as a component of their research. In regulatory applications that estimate costs and benefits of pollutant control strategies, satellite data have not yet been used to cover a large population, as required by national scale assessments. Currently, computer models and ground monitoring data are often combined to measure regional or national scale exposure levels. So what are the challenges facing more routine use of satellite data in health applications where spatial coverage plays an essential role?

There seem to be three major obstacles: uncertainty, accessibility and data quality. Epidemiological studies of the health effects of air pollution are an example. From a human exposure perspective, the best exposure data are those that represent the dosage, or to a less rigorous extent, the pollutant level in the microenvironment immediately surrounding the studied cohorts. Because such data are difficult to obtain, most large scale epidemiological studies use the ambient concentration as the exposure surrogate. Although there are many complicating factors when such a surrogate is used, this approximation has been widely employed in epidemiological studies. Satellite data, however, cannot provide surface level concentrations of air pollutants directly. Current approaches have been to obtain surface level atmospheric aerosols (Liu *et al.*, 2004; Morain 2011), but the results are yet to be used in epidemiological or assessment studies, most possibly because of the large uncertainties in the input data and processing algorithms.

Another challenge is the accessibility of satellite data for environmental epidemiologists. With the proliferation of satellite platforms, the task of using a combination of EO data sets has become intimidating for users lacking the expertise for analysing them (Patz 2005). This includes understanding that different satellite platforms measure different properties and phenomena of Earth's surface and its atmosphere in varying temporal and spatial resolutions (Kaufman *et al.*, 2002). In

most cases, satellite data need to be further processed, modelled, and interpreted to match health data, or be in GIS formats that are compatible with existing tools used routinely by epidemiologists. This challenge calls for developing data sets and tools that consider the special requirements of health studies and make the needed data accessible conveniently to users having little or no remote sensing background.

Beside accessibility, quality is also a major concern for EO data used in health studies. Health studies put high priorities on temporal resolution because of their fundamental research approaches. For instance, epidemiologists often use the time-series method to investigate the acute effect of environmental factors on a concerned health endpoint (Bell *et al.*, 2005). Satellite data, when used in such studies have to be associated with detailed data quality information, so that users can interpret their results properly. In other cases, such as cohort studies, long-term average data are used to derive the accumulated effect of environmental factors (Pope *et al.*, 2002; Jerrett *et al.*, 2009). Monthly averages are mostly aggregated from all valid data that extend from full temporal coverage to those having large data voids because of cloud cover or other quality issues. A careful study of the data quality section in metadata records is needed, so that proper caution can be exercised to avoid misinterpreting subsequent scientific analysis.

Finally, efforts are needed to reach out to the health community. Since 1985, NASA has hosted and participated in a series of workshops to solicit input from various health communities on potential applications of EO data in epidemiology, infectious diseases, and ecosystem health. These experiences and a review of literature (Beck *et al.*, 2000) have not lead to consensus regarding universal requirements for a remote sensing system in health applications. For example, environmental epidemiologists rely heavily on ground-based monitor data, which are considered ground verification. However, the narrow spatial and temporal coverage of monitoring sites imposes many limitations on these studies (Tong *et al.*, 2007; 2009). A system that is able to integrate information on pollutants, weather and other confounding factors will enhance risk assessment greatly (Patz, 2005).

### 3.1.3 Conclusion

It is well recognized that satellite-acquired environmental measurements hold great promise for epidemiologic applications. Limited by available measurements, many epidemiological studies rely on central monitor approaches to estimate exhort (i.e. force or impel in an indicated direction) exposures in the same area, which also bear their own inherent limitations. Because of their broad spatial coverage and continuous operations, satellite data hold great promise for alleviating some of the spatial data continuity limitations recognized by health communities. Numerous community efforts are being made to address these issues. It is expected that the future will see a further increase in health applications nurtured by more satellite products, reduced uncertainties, as well as user-oriented data services.

## 3.2 Using EO data for malaria surveillance

Major human diseases like malaria and dengue are sensitive to inter-seasonal and inter-decadal environmental and climatic changes (Thompson *et al.*, 1996). Monitoring variations in surface conditions such as temperature, rainfall and vegetation helps decision makers at Ministries of Health to assess the risk levels of malaria epidemics. The International Research Institute (IRI) for climate and society has developed products based on EO data to monitor these changes and provide information directly to decision makers. The mission of the IRI is to help societies to understand, anticipate and manage the impacts of seasonal climate fluctuations to improve human welfare and to better monitor environments, especially in developing countries. The URLs for this section are presented in Table 5.

### 3.2.1 The role of climate

Given its impact on populations and the gravity of its pathology, malaria remains one of the most significant infectious diseases. It is essentially an environmental disease since the vectors require specific habitats with surface water for reproduction and humidity for adult mosquito survival.

Table 5. URLs referenced in section 3.2.

Product/service	URL (accessed 18th January 2012)
A. MEWS	<a href="http://iridl.ldeo.columbia.edu/maproom/Health/Regional/Africa/Malaria/MEWS/">http://iridl.ldeo.columbia.edu/maproom/Health/Regional/Africa/Malaria/MEWS/</a>
B. IRI data library	<a href="http://iridl.ldeo.columbia.edu/SOURCES/USGS/LandDAAC/MODIS">http://iridl.ldeo.columbia.edu/SOURCES/USGS/LandDAAC/MODIS</a>

The development rates of both the vector and parasite populations are influenced by temperature. In sub-Saharan Africa the pattern of malaria transmission varies markedly from region to region, depending on climate and biogeography. The association between rainfall and malaria epidemics has been recognized for many decades (Christophers 1911), but while increasing precipitation may increase vector populations in many circumstances by increasing available anopheles breeding sites, excessive rains may also have the opposite effect by flushing out small breeding sites, such as ditches or pools (Fox 1957), or by decreasing the temperature, which in regions of higher altitude can hinder malaria transmission.

Temperature also plays an important role in the variability of malaria transmission. The development rate of mosquito larvae and the malaria parasite within the mosquito host is regulated highly by temperature. It is also one of the factors that influence the survival rate of mosquitoes. Generally, mosquitoes develop faster and feed earlier in their life cycle and at a higher frequency in warmer conditions. The plasmodium parasite in the mosquito also multiplies more rapidly at higher temperatures (Gilles 1993; NRC 2001). Humidity impacts the survival rate of the mosquito as well. Mosquitoes will generally not live long enough to complete their transmission cycle where relative humidity is constantly less than sixty per cent (Pampana 1969; Gilles 1993; NRC 2011).

### 3.2.2 Monitoring rainfall

In the majority of countries in Africa, monitoring rainfall is a major problem because there are too few rain gauges and very sparse coverage. Therefore, it is necessary to use rainfall estimations derived from satellite measurements. Using rainfall estimate products updated approximately every ten days through the Africa data dissemination service (ADDS), IRI has developed a web-based Malaria early warning system (MEWS) interface that enables users to gain a contextual perspective of the current rainfall season by comparing data to previous seasons and recent short-term averages (Table 5, A) (see Chapter 9 for details). The interface is in the IRI data library (Table 5, B) and takes the form of an online interactive map. It displays the most recent decadal rainfall map (Figure 2) over which national and district administrative boundaries and the epidemic risk zone can be overlaid, in this case as a guide rather than an absolute mask that excludes districts of local interest).

These visual features can be toggled on or off and the user can zoom in to any region for more clarity. In addition, the map can be downloaded in different formats compatible with common image analysis and GIS software such as ArcView® developed by Esri or HealthMapper developed by WHO (WHO 2012).

Decadal rainfall can be averaged spatially over a variety of user-selected areas, including administrative districts at 11 km<sup>2</sup>, 33 km<sup>2</sup>, 55 km<sup>2</sup> and 111 km<sup>2</sup> grids. By selecting a sampling area and a specific location of interest, four time-series graphs are generated by clicking the map at the area selected. Graphs are generated to enable an analytical comparison of recent rainfall with respect to that of recent seasons and the long-term series. A description of the time-series figures, the data used and their source is also provided (Grover-Koppe *et al.*, 2005). A newer version of this capability using the CPC morphing technique (CMORPH) developed by NOAA-CPC is scheduled to replace the current product with one that exhibits better agreement with field measurements (Joyce *et al.*, 2004; Dinku *et al.*, 2007).

### 3.2.3 Monitoring vegetation and water bodies

Vegetation type and growth stage play critical roles in determining mosquito abundance, irrespective of associated rainfall. The type of vegetation that surrounds breeding sites, and thereby provides

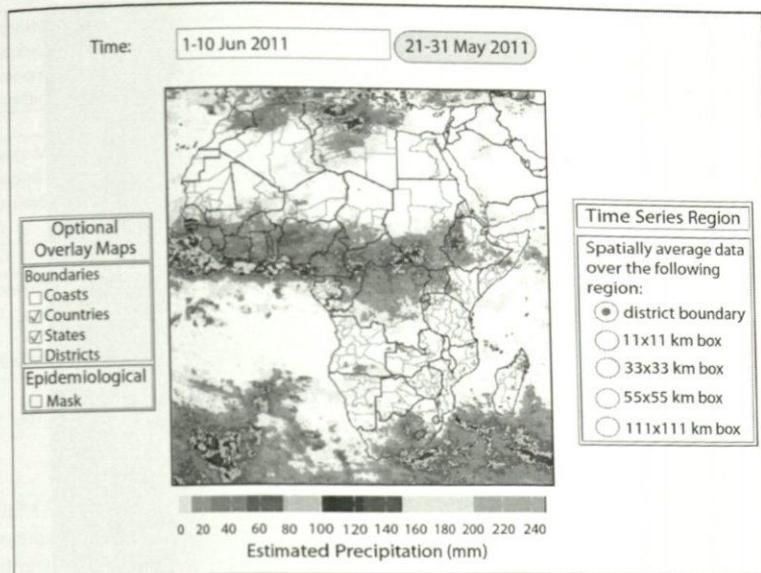


Figure 2. MEWS *clickable map* for rainfall monitoring 1–10 June 2011. (see colour plate 40)

potential resting and protection from solar radiation and desiccation, are important in determining mosquito abundance (Beck *et al.*, 1994). Surface water provides the habitat for the juvenile stages of malaria vectors (egg, larvae, and pupae). Monitoring the state of water bodies and wetlands is therefore important to identify sources of malaria vectors. To monitor vegetation and water bodies, high resolution MODIS data obtained at 250 m spatial resolution are optimal because they provide frequent observations and are available at no cost. Sixteen-day L3 Global 250M SIN GRID V004 vegetation indices can be downloaded from the USGS land processes DAAC and provided to users via the IRI data library. Users can download either the raw data as single spectral channels in blue, red, NIR and SWIR wavelengths, and in different formats compatible with common image analysis and GIS software, or as NDVI and EVI (Huete *et al.*, 2002). Using the online IRI data library, users can create a variety of secondary products. For instance, they can: 1) combine the different NIR-SWIR channels to generate tailored spectral indices for monitoring vegetation status in terms of moisture content (Ceccato *et al.*, 2002); 2) visualize a colour composite of the SWIR-NIR and Red channels (Red-Green-Blue) where vegetation appears in green, bare soil in brown and water in blue; 3) integrate the colour composite into GIS software with ancillary data such as roads and villages, as shown in Figure 3; 4) extract weighted averages of the different indices per GIS layers such as district boundaries or several other shape files; and 5) create long-term series of vegetation indices. Products such as these are already being used operationally by Ministries of Health for malaria control. They allow users to forecast the risk of malaria epidemics.

### 3.3 Internet-based heat evaluation and assessment tool (*I-HEAT*)

Over the past two decades heat waves have been responsible for more deaths in the US than any other natural hazard (Borden & Cutter 2008; NWS 2012). In 1995, at least 700 deaths were attributed to a heat wave in Chicago (Semenza *et al.*, 1996). Western Europe experienced an unusually intense heat wave in summer, 2003 that resulted in over 70,000 heat-related deaths (Robine *et al.*, 2008), and Paris alone attributed 4867 deaths to this heat wave (Dousset 2009). A report by the IPCC concluded that climate warming is unequivocal and most Earth scientists agree that global warming continues as the interglacial age deepens. Health professionals need to plan and prepare

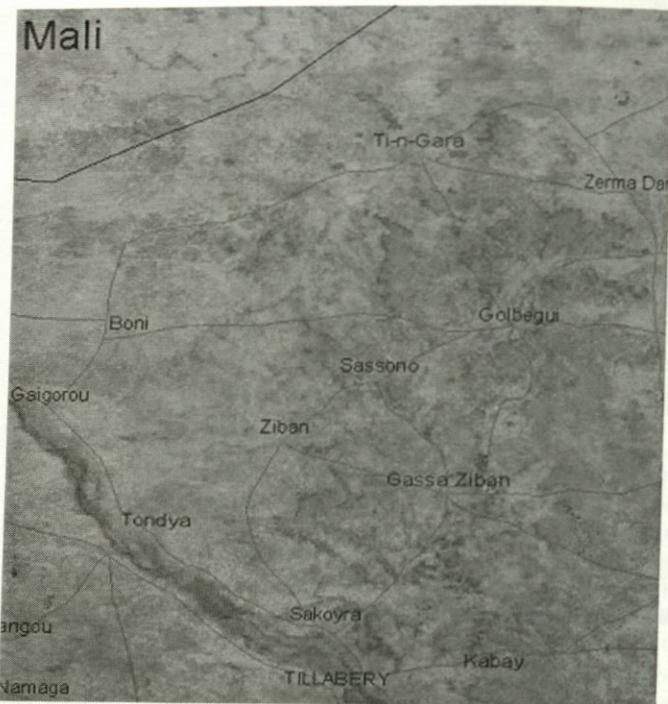


Figure 3. TERRA-MODIS colour composite RGB where the SWIR channel is displayed in red, the NIR channel in green and the visible red channel in blue. This image was acquired during the rainy season in north Niamey, Niger. It is roughly 200 km<sup>2</sup> in area. Small water bodies in blue are breeding sites for mosquitoes and locations where nomads water their cattle. (see colour plate 41)

for heat-related emergencies because future extreme events are expected to increase the frequency, duration and magnitude (Meehl & Tebaldi 2004; IPCC 2007).

The US global change research program (USGCRP) identified mapping and modelling tools for identifying populations vulnerable to climate changes as an important need. A report published by the Trust for America's Health Foundation concluded that special efforts should be made to address the impact of climate change on at-risk and vulnerable communities. The report recommends that all state and local health departments include public health considerations as part of their climate change preparedness. Yet, the report noted, only five States in the US developed a strategic climate change plan that included a public health response, and effective software tools are almost entirely lacking. Perhaps this is because there is less perceived health risk from gradual climate warming than from more immediate extreme weather events.

Impacts of climate variability on human health vary by region, demographics, duration of exposure and ability to adapt or cope with the changes (Ebi & Semenza 2008; English *et al.*, 2009). Elderly populations over sixty-five years, for example, are particularly vulnerable to the effects of extreme heat in part because of physiology, disabilities and medications (Semenza *et al.*, 1996; Ebi & Meehl 2007; Kovats & Hajat 2008; Schifano *et al.*, 2009). To identify populations vulnerable to heat waves, demographic data such as age and poverty status need to be linked with local-scale environmental data such as land cover and temperature (English *et al.*, 2009; Rinner *et al.*, 2010). An extensive review of heat-wave impacts on human health concluded that geospatial and remote sensing technologies provide a key capability for mitigating heat-related deaths. In particular, understanding the dynamics between urban heat islands and vulnerable populations is critical for reducing heat-related deaths (Wilhelmi *et al.*, 2004).

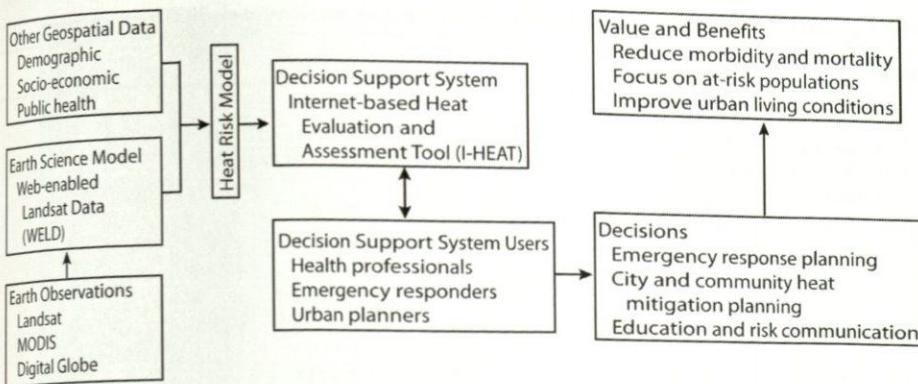


Figure 4. I-HEAT system diagram.

### 3.3.1 I-HEAT system description

The Internet-based heat evaluation and assessment tool (I-HEAT) is a new software system to provide health professionals and risk assessors with advanced geospatial web-based tools for preparing and responding to emergency heat events, developing mitigation strategies, and educating the public. The system couples demographic and environmental data obtained from Landsat data and imagery with browser-based software to provide health professionals with a tool to model and map heat-related morbidity and mortality risks at the neighbourhood level (Figure 4). Landsat data will be integrated with demographic, socio-economic, and health data in a heat-risk model that incorporates elements of the vulnerability mapping (Table 6).

Local surface temperature derived from EO digital data is a significant predictor of heat-related mortality (Vandentorren *et al.*, 2006; Smargiassi *et al.*, 2009). Landsat provides these data at 30–120 m spatial resolution and temporal repeat cycles of eight to sixteen days for modelling heat risk in urban environments (UHI 2010). Landsat data are used routinely by environmental scientists, yet their potential has not been fully realized by health practitioners, perhaps because of the difficulty and expense of integrating the data, either by fusing or assimilating them into applications by non-technical users. Only a few studies have demonstrated the utility of these data in heat risk modelling (Vandentorren *et al.*, 2006; Smargiassi *et al.*, 2009).

Although Landsat data are available at no cost through the USGS Earth Resources Observation and Science (EROS) Centre, the data are not provided in a format that can be imported and used easily without intermediate analysis by analysts with access to high-end computing facilities. Products from Web-Enabled Landsat Data (WELD) are needed to integrate them into user-friendly software applications. WELD provides a nationally consistent data source that has been temporally composited, mosaicked, and converted to geo-physical and bio-physical products (Roy *et al.*, 2010). One of the initial goals of the I-HEAT project is to determine the feasibility of using land surface temperature and vegetation measures from WELD for heat risk assessment.

Managing and processing geo-spatial and image data requires technical expertise that often is not available in many health organizations. I-HEAT will be designed for use by health officials who require a user-friendly tool to access and analyse WELD and other geospatial data used in Table 6. Geospatial mapping and modelling can be intensive computationally, especially if satellite image data are incorporated. Web-based software tools enable users to access and process these large data sets without acquiring and maintaining computer resources in-house. They also stimulate better collaboration between, and among health organizations to ensure accessibility to the software and data security. In addition, developers can maintain and update software easily to meet various user requirements in a single secured environment, thus reducing the burden on health professionals to maintain desktop applications.

Table 6. Selection of variables in a heat risk model (adapted from Reid *et al.*, 2009).

Category/variable	Source
Land cover	
Vegetation greenness (NDVI)	Landsat (WELD)
Land surface temperature	Landsat (WELD)
Demographic	
Per cent population below the poverty line	Census
Per cent population with less than a high school diploma	Census
Per cent population of a race other than white	Census
Per cent population living alone	Census
Per cent population $\geq 65$	Census
Per cent population $\geq 65$ living alone	Census
Per cent population $\geq 65$ single	Census
Diabetes prevalence	
Per cent population ever diagnosed with diabetes	Behavioural Risk Factor Surveillance System
Air conditioning	
Per cent households without central AC	American Housing Survey
Per cent households without any AC	American Housing Survey

I-HEAT is designed initially for Detroit, Michigan. Detroit lacks a heat wave health warning system and has significant racial/ethnic and socio economic disparities in heat exposure and heat-related health effects (O'Neill *et al.* 2003; Schwartz 2005), and climate change scenarios suggest the area will experience high temperature increases (Meehl & Tebaldi 2004). Surveys to assess the efficacy of I-HEAT include vulnerable residents, local governmental officials and community leaders to quantify perceptions of heat risk, and to identify the prevention and intervention programs needed. Software for I-HEAT is designed for two user groups: those monitoring heat risks to develop response strategies; and those modifying the underlying risk model. Both groups receive displays of a web-delivered map that provides an intuitive representation of geographic patterns (Figure 5). Street maps and satellite images are provided for easy reference. The design allows users to switch easily between the model and results display. The primary data displayed are at urban residential scales so users can view underlying relationships among risk model variables: surface temperature, land cover, and demographic factors. The modelling interface allows user to adjust coefficient weightings to compare different outcomes heuristically.

### 3.3.2 Summary

Over the past two decades heat waves have resulted in many deaths in the US and globally. Extreme heat events are expected to rise in frequency and metropolitan and regional health authorities need to prepare for the needs of vulnerable populations. I-HEAT will enable these authorities to prepare and respond more effectively to heat emergencies by planning and targeting intervention programs. The system supports queries such as: show residential areas with low air conditioning availability, high proportion of elderly populations, and daytime temperatures during heat events in excess of 90°F (32°C); and, where should first responders be based during heat events to better serve vulnerable populations? It is anticipated that future I-HEAT developments will incorporate real-time heat warning capabilities.

### 3.4 An IT system for interoperable multi-resolution dust modelling

Environmental forecast products are valuable to public health professionals for developing alerts and planning resource allocations. Their value is in direct proportion to the following characteristics of the modelling system: 1) timeliness of the forecast products, particularly in terms of lead time to identify future events; 2) spatial resolution of forecasts relative to county, zip-code, census tract or

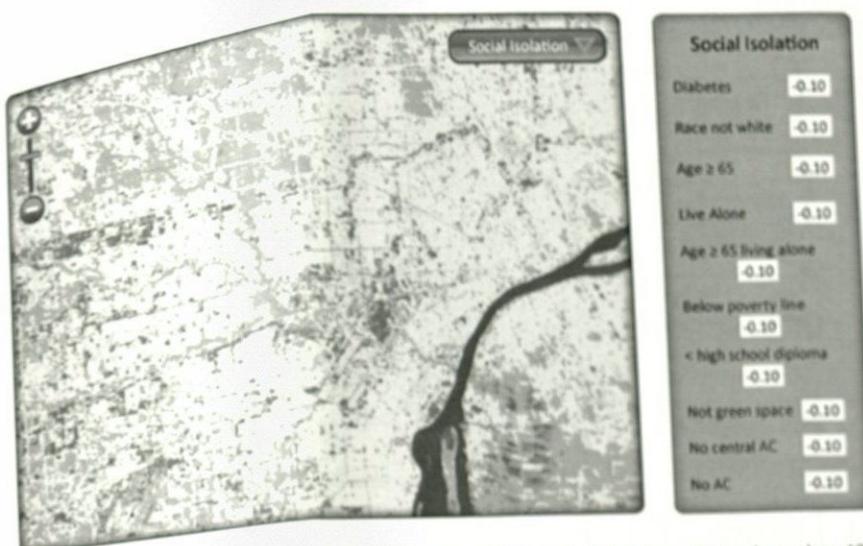


Figure 5. The I-HEAT interface showing a heat-risk map of Detroit, Michigan. (see colour plate 42)

Table 7. URLs referenced in section 3.4.

Product/ service	URL (all accessed 18th January 2012)
A. PHAiRS	<a href="http://phairs.unm.edu">http://phairs.unm.edu</a>
B. SOAP	<a href="http://www.w3.org/standards/techs/soap#w3c_all">http://www.w3.org/standards/techs/soap#w3c_all</a>
C. OGC web coverage service	<a href="http://www.opengeospatial.org/standards/wcs">http://www.opengeospatial.org/standards/wcs</a>
D. Global Forecast Service	<a href="http://www.nco.ncep.noaa.gov/pmb/products/gfs/">http://www.nco.ncep.noaa.gov/pmb/products/gfs/</a>
E. OGC web mapping service	<a href="http://www.opengeospatial.org/standards/wms">http://www.opengeospatial.org/standards/wms</a>

block group that are the preferred analytical and alert units for public health surveillance; 3) forms and formats of products that have maximum utility for public health system users; and 4) accuracy of the timing and location of events. This section describes a feasibility study for developing a dust forecasting system having improved system performance for items 1–3, and establishes a framework for streamlined evaluation of item 4, all while building upon dust forecasting model components developed over the past decade of research and application development projects. Two health application projects contributed components to this environmental forecast model: the public health applications in remote sensing (PHAiRS) project; and the interoperability and high-performance computing testbed project to expand the capabilities of PHAiRS model components. Table 7 is a list of URLs referenced in this sub-Section.

The PHAiRS project (Table 7, A) included the University of Arizona, the University of Malta, and the University of New Mexico, to produce an enhanced version of DREAM (Nickovic *et al.*, 2001) for the south western US. Resulting model outputs were made available to public health surveillance personnel and decision makers through a set of web services that deliver data via simple object access protocols (SOAP) (Table 7, B). The interoperability and high-performance computing (HPC) testbed advanced the products and services developed through PHAiRS by expanding the number of modelled particle size bins from four to eight, and by migrating from DREAM/eta to HPC's version of DREAM/nmm. The 8-bin DREAM/eta and DREAM/nmm models are the core systems for developing the interoperable, multi-resolution IT system described here.

The project's goal was to perform a test implementation of the modelling systems and to assess their interacting performance characteristics and source data through open standards-based

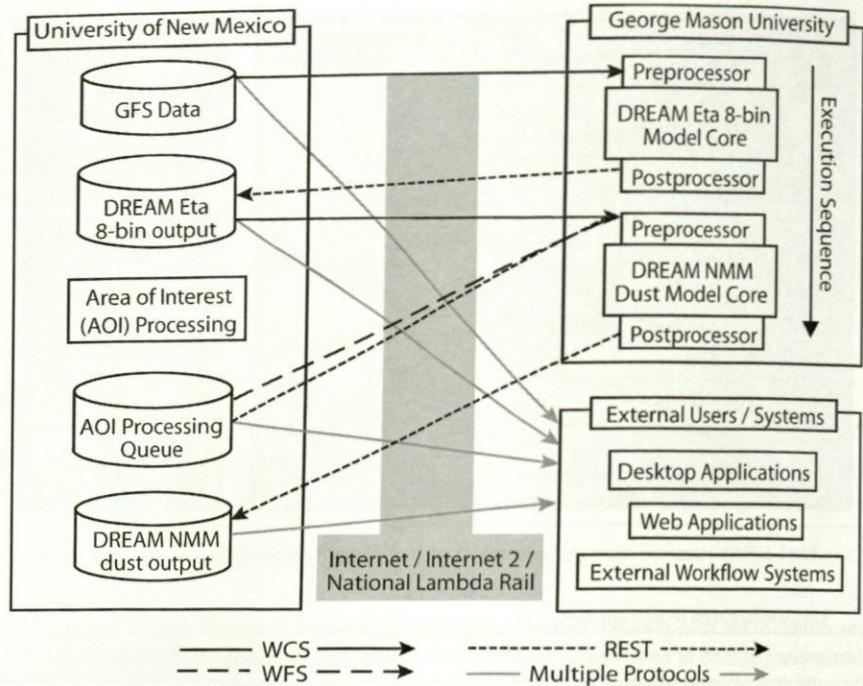


Figure 6. System integration workflow.

interoperable data services. The potential benefits of such a system are: 1) improved flexibility and an improved capability for integrating new data sources into the modelling system; 2) an ability to separate data management and publication physically from modelling; and 3), insert a degree of platform and file format independence between system components (i.e. only having to know the protocol for requesting data from a service, and not having to know a model's operating system or source data format for a particular data service).

The system integration approach is illustrated in Figure 6. The overall integration model is one in which all persistent data are stored and published via Open Standards at the Earth Data Analysis Center (EDAC) at the University of New Mexico, and all modelling is performed on separate systems at George Mason University (GMU). Data exchange between EDAC and GMU was accomplished through two primary OGC web service standards, and OGC web coverage services (Table 7, C) and the representational state transfer (REST) web service model, which in turn is based on the Internet standard hypertext transfer protocol (HTTP) (Fielding 1999; Fielding 2000; Whiteside & Evans 2006). These standards provide the connectivity between separate data management and processing components in the system.

Individual components performing the integrated modelling steps include the data storage and management systems represented by the cylindrical objects on the left in Figure 6. They provide persistent storage of the global forecast system (GFS) data (Table 7, D); low-resolution DREAM/eta model outputs; calculated areas of interest for which high-resolution model runs should be executed; and high-resolution DREAM/nmm model outputs. All three sets of model products (GFS, DREAM/eta, and DREAM/nmm) are published both within the system and for external users via WCS, while the calculated areas of interest are published via REST. The modelling and data processing components are represented by the rectangular boxes on the right in Figure 8. They perform model runs and analytical functions. For the modelling components hosted at GMU, model

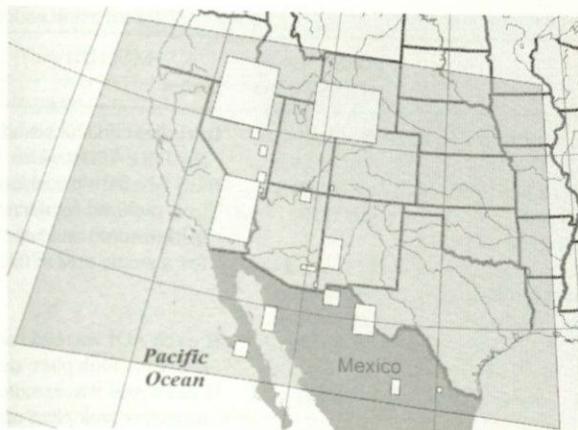


Figure 7. Feasibility study model domain and areas of interest bounding boxes for high-resolution model runs.

pre-processors were modified to enable dynamic acquisition of initialization and boundary conditions from remote data servers via WCS and REST, and post-processors were modified to push model outputs to the remote data storage systems at EDAC. In both cases modification of the pre- and post-processors allow streamlined integration of the existing DREAM/eta and DREAM/nmm modelling cores within a broader services-oriented architecture.

The DREAM/eta and DREAM/nmm modelling cores remained largely unchanged in this project. The primary changes consisted of parameterising the models to allow modifications of the model domain without having to recompile the model. Figure 7 shows the full model domain. A complete set of model runs consists of full-domain low-resolution (50 km) runs of DREAM/eta, from which sub-domain regions of elevated dust concentrations are derived by the areas of interest processing component run on the servers at EDAC. The set of areas of interest for a specific low-resolution model run are then retrieved by the DREAM/nmm server for execution of individual areas of interest at higher spatial resolution. This targeted execution of a high-resolution model only for areas for which a more detailed analysis is required reduces the overall computational requirements significantly for generating a comprehensive picture of a large model domain.

### 3.4.1 Data sources

A variety of sensor data, model, and geospatial data were employed. MODIS land cover, from which potential dust production sites are identified, are obtained from both MOD12Q1 and MCD12Q1 products using the *Barren/Sparsely Vegetated* class. Digital elevation data were obtained from the global digital elevation and topography data set (GTOPO30); and aerodynamic surface roughness data were derived from a table look-up based on land cover developed by Stennis Space Center (SSC) (Sanchez 2007; Morain & Sprigg 2008). GFS data at one degree resolution are available every six hours. These data constitute the key, externally provided, initial and boundary condition parameters used by both modelling cores.

In previous NASA applications projects, data and products have been stored locally with the model and used as fixed data sources that are linked statically to the model through direct reconfiguration of the model code. This static approach was overcome in this project by using modified model pre-processors that use WCS for acquiring essential GFS data for a specific model run. This approach provides two reductions in data management and transfer costs. First, the current collection of over 1.3 terabytes of GFS data stored at EDAC does not need to be replicated at GMU to support local modelling. Second, WCS allows transfer of only the subset of needed meteorological parameters from each GFS product instead of transferring complete data files. Specifically,

Table 8. Feasibility test data transfer times.

Activity	Time (HH:MM:SS)	Notes
Retrieval of GFS parameters from EDAC	00:01:32	Initialization and boundary conditions for DREAM/eta from WCS
Delivery of DREAM/eta results to EDAC	00:04:37	REST-based submission
Retrieval of DREAM/eta and AOI data from EDAC servers	00:03:50	Time required for the retrieval of the required initialization and boundary condition data for a single area of interest DREAM/nmm run
Delivery of DREAM/nmm results to EDAC servers	00:01:30	
Minimum transfer time	00:11:59	If each AOI was executed in parallel and transfers took place simultaneously
Maximum transfer time	01:42:09	If each AOI was executed sequentially and transfers took place one after another

only five meteorological parameters are required to run DREAM/eta: geopotential height, relative humidity, temperature, u winds and v winds (Benedict *et al.*, 2010). These five parameters, when retrieved for a 24-hour model run, represent a file transfer of about 1.2 megabytes, compared to a transfer size of 360 megabytes to transfer the complete set of source GRIB1 files.

#### 3.4.2 Project outcomes

The feasibility test was developed and executed for a specific set of dust events that took place on 1 July 2007 as an end-to-end test of the performance characteristics. The test included WCS services based upon the THREDDS data server and REST services hosted at EDAC for model submission to EDAC's servers and retrieval of areas of interest; and DREAM/eta and DREAM/nmm modelling systems hosted at GMU. The test measured data transfer times between EDAC and GMU over Internet-2, and model execution times for both DREAM/eta and DREAM/nmm on the servers maintained by GMU. The test was performed over the DREAM/eta domain shown in Figure 9. Eighteen areas of interest were identified in the test, each of which was used to execute a high-resolution (3 km) DREAM/nmm model run. Data transfer times are presented in Table 8 with the minimum and maximum data transfer times summarized at the bottom of the table. One can see that individual transfer times are relatively short over the Internet 2 connection between EDAC and GMU; and, given a scenario where data for individual areas of interest execution on separate systems takes place, the minimum transfer time for parallel transfer and execution is just under twelve minutes.

Model execution times are summarized in Table 9 for both the full-domain DREAM/eta and eighteen individual areas of interest for the DREAM/nmm model. As with the data transfer times discussed above, while the total execution time for all of the individual models could be quite long if executed sequentially over twelve hours, if the DREAM/nmm model runs were executed simultaneously on separate computers, the total execution time for the DREAM/eta and all eighteen DREAM/nmm areas of interest model runs is only 1.75 hours. This total execution time is significantly shorter than the estimated seventy-two hours that it would take to execute a twenty-four hour DREAM/nmm model run over the entire model domain.

These results suggest two conclusions: 1) that the interoperable nested modelling approach is feasible as an alternative to monolithic model integration frameworks in which specialized data transfer protocols and connectivity models are employed; and 2) significant gains in accelerated access to multi-resolution model outputs can be obtained through parallel execution of independent high-resolution areas of interest models when compared to the time required to execute a full-domain, high-resolution model.

Table 9. Model execution times for DREAM/eta and DREAM/nmm models.

Model run type	Time (HH:MM)	Notes
DREAM/eta	0:20:00	Full 37 × 20 degree model domain at a ~50 km resolution
DREAM/nmm	Min: 00:27 Max: 01:25 Mean: 00:39 Median 00:32 Sum: 11:43 n = 18	Summary statistics for 18 NMM-dust runs (3 km × 3 km) for the AOIs identified following the 1 July 2011 DREAM/eta model run
Minimum total execution time	01:45	Combined DREAM/eta and DREAM/nmm execution time if all DREAM/nmm AOI model runs are executed simultaneously on separate computers
Maximum total execution time	12:03	Combined DREAM/eta and DREAM/nmm execution time if all DREAM/nmm AOI model runs are executed sequentially on a single server

Overall, the outcome of this analysis is that this approach has potential for accelerating delivery of actionable dust forecasts to public health officials, while also establishing open standards based services for performing automated QA/QC on products generated by the system. Furthermore, the simplified parallel execution model has significant potential for deploying scalable modelling systems into commodity computing environments (that is, cloud computing), providing scalable modelling capabilities that do not require investment in fixed computing infrastructure, but instead take advantage of the expanding availability of on-demand computational resources.

### 3.5 Health data from SEDAC

The Socioeconomic Data and Applications Center (SEDAC) is operated by the Center for International Earth Science Information Network (CIESIN) at Columbia University. It provides data and services related to human interactions with the environment, and in particular demographic and socioeconomic data that can be integrated with EO data and imagery. A number of SEDAC data and information resources have been used extensively in public health research and surveillance, especially those related to the spatial distribution of population and associated demographic characteristics. SEDAC also provides a number of interactive tools and resources for visualization and analysis of interdisciplinary data useful to a wide range of users concerned with public health issues. The sub-sections below describe a variety of data sets and services for generating health-related information on-the-fly from numerous URLs. Table 10 is a list of these services and their access URLs.

#### 3.5.1 Basic services

SEDAC's interactive map client provides visualization of various geospatial data layers available from SEDAC (Table 10, A). It is organized by data set collections, selected interdisciplinary topics and major world regions. It retrieves data from SEDAC map servers through open geospatial standards, and overlays them for visual analysis. Users can perform basic map operations by turning base layers and overlays on or off, and by zooming into areas of interest. Key layers of interest to the public health community include data on population distribution, land cover and land use, infant mortality, child malnutrition, air and water quality, and natural hazards. Planned developments include the ability to subset directly and download data to retrieve layers from external servers, and to integrate spatial queries and other analytic functions. SEDAC's data layers are also available to other map clients that can access data through OGC standards.

The population estimation service (PES) (Table 10, B) estimates population totals and related statistics within a user-defined region. It enables one to obtain estimates of the number of people

Table 10. URLs referenced in section 3.5.

Products/services	URL (accessed 18th January 2012)
A. Interactive map client	<a href="http://sedac.ciesin.columbia.edu/maps/client">http://sedac.ciesin.columbia.edu/maps/client</a>
B. Population estimation service	<a href="http://sedac.ciesin.columbia.edu/gpw/wps.jsp">http://sedac.ciesin.columbia.edu/gpw/wps.jsp</a>
C. SEDAC TerraViva! Viewer	<a href="http://sedac.ciesin.columbia.edu/terraVivaUserWeb">http://sedac.ciesin.columbia.edu/terraVivaUserWeb</a>
D. Map gallery	<a href="http://sedac.ciesin.columtia.edu/maps/gallery/browse">http://sedac.ciesin.columtia.edu/maps/gallery/browse</a>
E. Photostream	<a href="http://www.flickr.com/photos/54545503@N04/">http://www.flickr.com/photos/54545503@N04/</a>
F. Global rural urban mapping project	<a href="http://sedac.ciesin.columbia.edu/gpw">http://sedac.ciesin.columbia.edu/gpw</a>
G. Global poverty mapping project	<a href="http://sedac.ciesin.columbia.edu/povmap/">http://sedac.ciesin.columbia.edu/povmap/</a>
H. United States census grid data	<a href="http://sedac.ciesin.columbia.edu/usgrid/">http://sedac.ciesin.columbia.edu/usgrid/</a>
I. Population, landscape, and climate estimates	<a href="http://sedac.ciesin.columbia.edu/place">http://sedac.ciesin.columbia.edu/place</a>
J. SeaWiFS chlorophyll concentrations	<a href="http://sedac.ciesin.columbia.edu/es/seawifs.html">http://sedac.ciesin.columbia.edu/es/seawifs.html</a>
K. Environmental performance index	<a href="http://sedac.ciesin.columbia.edu/es/epi/">http://sedac.ciesin.columbia.edu/es/epi/</a>
L. Anthropogenic biomes	<a href="http://sedac.ciesin.columbia.edu/es/anthropogenicbiomes.html">http://sedac.ciesin.columbia.edu/es/anthropogenicbiomes.html</a>
M. Confidentiality in geospatial data applications	<a href="http://sedac.ciesin.columbia.edu/confidentiality">http://sedac.ciesin.columbia.edu/confidentiality</a>

residing in specific areas quickly without having to download and analyse large amounts of spatial data. The service accepts polygons that define areas of interest and returns population totals, land area, quality measures, and basic parametric statistics for the requested polygons based on version-3 of SEDAC's gridded population of the world (GPWv3) data. The service can be used to estimate populations at risk of exposure to a public health threat or natural hazard event or episode.

The PES service is accessible through three standard protocols used by many online map tools and clients: the OGC WPS standard, a REST interface, and a SOAP interface. Standards-based clients such as uDig are able to submit requests using the OGC WPS. Users of ArcGIS software from Esri can submit requests through SOAP. The REST interface is intended for use with lightweight java script clients. The parametric statistics returned for each polygon include the count (number of grid cells used in the analysis), minimum population count, maximum population count, range of population counts, mean population counts, and standard deviation of population counts. Two measures of data quality are included in the service results. The first measure reflects the precision of the input data and the second indicates when the requested polygons were too small in area compared with the underlying input data to produce reliable population statistics. SEDAC provides a simple client interface and related interfaces that allow users to submit a single polygon for analysis. More complex queries can be submitted through a GIS software package that supports spatial queries through one of the three supported protocols.

*Terra Viva! SEDACViewer* (Table 10, C) is a Windows-based standalone software application that provides a powerful data-viewing engine and tools to visualize and integrate hundreds of different socioeconomic and environmental variables and layers, including a range of remote sensing data sets. The 2011 version has been updated with several new SEDAC data sets and includes fifty-one ready-made maps, ten GeoData indicator collections with hundreds of variables, and other features such as scatter plots, tabular data display, map image production, and Web-based download of additional data layers.

The SEDAC TerraViva! Viewer is unique for many reasons. It is easy to use and does not require an Internet connection. Its global data viewing engine enables users to open multiple windows to view multiple layers side-by-side, rather than as overlays. These windows may be linked to show different layers for the same area or unlinked to show the same layer at different spatial scales or using different geographic projections. The gazetteer helps to pinpoint cities, states, provinces, countries, water bodies, and other locations. The Viewer offers a library of maps by theme: population distribution, land cover, physical geography, and more. Users may also create customized maps and charts based on their data and areas of interest.

SEDAC's map gallery provides ready-made maps based on key SEDAC data resources addressing a wide range of topics relevant to the public health community (Table 10, D). Maps have been created for individual countries, selected regions, and for the globe as appropriate for each data set or data collection. Most maps are available both in jpeg and PDF formats, and are made available under a creative commons 3.0 attribution license as long as acknowledgement is given. The maps are also available directly via Flickr through the photo-stream service (Table 10, E). Maps relevant to public health applications include those developed for the GPW and global rural-urban mapping project (GRUMP) collections, the indicators of coastal water quality data set, the US Census Grids collection, the poverty mapping collection, and the environmental performance index data sets.

### 3.5.2 Data and information for health applications

GPWv3 and GRUMP collections represent complementary efforts to improve understanding of human spatial distributions across the globe (Table 10, F). GPWv3 provides estimates of population totals and densities on a latitude-longitude grid for 1990, 1995, and 2000, based on data available in 2000 from national population censuses and related sources. GPWv3 has a nominal resolution of 2.5 arc minutes. Grid cells are about 21 km<sup>2</sup> at the equator or 15 km<sup>2</sup> at 45°N or 45°S, and use the best available subnational boundary and population data for more than 125,000 subnational administrative units to determine population distribution within countries at this grid scale. GRUMP attempts to better delineate urban versus rural population distribution by incorporating data on urban extent from satellite observations of night lights and other supplementary data sources. GRUMP provides population counts and densities for 1990, 1995, and 2000 on a thirty arc-second grid (about one square kilometre). GRUMP also includes data on human settlements and urban extent.

Several complementary data sets have also been developed. GPW future estimates (GPWfe) projects population distribution grids for 2005, 2010, and 2015 using simple population growth rates. The purpose of these projections is to show a scenario of future spatial distribution for population at a subnational resolution, assuming a continuation of recent demographic patterns (Balk *et al.*, 2005). The low elevation coastal zone urban-rural estimates (LE CZ-URE) combine GRUMP data with SRTM elevation data to produce population estimates on a thirty arc-second grid for populations resident on coastal lands less than ten meters in elevation above mean sea level (McGranahan *et al.*, 2007).

These population data sets have been used widely in public health research and surveillance, as they provide information on the population potentially exposed to infectious disease, environmental stresses and disease. Combining the GPW with satellite-derived ground-level values of fine particulate matter (PM<sub>2.5</sub>), scientists have been able to estimate global long-term exposures to this air pollutant (van Donkelaar *et al.*, 2010). GPW data also have been used to determine the effect of population density and population growth on global temporal and spatial patterns of emerging infectious diseases (Jones *et al.*, 2008). GPW and GRUMP data were used in several studies on malaria incidence and the effectiveness of prevention programs (Teklehaimanot *et al.*, 2007; Hay *et al.*, 2010; Tatem *et al.*, 2010). Gridded population counts and population density from GRUMP enabled estimation of the number of children under age five at risk of malaria in a study on insecticide-treated net coverage in Africa (Noor *et al.*, 2009). In modelling the interaction of short-distance commuting and long-distance airline patterns and their impacts on global epidemics, both GPW and GRUMP were used to divide the world into a grid with cells that could be considered at different resolution levels (Balcan *et al.*, 2009). These are only a few of many health-related studies that have utilized these gridded population data sets as an integral part of their analyses.

#### 3.5.2.1 Poverty mapping

The Global Poverty Mapping project seeks to enhance current understanding of the global distribution of poverty and the geographic and biophysical conditions of where the poor live (Table 10, G). The project aims to assist policy makers, development agencies, and the poor themselves in designing socioeconomic and health interventions to reduce poverty and achieve the Millennium Development Goals. Subnational, spatially explicit, poverty data sets are provided for selected



Figure 8. Subnational infant mortality rate data serves as one proxy in mapping global poverty. (see colour plate 43)

proxy measures of poverty at global and national scales. The global data are of varying resolution, but primarily coarse; the national data sets are at considerably higher-resolution. Data catalogues for Poverty Mapping describe the variables available in each data set, and the underlying spatial, survey, and census data sets used to construct the integrated collection. Metadata records also provide details on source data and methods.

At a global scale, poverty is usually represented by national indicators such as gross domestic product or population living on less than one US dollar/day. These indicators are not available at a subnational level for most countries. The poverty mapping site provides global estimates of poverty based on subnational infant mortality rates and child malnutrition data, recognizing that both are proxies for poverty and welfare rather than direct measures (Figure 8). Data were drawn from demographic and health surveys, multiple indicator cluster surveys, national human development reports, and other sources. These data were first linked to boundary data for their reporting regions and then translated to a common grid (Storeygard *et al.*, 2008).

Numerous methods have been used to construct estimates of poverty that are finely resolved. These methodologies utilize both indirect and direct estimation techniques. Indirect estimation includes small area estimates of poverty and inequality, poverty and food security country case studies, and combining surveys with remote sensing data. Direct estimation is based on the unsatisfied basic needs data set. Poverty and welfare are generally measured through proxy consumption variables like estimates of expenditures and consumer goods, or basic needs like sanitation, water, housing, and education. Data sets representing both approaches are available in this collection.

### 3.5.2.2 Populations near US superfund sites

In response to a request by the National Institute of Environmental Health Sciences (NIEHS), CIESIN conducted an assessment of populations living in proximity to superfund National Priorities List (NPL) sites for 2000 (Golden *et al.*, 2008) (Figure 9). This assessment improved on earlier studies in several ways: 1) SEDAC's US census grids data set was used to determine population totals and demographic breakdowns in proximity to NPL sites (Table 10, H); 2) a methodology was applied to eliminate double-counting of populations in proximity of more than one NPL site; 3) the study used a majority rule to determine when population data should be included in the summation; 4) the total population in proximity to two or more Superfund sites was estimated; and 5) these analyses were conducted twice: once for populations living within a one-mile buffer, and again for those within four miles of the site.

SEDAC provides access to the assessment report and to the three key spatial data sets in shapefile format used for the analyses: 1) Agency for Toxic Substances and Disease Registry (ATSDR) hazardous waste site polygon data, 1996; 2) US/EPA's NPL point data with CIESIN modifications, 2008; and 3) ATSDR hazardous waste site polygon data with CIESIN modifications, 1996. In the future, an online interactive map service with information on vulnerable populations in proximity to Superfund sites across the US will also be available to the public.



Figure 9. US population and NPL superfund sites with four mile buffers. (see colour plate 44)

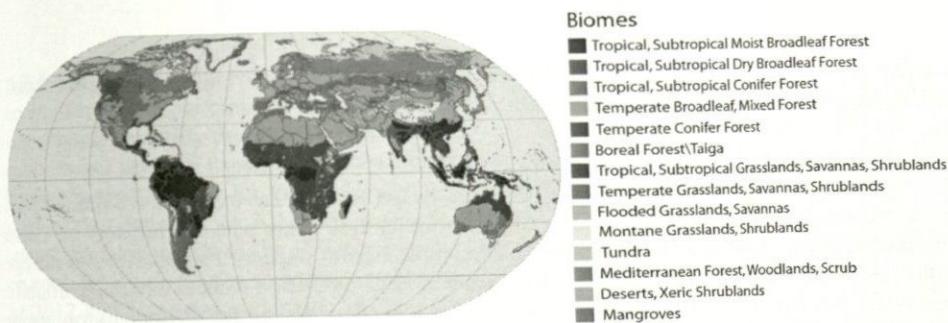


Figure 10. PLACE II provides population and land area distribution for classes of several different demographic, physical, biological, and climatic variables, including Biomes. (see colour plate 45)

### 3.5.2.3 Population, landscape and climate estimates

The population, landscape and climate estimates (PLACE II) data set contains country-level measures of the spatial characteristics of 228 nations to researchers for whom national aggregates are more useful than GIS data (Table 10, I). PLACE II provides the numbers and percentages of people and the land area (square kilometres and percentages) represented within each class of a number of demographic, physical, biological, and climatic variables for each country around the world, for the years 1990 and 2000. These variables include biomes, climate zones, coastal proximity zones, elevation zones, and population density zones as shown in Figure 10. The full data array of nearly 300 variables, tabulated by country, is available for download in spread-sheet format, together with supporting documentation. The PLACE II map collection displays examples of input variables and country dynamics via more than forty maps at global, continental, and national scales.

The methodology used to develop PLACE II has been refined since the first version of PLACE. PLACE II employs population estimates from version 3 of the gridded population of the world data set (GPWv3) for both 1990 and 2000, in order to assess trends and population shifts through time. Thematic classes and land area estimates by country were held constant, to provide a uniform spatial geography across the decadal span. PLACE II also improves upon the spatial processing workflow to better account for updated coastlines and country boundaries used in GPWv3. Codebooks describing each variable used are provided along with a methodology paper (SEDAC 2007). PLACE II enables users to address such questions such as: 1) Which countries have the highest percentages of their population living within 100 or 200 km of a marine coast?; 2) What proportion of the

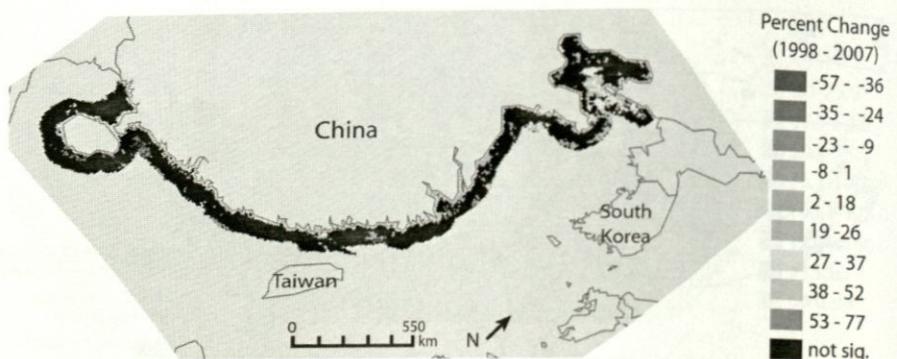


Figure 11. Percentages of change in chlorophyll concentrations as an indicator of algal biomass along the China coast, 1998–2007. (see colour plate 46)

Kenyan population lives at an elevation lower than where malaria-carrying mosquitoes are found?; and 3), In which climate types do most of the US population live?

#### 3.5.2.4 Indicators of coastal water quality

The flow of nutrients into coastal waters from land-based sources has increased around the world in recent decades. The resulting change in water quality has many potential impacts on coastal and marine ecosystems that in turn could lead to public health impacts. Phosphorus and nitrogen contribute to enhanced algae growth, and subsequent decomposition reduces oxygen availability to benthic life forms. Changes to nutrient loadings can also change the phytoplankton species composition and diversity. In extreme cases, eutrophication can lead to hypoxia, oxygen-depleted dead zones, and harmful algal blooms (see Chapter 9).

Measuring chlorophyll concentrations as an indicator of algae biomass may be a way to assess coastal water quality and changes over time. Figure 11 shows chlorophyll- $\alpha$  concentrations derived from SeaWiFS (Table 10, J). The data have been used to analyse trends over a ten year period (1998–2007), helping to identify near-coastal areas with improving, declining, and stable chlorophyll concentrations that can provide guidance for environmental management decisions. In Figure 11, black indicates no significant change; orange 53–77 per cent; light orange 38–52 per cent; yellow, 27–37 per cent; light green, 19–26 per cent; and green, 2–18 per cent. Areas of declining concentrations are shown in light blue, –1 to –8 per cent; medium blue, –9 to –23 per cent; blue, –24 to –35 per cent; and dark blue, –36 to –57 per cent. The website also provides a description of the global data set and how it was derived. Raster data are available for download in Esri GRID and GeoTiff format, and the indicators of change are provided in tabular format. Ancillary data are provided in Esri GRID and shape-file formats.

Yearly average concentrations of chlorophyll- $\alpha$  in nanograms/metre<sup>3</sup> are based on annual composites of SeaWiFS satellite data provided by GSFC and GeoEye at a resolution of nine km<sup>2</sup>. Level-3 products were derived from true-colour images generated from sub-sampled, calibrated, Rayleigh-corrected level-2 data, which were derived from raw radiance counts by applying sensor calibration, atmospheric corrections, and bio-optical algorithms. HDF files were converted to Esri GRID format and coastline buffers were clipped by country and exclusive economic zone. For defining coastal zones, the first ten kilometres of coastal waters were excluded because of the potential for bottom-reflectance or suspended sediments affecting the satellite measurements in close proximity to the coast. The extent of coastal zones was limited to 100 km offshore as an arbitrary limit beyond which impacts from land-based sources on oceanic eutrophication are unlikely.



Figure 12. 2012 Environmental performance index scores. Dark green: strongest index; light green, strong; ivory, modest; orange, weaker; red, weakest; and grey, no index (Courtesy CIESIN & YCELP). (see colour plate 47)

### 3.5.2.5 Environmental performance index 2010

The environmental performance index (EPI) ranks 132 countries on their environmental performance based on twenty-two indicators in ten policy categories including environmental health, air quality, water resource management, biodiversity and habitat, forestry, fisheries, agriculture, and climate change (Table 10, K). Figure 12 shows EPI rankings for 2012. The website provides access to materials relating to the 2012 EPI, including a summary for policy makers, a detailed report, country profiles, access to the EPI component data and metadata, and an interactive visualization tool. The website also includes an archive of data and reports for the 2006 and 2008 EPIS. The 2012 EPI was developed by the Yale Center for Environmental Law and Policy (YCELP) and CIESIN, in collaboration with the World Economic Forum and the Joint Research Centre of the European Commission. The 2012 EPI centres on two broad environmental protection objectives: reducing environmental stresses on human health, and promoting ecosystem vitality and sound natural resource management. It utilizes a proximity-to-target methodology focused on a core set of environmental outcomes linked to policy goals (YCELP *et al.*, 2010).

Each nation is benchmarked against others that are similarly situated with groupings, based on geographic regions, level of development, trading blocs and demographic characteristics. Many of the EPI indicators are directly relevant to public health. These include: 1) *child mortality*: the probability of dying between one's first and fifth birthdays per 1000 children at age one; 2) *access to sanitation*: percentage of a country's population that has access to an improved source of sanitation; 3) *access to drinking water*: percentage of a country's population that has access to an improved source of drinking water, that is: piped water into dwelling, plot or yard; public tap/standpipe, tubewell/borehole, protected dug well, protected spring, and rainwater collection; 4) *indoor air pollution*: percentage of population using solid fuels in households; 5) *particulate matter*: population weighted exposure to PM<sub>2.5</sub> in micro grams per cubic meter; 6) *SO<sub>2</sub> emissions per capita*: amount of SO<sub>2</sub> in kilograms per person; 7) *SO<sub>2</sub> emissions per \$ GDP*: amount of SO<sub>2</sub> per 2005 constant dollar USD; 8) *change in water quantity*: area weighted per cent reduction of mean annual river flow from "natural" state owing to water withdrawals and reservoirs; 9) *pesticide regulation*: based on a twenty-two point legislative scale on status of pesticide use and toxic chemical controls; and, 10) *renewable electricity*: renewable electricity production as a percentage of total electricity production.

### 3.5.2.6 Anthropogenic Biomes

Anthropogenic biomes (aka anthromes or human biomes) describe the terrestrial biosphere in its contemporary, human-altered form using global ecosystem units defined by patterns of sustained

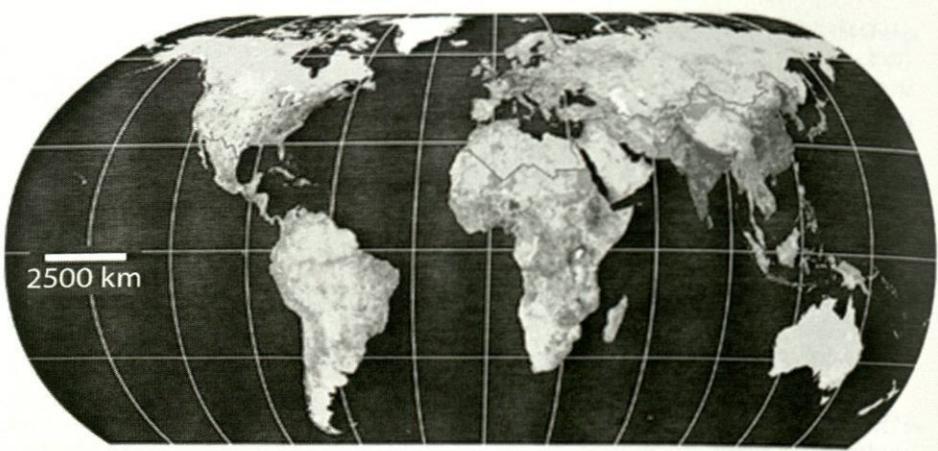


Figure 13. Anthromes based on population density, land use and vegetation cover. Blue, purple and red colours define heavily cultivated regions in Asia and Europe; greens and yellows are semi-natural and croplands; pale to darker orange colours are defined as rangelands; and pale green to grey areas are "wild" lands (Courtesy Brandon, 2010). (see colour plate 48)

direct human interaction (Figure 13). It delineates twenty-one anthropogenic biomes based on population density, land use and vegetative cover. Anthropogenic biomes are grouped into six major categories: dense settlements, villages, croplands, rangeland, forested and wild lands (Ellis & Ramankutty 2008). A full description of these biomes is provided in their website (Table 10, L). Users may download the data as one global grid or in six separate grids for the populated continents. The data are available in GeoTif and GRID formats. Data may be useful to public health surveillance and research by providing contextual information important for understanding human-environment-disease interactions.

#### 3.5.2.7 Confidentiality issues in geospatial data applications

The synthesis of geospatial data with socioeconomic and medical data could lead to many benefits for society, especially in terms of improving public health. Geographic information systems are powerful integrating technologies capable of bringing together information from a variety of sources, including remotely sensed data from instruments aboard aircraft and orbiting satellites and precise spatial coordinates from GPS instruments. The analytical potential of linking spatially explicit data with health surveys and other demographic and behavioural data is great. However, location-specific data at the household or even neighbourhood level may provide sufficient information so that the identity of study subjects can be determined either directly or indirectly.

In the field of public health, good science and successful policies depend on developing effective strategies to balance the rights of individuals with the needs of the community. In order to understand, diagnose, monitor, treat, and prevent diseases and injuries that harm sectors of the population, it is necessary to enlist the cooperation of those at risk. Only with detailed information on individual exposures, behaviour, and socio-demographic and health conditions can researchers begin to understand the aetiology of illnesses. Survey and study data combined with extensive georeferenced data from multiple projects across diverse disciplines can further reveal the dynamic interactions of environment, infrastructure, populations, and disease. However, such linkages also have the potential of disclosing the identities of individual study participants. Some approaches that have been implemented to protect confidentiality while still providing data access include: aggregation, masking, and suppression; research contracts with confidentiality clauses and disclosure penalties; safe houses for restricted data access by approved users; and, protected online data-sharing co-laboratories (Golden *et al.*, 2005).

Table 11. URLs referenced in section 3.6.

Product/service	URL (all accessed 18th January 2012)
A. Distributed Active Archive Centers	<a href="http://esdis.eosdis.nasa.gov/dataaccess/datacenters.html">http://esdis.eosdis.nasa.gov/dataaccess/datacenters.html</a>
B. DAACs and data archives	<a href="http://nasadaacs.eos.nasa.gov/bout.html">http://nasadaacs.eos.nasa.gov/bout.html</a>
C. Mirador	<a href="http://mirador.gsfc.nasa.gov/">http://mirador.gsfc.nasa.gov/</a>
D. OPeNDAP	<a href="http://disc.sci.gsfc.nasa.gov/services/opendap/">http://disc.sci.gsfc.nasa.gov/services/opendap/</a>
E. IDV	<a href="http://www.unidata.ucar.edu/software/idv/">http://www.unidata.ucar.edu/software/idv/</a>
F. McIDAS-V	<a href="http://www.ssec.wisc.edu/mcidas/software/v/">http://www.ssec.wisc.edu/mcidas/software/v/</a>
G. Panoply	<a href="http://www.giss.nasa.gov/tools/panoply/">http://www.giss.nasa.gov/tools/panoply/</a>
H. Ferret	<a href="http://ferret.wrc.noaa.gov/Ferret/">http://ferret.wrc.noaa.gov/Ferret/</a>
I. GrADS	<a href="http://www.iges.org/grads/">http://www.iges.org/grads/</a>
J. GrADS data server	<a href="http://disc.sci.gsfc.nasa.gov/services/grads-gds">http://disc.sci.gsfc.nasa.gov/services/grads-gds</a>
K. Simple subset wizard	<a href="http://disc.gsfc.nasa.gov/SSW/">http://disc.gsfc.nasa.gov/SSW/</a>
L. Giovanni	<a href="http://giovanni.gsfc.nasa.gov">http://giovanni.gsfc.nasa.gov</a>
M. Giovanni documentation	<a href="http://disc.sci.gsfc.nasa.gov/giovanni/additional/users-manual">http://disc.sci.gsfc.nasa.gov/giovanni/additional/users-manual</a>

The SEDAC website on confidentiality issues in geospatial data Applications (Table 10, M) promotes awareness of confidentiality and privacy issues related to the integration of geo-referenced data from the natural, social, and public health sciences. It provides access to resources that analyse these issues and that offer concrete tools, techniques, and policies for safeguarding confidential information. It is intended for natural, social, and public health scientists, educators, data managers, and decision makers who use and disseminate geospatial data, especially remote sensing data.

### 3.6 Health data and information services at GES DISC

As noted in Section 2.0, public health and environmental scientists have become increasingly resourceful in using satellite sensor data sets to facilitate their research. Although a core set of environmental parameters has emerged, many other parameters are utilized as a result of particular research interest and/or the need of specific projects. In addition there is an increasing number of choices about where to access and utilize specific geophysical measurements. As a result, project personnel often need to become students of the data products they employ to understand the assumptions embedded in data creation, data quality, and data validation. Maximizing the combination of the data user's knowledge of the product, and data services that more easily provide that knowledge will yield the most refined research results.

NASA, through its Earth Observing System Data and Information System (EOSDIS) program, has developed, and continues to evolve, information technologies that facilitate maximum return on its Earth science data investments. The GES DISC is one of twelve Earth science data centres, aka Distributed Active Archive Centres (DAACs). Each specializes in one or more specific Earth science disciplines. URLs associated with this section are listed in Table 11

GES DISC specializes in archiving, distributing, stewarding, and providing user data access and analysis services for remote sensing data associated with global atmospheric composition, atmospheric dynamics, hydrology, precipitation, and global modelling data. In addition, GES DISC has developed several tools and services that promote easier use and usability of Earth science data and information. As data and information management needs of science researchers have become more sophisticated, services have taken advantage of maturing information technologies to develop and implement tools and services that help researchers extract the information they seek from data.

Although this discussion focuses on the products and services residing at GES DISC, the same, or similar data and services, are available from other DAACs and data archives (Table 11, A & B). For example, the source of MODIS data products shown in Table 12 to support public health

Table 12. Types of environmental parameters and sensors that provide data.

Environmental parameter	Sensor measurements
Precipitation	TRMM; Aqua/AIRS (daily); Aqua/AIRS (monthly); GLDAS (monthly); MERRA (monthly 2D & 3D, chem., hourly 2D & 3D-3 hourly)
Relative humidity	MODIS (daily); Aqua/AIRS (daily); Aqua/AIRS (monthly); Aura MLS
Water runoff	GLDAS (monthly); MERRA (monthly 2D and 3D); MERRA (monthly chem.); MERRA (hourly 2D, 3D, and 3-hourly)
Vegetation indices	MERRA (monthly 2D and 3D); MERRA (hourly 2D, 3D, and 3-hourly)
Soil moisture	GLDAS (monthly); NEESPI (daily, monthly); MERRA (monthly 2D and 3D); MERRA (hourly 2D; 3D, and 3-hourly)
Surface temperature	Aqua/AIRS (daily, monthly); Aura MLS; MERRA (monthly 2D and 3D; hourly 2D, 3D, and 3-hourly); GLDAS (monthly)
Air quality	Aura TES
Aerosols	AEROSOL (daily, monthly); Aura (OMI L3, L2G); MODIS (daily, Monthly); MISR; TOMS; MERRA (monthly 2D, 3D; monthly chem.); MERRA (hourly 2D, 3D, and 3-hourly)
Wind	MERRA (monthly 2D, 3D; monthly chem.); MERRA (hourly 2D, 3D and 3-hourly); GLDAS (monthly)

research can be found at several locations. The remainder of this section describes the data and information services offered by GES DISC that are relevant to public health research.

### 3.6.1 *Health related Earth science data sets at GES DISC*

Various remote sensing measurements identified by public health research projects reside in many NASA and non-NASA archives. In Table 12 previously used and potential data sets relevant to public health research available from GES DISC are listed by measurement and remote sensing instrument. In addition, relevant modelled data of global geophysical parameters are included in the modern era retrospective-analysis for research and applications (MERRA), and other data sets generated by GSFC's global modelling and assimilation office (GMAO), the North American land data assimilation system (NLDAS), and the global land data assimilation system (GLDAS). Data products generated by GSFC's Hydrological Sciences Branch are also listed. Modelled data provide additional parameters useful to environmental health studies.

Whereas, several GES DISC products listed in Table 13 have proven to be useful in relating environmental factors to public health conditions, others have potential use to complement or verify EO data and ground based measurements currently being captured, or may be useful depending on the temporal or spatial resolutions required of the research.

### 3.6.2 *Data search, access, exploration, and discovery services*

Acquiring NASA data and information has been expedited greatly with advances in information technology and implementation of promising technologies to perform pre-research functions quickly, such as searching for specific data, exploring terabytes of data, and visualizing chosen data of interest. The use and usability of data initially generated for science research has expanded to applications research where data users are typically less familiar with EO data than with their usual data and data sources. Information services have facilitated inclusion of sensor data into applications R&D, and in particular public health research. The following GES DISC services have been developed and continue to evolve to enhance the access of data and glean information from the data on the researcher's behalf.

#### 3.6.2.1 Data search and access

Mirador is an Earth science data search tool developed at GES DISC (Table 11, C). It has a simplified, clean interface that combines a free text data set search with a relational database that

Table 13. Satellite data sets, parameters measured, spatial resolutions, and lengths of record.

Data set*	Measurement	Spatial resolution	Temporal resolution
TRMM	Precipitation	0.25 – 1 deg	1997-present
GPM	Precipitation		Starting in 2003
AIRS	Precipitation; surface air temperature	50 × 50 km ; 1 × 1 deg	2002-present
GLDAS	Precipitation; water runoff; soil moisture; surface air; temperature; wind	1/4 × 1/4 deg; 1 × 1 deg	1979-present
NLDAS	Precipitation; water runoff; vegetation index; soil moisture; surface air temperature; wind	1/8 × 1/8 deg	1979-present North America
MERRA	Precipitation; water runoff; vegetation index; soil moisture; surface air temperature; wind; aerosols	1.25 × 1.25 deg; 2/3 × 1/2 deg	1979-present
NEESPI-	Soil moisture	1 × 1 deg	2002-present
AMSR-E			
LPRM using	Soil moisture	25 × 25 km	2002-present
AMSR-E			
TOVS	Surface air temperature	1 × 1 deg	1984–1995
OMI	Aerosols; solar irradiance	24 × 13 km; 1 × 1 deg; 1/4 × 1/4 deg	2004-present
TOMS	Aerosols	1.25 × 1 deg	2004-present
DUST-DISC	Aerosols	4 × 4 km; 1/2 × 1/2 deg	1997-present
GOCART	Aerosols	2.5 × 2 deg	2000–2007
GSSTF2b	Wind	1 × 1 deg	1998–2008
SORCE	Solar irradiance		2003-present

\*For full names of data sets see Acronyms.

stores file information, as well as gazetteer locations for locations and geophysical events. Other features include quick response, spatial and variable sub-setting, data file hit estimator, gazetteer (geographic search by feature name capability), and an interactive shopping cart. The Mirador starting page has a keyword data search that requests three basic search entries: Keyword (e.g. instrument, parameter), date range for which data are desired, and geographic location for which data are desired. Location can be entered manually by pointing, clicking and dragging across a desired area on the map. All data sets residing at GES DISC can be located through Mirador and documentation for data and services is available.

Two other Mirador tabs are the *Projects* and *Science Areas* tabs. The former lists and describes all available data sets by sensor source or assimilation model. The latter provides data for Earth science categories and further breaks the categories into measurements that can be searched and accessed. The system continues to evolve as user communities request new services and data access.

### 3.6.2.2 OPeNDAP

The open source project for a network data access protocol (OPeNDAP, Table 11, D) provides alternative access to individual variables within data sets in a form usable by many tools, such as the integrated data viewer (IDV) (Table 11, E), McIDAS-V (Table 11, F), Panoply (Table 11, G), Ferret (Table 11, H) and GrADS (Table 11, I). This system provides software that makes local data accessible to remote locations regardless of local storage format. OPeNDAP software is available at no cost. GES DISC has installed OPeNDAP to provide users access to data that can be used easily in their local data analysis tools. Currently, health related data sets associated with the AIRS, TRMM, TOMS, and OMI instruments and the MERRA modelled data sets are accessible through OPeNDAP. Within the core of OPeNDAP is the GrADS data server GDS. This is a stable, secure data server that provides sub-setting and analysis services across the internet (Table 11, J). It is used

for data networking to make local data accessible to remote locations. Currently TRMM, MERRA, GLDAS, and NLDAS data products are available through GDS.

#### 3.6.2.3 Web mapping service

The OGC/WMS described in Section 3.4 is a service that provides map depictions for any spatial area over the network via a standard protocol, enabling users to build customized maps based on data coming from a variety of distributed sources. GES DISC offers WMS layers for AIRS surface temperature, TRMM rainfall, and OMI atmospheric parameters. WMS requests return an XML file that can be read by an OGC compliant client that includes many GIS tools. Subsequent WMS GetMap requests return image files to be rendered in the client. WMS-produced maps are generally rendered in a pictorial format such as PNG, GIF or JPEG. Occasionally they are rendered as vector-based graphical elements in scalable vector graphics (SVG) or web computer graphics metafile (WebCGM) formats (de la Beaujardiere 2006).

#### 3.6.2.4 Simple Subset Wizard

GES DISC has teamed with nine of other EOSDIS Data Centres to develop the simple subset wizard (SSW) (Table 11, K). This feature provides a unified user interface for submitting spatial subset requests for data across EOSDIS data centres. Geophysical parameters (keywords) related to public health can be searched by data centre.

### 3.6.3 *Information exploration and discovery*

#### 3.6.3.1 Giovanni

Giovanni is an online tool allowing application users to explore, visualize, inter-compare, and analyse EO data easily using only a Web browser. It has proven to be a valuable tool for locating and visualizing data, whether the user knows the specific data set of interest, or is seeking measurements with particular signatures. Starting with basic information: Location, date/time, parameter, and type of visualization, users can explore, view, and download specific data for further analysis. If the viewed data are not of interest, users can repeat the operation with different basic information. More advanced operations and visualizations are also available, such as multi-data comparisons, scatter plots, time series, animations, and more. Examples are shown in Figures 14–16. Currently, over forty project and discipline specific versions of Giovanni exist (Table 11, L).

Giovanni contains several hundred parameters from GES DISC data holdings, plus MISR AOD and data from the Tropospheric Emission Spectrometer (TES). The latter are distributed from the Langley Research Center (LaRC, Atmospheric Science Data Center (ASDC)). The MODIS aerosol and water vapour products are archived and distributed from LAADS at GSFC. As an example, precipitation data available through Giovanni are listed under measurements relevant to public health research. These include TRMM; Aqua/AIRS (daily and monthly); GLDAS (monthly); MERRA (hourly -2D, -3D; monthly-2D and -3D; monthly chemistry; and 3-hourly). Documentation is available (Table 11, M).

#### 3.6.3.2 Data Quality Screening Service

The data quality screening service (DQSS) filters bad-quality data points from a data file by using quality variables included within the file (Lynnes 2010). This service is extremely useful in applications research because it uncovers and utilizes higher quality data efficiently to give more accurate results. DQSS is available through the Mirador search tool. It generates output files that have the same structure as the input, albeit with a few extra variables, and is thus usable in any tool made to work with the original product. The service replaces data arrays with the corresponding filtered arrays. Data values deemed unusable are set to *fill values*. The output file also includes the quality mask used for filtering and the original (unscreened) data values in additional data objects. The DQSS steps shown in Figure 17 are accessible through Mirador for an AIRS data set. MODIS and MLS data sets are planned.

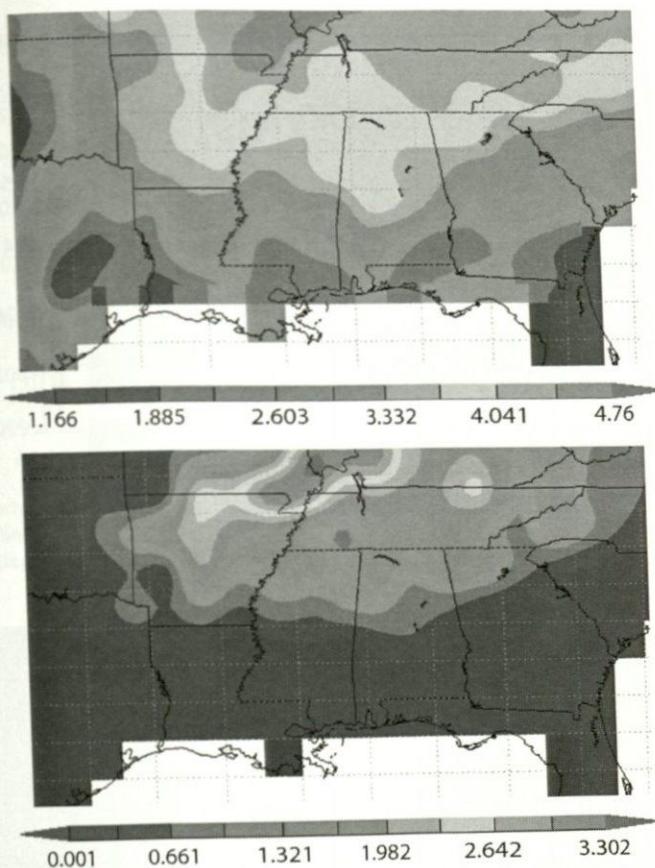


Figure 14. (Top) GLDAS monthly soil moisture and (Bottom) monthly surface runoff for the south eastern US between April and June 2011. Soil moisture ranges from <1.2 in purple and blue areas to more than 4.5 in the orange and red areas. Surface runoff values range from <0.6 to >2.6. (see colour plate 49)

#### 4 ACCESSING NASA EARTH SCIENCE INFORMATION

Section 2 provided the breadth of public health research that takes advantage of the availability of remote sensing data. Section 3 described the tools and services available to access remote sensing data for this research. Section 4 provides an in depth look at three projects describing in greater detail how remote sensing data are used to enhance public health applications. The first study by Doctors Tilburg and Zeeman addresses precipitation events to predict coastal water quality. The second study by Doctor Zaitchik addresses the value of remotely acquired data for developing a malaria early warning system in the Peruvian Amazon; and the third, by Doctor McClure links environmental data with a national public health cohort study to enhance public health decision making.

##### 4.1 Precipitation events for predicting coastal water quality

Rainfall and runoff are associated with the rise in water-borne diseases (Curriero *et al.*, 2001). Heavy rainfall events are more likely to lead to marked decline in microbiological quality of inland and marine recreational waters and drinking water supplies as a result of heavy runoff (Fayer *et al.*, 2004). Rain events result in movement of land-based pollutants to streams and rivers. These pollutants travel downstream with river flow to the coastal ocean and adjacent waters, which

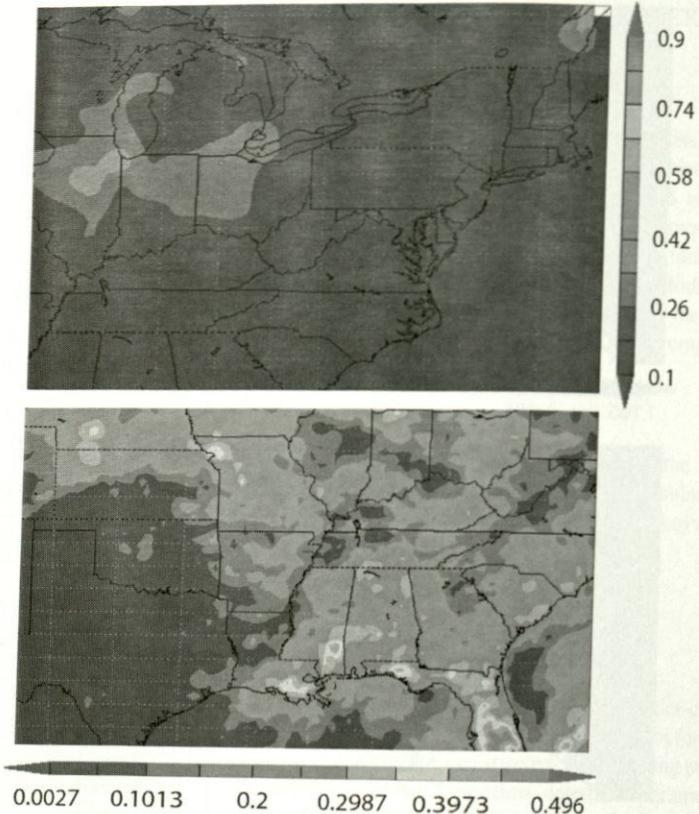


Figure 15. (Top) MODIS AOD values range from lowest in purple (<0.1) to highest in pale green (>0.3). (Bottom) Average TRMM precipitation rate at 0:00 UTC 22 June to 22 August 2011. Purple, blue, turquoise represent <0.002 to 0.2; green, yellow represent >0.2 <0.39; and orange, red >0.39 <0.49 mm/hr. (see colour plate 50)

results in reduced water quality events (RWQEs). In the north eastern US, climate change scenarios typically indicate an increase in overall precipitation, which would lead to larger river discharge (IPCC 2007). Increased discharge would then lead to greater pathogen loading in rivers and coastal waters from anthropogenic sources due to runoff from contaminated sites. Consequently, a method for predicting RWQEs has become crucial for effective water resource management. The central objective of this study is to determine if a relatively simple model using easily accessible remotely-sensed data can be used to predict RWQEs along the coast of Maine. Information from this study can be used to enhance every day decision making at the state and county level or aid in long term plans for water treatment plants in the region.

The most common method for testing water for RWQEs due to the presence of pathogens is to measure the concentration of *Escherichia coli* and total coliforms, which are bacteria that are normally present in the intestinal tract of humans and other animals and are used as indicators for recent faecal contamination. Unfortunately, the method is costly and time-consuming. Since most states do not currently have the ability or resources to quickly test for *E. coli* and Total Coliforms (typically 18–24 hrs.), proxies that can predict RWQEs are needed. A common method in the state of Maine is river discharge since discharge is highly correlated with upstream precipitation and the associated runoff within watersheds (A. Bouravkovsky pers. comm.). However, bulk measurements of discharge are not able to differentiate between run-off from high pollution areas that would tend

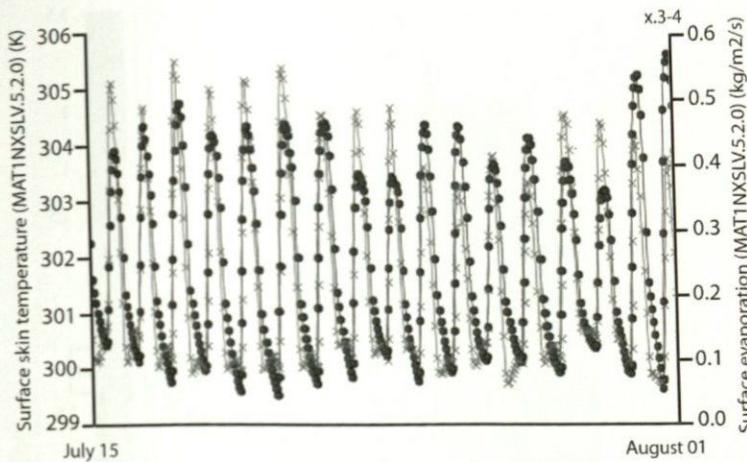


Figure 16. Hourly area-averaged temperature time series along the Texas/Louisiana coast 15July to 1August 2011. The blue line is surface skin temperature ranging from 299°K to 306°K, and the red line is surface evaporation ranging from 0.0 to 1.2 kg/m<sup>2</sup>. (see colour plate 51)

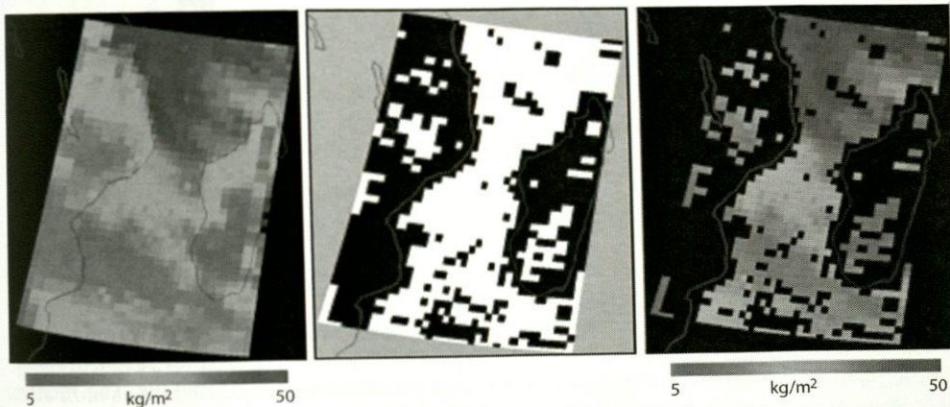


Figure 17. Total column precipitable water (kg/m<sup>2</sup>) from AIRS near Madagascar on 4 June 2009. (Left) raw data; (Center) quality mask; (right) post screen out-put. Red is  $\pm 5$  km<sup>2</sup>; blue is  $\pm 50$  kg/m<sup>2</sup>. (see colour plate 52)

to result in RWQEs and run-off from less developed, pristine areas that would not affect water quality during high run off events. Additionally, not all rivers are equipped with river gauges to measure discharge. A method that can incorporate the source of run-off in its predictions of RWQEs would likely be more accurate and dependable.

Here, we describe the methods used to create a model that predicts RWQEs using remotely-sensed precipitation data and then test the accuracy and usefulness of that model using *in-situ* observations of water quality. The ease of access and the regular spacing of the remotely acquired precipitation data are likely to lead to timely and accurate predictions of RWQEs that were not available by earlier methods.

The model was used to examine water quality in three different watersheds: the Saco, Kennebec, and Androscoggin River watersheds located in Maine and New Hampshire (Figure 18). All empty into the Gulf of Maine, a marginal sea in the northwest Atlantic Ocean. Maine is an ideal test

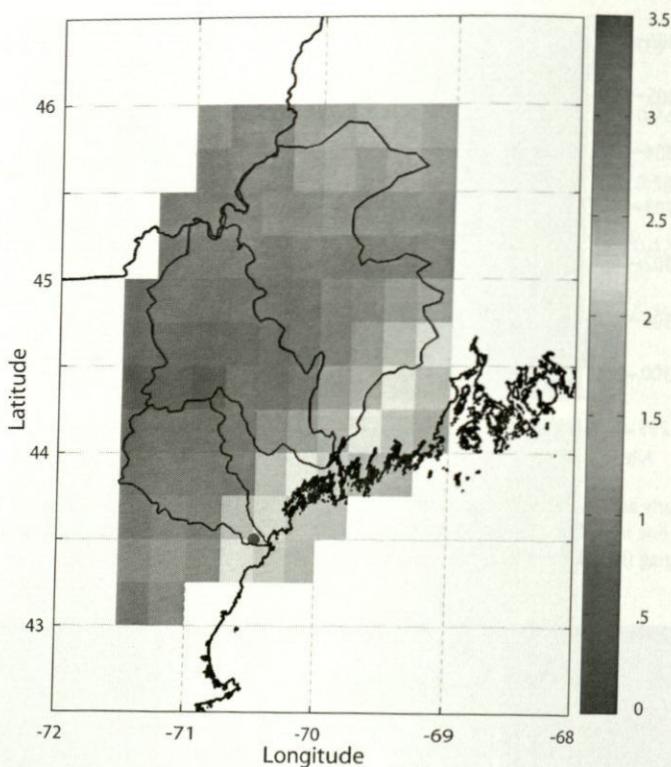


Figure 18. Mean daily precipitation (mm/day) observed for TRMM grid cells between 1998 & 2009. Black lines represent boundaries of three different watersheds. Blue circle is a water sampling site in Saco River Watershed. (see colour plate 53)

bed for this project because of its large number of rivers, varied watersheds, and strong reliance on coastal waters for both recreation and seafood. In addition, state agencies have little access to biological testing and are in need of new methods for the detection of RWQEs. The relatively low computational cost of the model (it can run on any laptop computer) allows one to examine different scenarios and instances, and their effects on water quality along the Atlantic coast.

#### 4.1.1 Data sets

To achieve the main objective of the project, both remote-sensing and *in-situ* data were used. Observations for the project were collected over a twelve year time period from 1998–2009 from TRMM's TMI sensor. It is able to distinguish precipitation variations at a grid cell resolution of  $0.25^\circ$ . Figure 18 shows the location of the different TRMM grid cells and outlines of the three watersheds. The colours within each cell represent the mean daily precipitation measured over the twelve year time period. Since the long-term goal of this project was to produce an easily accessible tool to users who might be unfamiliar with satellite data, no effort was made to process the downloaded data, but instead to determine how minimally manipulated data could be used to predict RWQEs. The TRMM data time series is available from Giovanni. The data were manually downloaded and processed in MATLAB. Eventually, the method will be exported to an Excel spread-sheet or some other easily accessed software tool. The large area of the watersheds (ranging from 4410 to 9100 km<sup>2</sup>) allowed for ten to twenty-five TRMM grid cells to occupy each watershed, providing details of small-scale spatial patterns of variability across a watershed. Data for water quality were obtained from state agencies to both calibrate and validate model outputs.

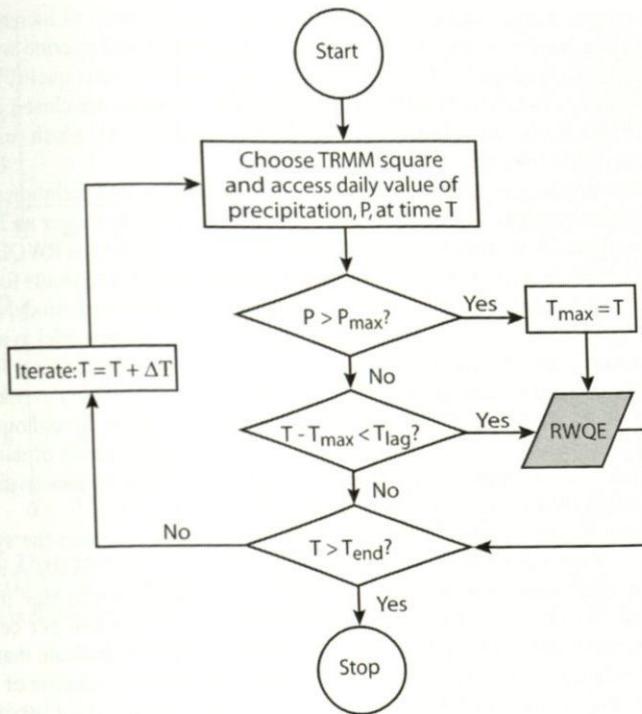


Figure 19. Flow diagram used to predict RWQEs from precipitation data.

#### 4.1.2 Comparison methods

Any method that predicts RWQEs must balance two competing goals: (1) accurate prediction of individual reduced water quality events and (2) reduction of false alarms (i.e. a predicted RWQE that did not materialize). Resource managers depend on knowledge of water quality in the coastal regions to make decisions in order to protect consumers of shellfish, users of recreational areas, and others from pollutants in the water. However, they are also under pressure to avoid costly shutdowns that remove revenue from beach-goers, consumers of shellfish, and others from local economies. Closing a healthy shellfish bed or beach due to an inaccurate prediction of reduced water quality (i.e. a false alarm) can have devastating consequences for local businesses. Consequently, resource managers need access to a reliable method that predicts as many RWQEs as possible with as few false alarms as possible.

To examine the link between remotely acquired precipitation data within a watershed to RWQEs downstream of the precipitation locations, the question was: *What amount of precipitation at some location within the watershed will result in RWQEs at a particular location downstream?* The general model used to predict the presence of RWQEs from precipitation data is represented in a flow diagram in Figure 19. Three variables were used to determine the optimum settings to predict RWQEs: (1)  $P_{\max}$ , the amount of rainfall that defines an RWQE-producing precipitation event is measured in mm/day over a two to twenty day period; (2) the location of TRMM grid cell; and (3),  $T_{\text{lag}}$ , the time period over which a particular precipitation event can effect a region measured over a three to seven day period.

To measure the accuracy of the model and determine the optimum amount of precipitation, ideal location, and time lag of precipitation for each water quality site, an examination was performed on the ability of the model to predict known RWQEs with the current method (river discharge) and to avoid false alarms. Maine's current method of predicting reduced water quality events relies

on measured discharge. Since higher river discharge is a direct result of increased run-off, there is a strong connection between measured discharge and coastal and riverine water quality. Each river is assigned a discharge that will result in a predicted reduced water quality event. When river discharge exceeds this value, shellfish beds and recreational areas are closed and testing of the region begins. For this study, we used a value of 1000 ft<sup>3</sup>/s (28 m<sup>3</sup>/s), which resulted in the most accurate prediction of RWQEs in the coastal ocean.

RWQEs can occur due to a number of different mechanisms, such as pollution events or wastewater treatment plant malfunctions. Although precipitation is likely to trigger an RWQE, there will be a number of precipitation events that do not result in RWQEs as well as RWQEs that occur even after no rainfall. To provide a quantitative measure of model skill that accounts for these additional events and to determine the level of significance of agreement between the model and observations, we used a simple randomization test, which compared the ability of the model to predict events and avoid false alarms with distributions from a large number of synthetic data sets. The synthetic data sets were constructed by determining the number of different precipitation events in the observed data set that exceeded  $P_{max}$  ( $n$  events) and then randomly placing those throughout the time period. The random data set is then used to predict RWQEs. Once a large number of data sets have been created for each water quality measurement location (100–1000), a frequency distribution of the percentage of predicted RWQEs and false alarms can be constructed.

Comparison of the model predicted RWQEs with the distribution from the synthetic data set provides an estimate of the significance of the model skill (Bishop *et al.*, 2010). A model prediction that is greater than ninety-five per cent of the synthetic prediction percentage indicates that the model skill outperforms random chance at a significance of ninety-five per cent. Model false alarms that are less than ninety-five per cent of the synthetic data set indicate that the model outperforms random chance at a significance of ninety-five per cent. An example of this is shown in Figure 20 for the location in the Saco River watershed shown in Figure 18 for two different TRMM cells. Examination of Figure 21 reveals that the location of the precipitation events has a very strong effect on the accuracy of the method. Precipitation within TRMM square seven provides no better predictive skill than random guessing. The fraction of the number of observed RWQEs successfully predicted by precipitation within TRMM square seven (0.28) is less than all of the synthetic data sets in which the time of RWQEs were randomly chosen throughout the time period, while the percentage of false alarms (0.67) is higher than eighty-five per cent of the synthetic data sets. Precipitation within TRMM square eighty-five can be used to predict RWQE (0.67 prediction rate and 0.45 false alarm fraction) with much greater accuracy than random chance, providing a valuable tool for resource managers. The river discharge model using a discharge threshold of 28 m<sup>3</sup>/s produces similar results, indicating that satellite recorded precipitation data can provide as accurate a predictive capability as more costly river discharge data. Note that  $P_{max} = 8$  mm/day and  $T_{lag} = 3$  days for both TRMM cells and random synthetic data sets. A perfect method would predict all of the RWQEs (Fraction = 1.0) and have no false alarms (Fraction = 0.0). Varying the amount of precipitation that can trigger an RWQE also results in quite different effects on RWQE prediction and false alarms (Figures 21 & 22). A low threshold ( $P_{max}$ ) results in the method predicting almost all of the observed RWQEs, while a high  $P_{max}$  results in the method missing a large number of observed RWQEs. Interestingly, an increase in  $P_{max}$  does not result in a monotonic decrease in the number of false alarms, indicating that large precipitation events do not always result in RWQEs.

#### 4.1.3 Conclusions

The use of readily available precipitation data at a spatial resolution of eighteen to twenty five kilometres may provide an ideal method for water resource managers to aid in their prediction of RWQEs. However, the method must be tested and the model calibrated before it can be used by resource managers. Examination of the method's statistical significance using a randomization test suggests that the method has promise. TRMM data provide for an easily accessible, effective method of predicting water quality, although the model still requires a degree of calibration for individual downstream sites.

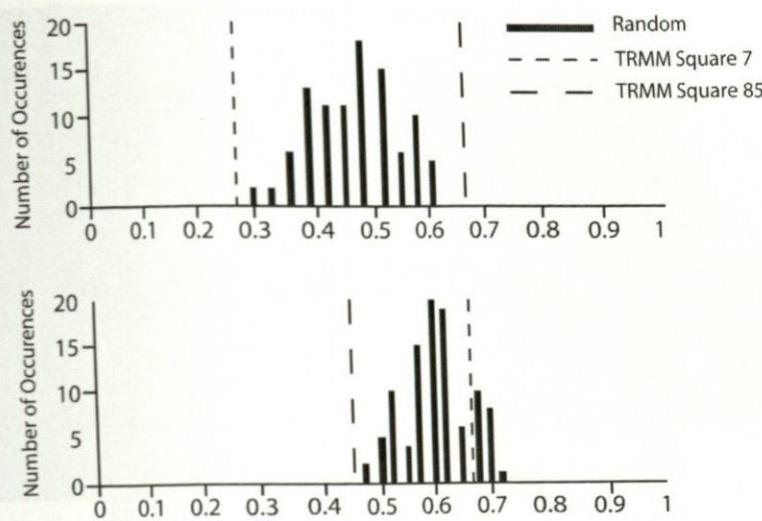


Figure 20. (Top): The fraction of RWQEs (x-axis) preceded by high precipitation events. The dashed vertical lines are for TRMM square 7 (left) and TRMM square 85 (right); (Bottom): The fraction of false alarms predicted by high precipitation events. The dashed vertical lines are for TRMM square 85 (left) and TRMM square 7 (right).

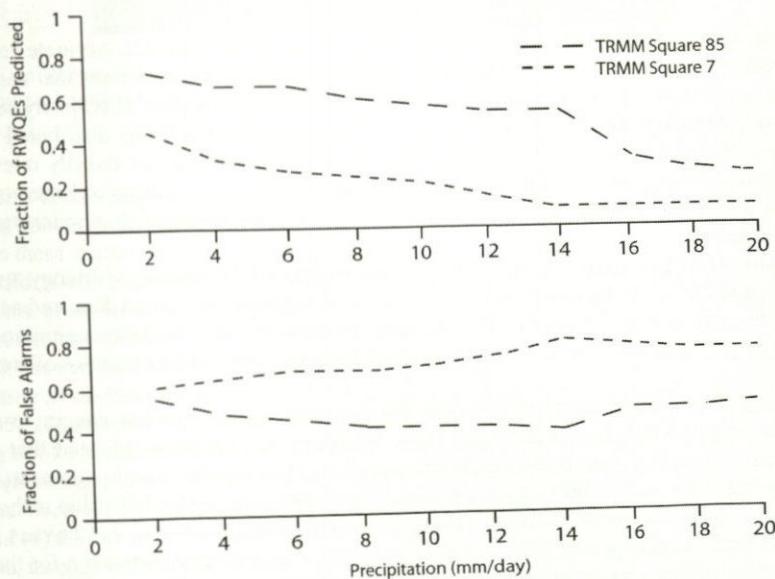


Figure 21. (Top): The fraction of RWQEs (y-axis) predicted by high precipitation events as a function of  $P_{\max}$ ; (Bottom): The fraction of false alarms for high precipitation events as a function of  $P_{\max}$ .

#### 4.2 Developing a detection and early warning system for malaria risk in the Amazon

This application makes use of satellite-derived data in a number of ways. The first and most obvious motivation for satellite analysis is the fact that it is very difficult to perform land surveys

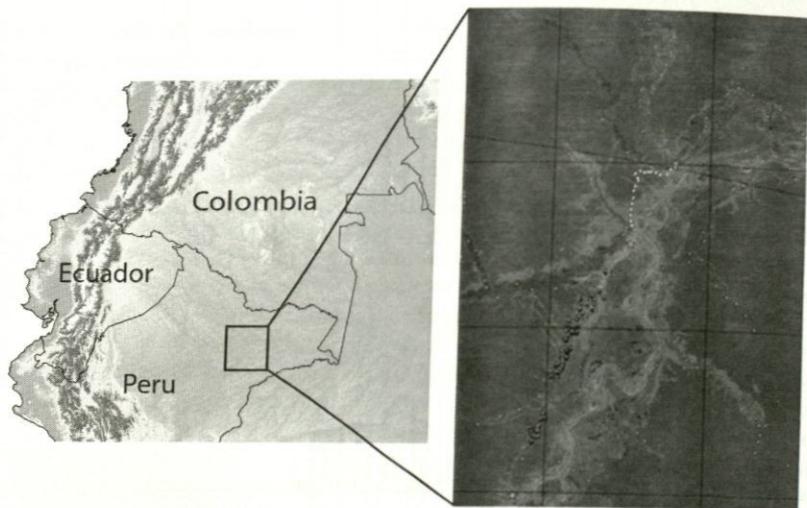


Figure 22. Portion of the Peruvian Amazon, shown on a 432-RGB Landsat TM composite from 22 August 2005. Red and yellow points show mosquito collection sites along the Nauta-Iquitos and Iquitos-Mazan roads, respectively. Green points are settlements. (see colour plate 54)

over thousands of square kilometres in the Peruvian Amazon (Figure 22). Accurate characterization of evolving land cover, of shifting rivers, and of topographic variations that can influence hydro-ecology and mosquito breeding sites, all require the application of remotely sensed data. Second, satellite platforms provide novel observations to inform spatially distributed hydrological and meteorological analysis over large areas. These observations are directly relevant to the characterization of mosquito breeding habitats, and they represent an unprecedented opportunity to integrate observation-based eco-hydrological analysis into disease risk monitoring and early warning systems.

In the context of this application, satellite observations of landscape, hydrology, and climate are used both in offline analysis of *in situ* mosquito and malaria case count data, and as inputs for the land surface models at the core of the Peruvian Amazon frontier land data assimilation system (PAF-LDAS) being implemented to integrate data streams relevant to estimates of transmission risk.

One example of offline satellite-based analysis is the use of Landsat images to characterize LCLUC around mosquito monitoring sites. Previous studies have indicated that counts of mosquitoes at monitoring sites in the Peruvian Amazon can be correlated with patterns in surrounding land cover (Vittor *et al.*, 2009). One proposed explanation for this relationship is that changes in land cover, including deforestation, change local hydrology in a way that favours mosquito breeding; indeed, the promise of this explanation serves as a primary motivation for the coupled weather/hydrology/land cover analysis at the heart of the current project. As a starting point, then, it is important to confirm that the relationship between land cover and mosquito density holds in the current study area for recently established mosquito collection sites. Next, the spatial scale of land cover influence on mosquito density must be further investigated to inform skilful extrapolation of the method to regions that lack active mosquito monitoring sites. To accomplish these analyses, Landsat-derived land cover characteristics within a radius of 250 m to 10 km of mosquito collection sites are being tested as predictive variables for mosquito counts, both as independent predictors and as inputs to a multivariate statistical model. Along this same line of questioning, the investigators are examining temporal relationships between land use change and mosquito presence: are

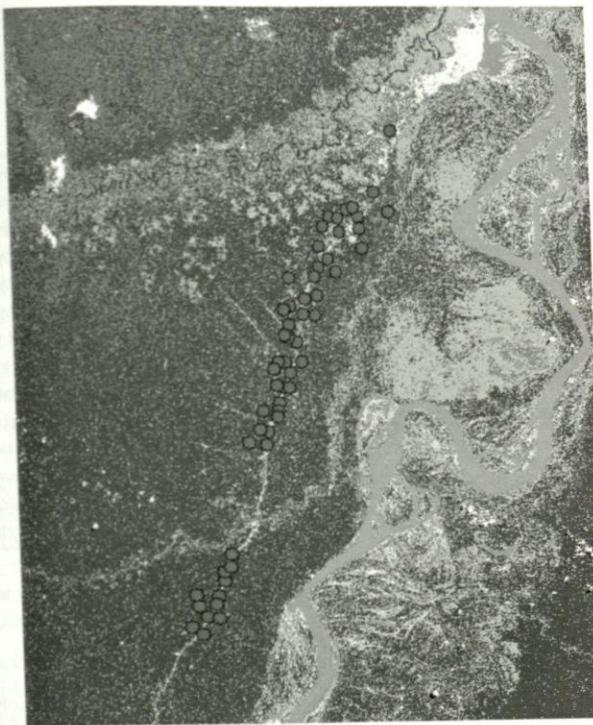


Figure 23. Collection sites reported along the Nauta-Iquitos road, overlain on a supervised classification of a Landsat image. Iquitos is located in the northeast corner of the image. (see colour plate 55)

the mosquito counts largest in sites near recently cleared forest or in sites surrounded by long-established human settlements? The archived Landsat record offers an unparalleled opportunity to investigate these spatio-temporal dynamics, as shown in Figure 23.

Initial efforts with single-scene Landsat image classifications have proved to be quite promising. Nevertheless, given the known complications in detecting tropical forest disturbance from satellite over larger scales (Oliveira *et al.*, 2007), the research team is prepared to employ a number of more sophisticated land cover analysis techniques, including: (1) supervised and unsupervised classification of stacked seasonal Landsat images from a single year (Ichii *et al.*, 2003; Matricardi *et al.*, 2005), (2) unsupervised classification of Landsat image stacks from multiple years, in which clustering of like pixels can distinguish between pixels that have undergone change and those that have not (Steininger *et al.*, 2001), (3) spectral end-member un-mixing of single-scene Landsat images that accounts for green vegetation, senescent vegetation, and bare ground (Oliveira *et al.*, 2007), and (4) Fourier cycle similarity analysis of MODIS 16-day NDVI composites for each year of analysis, through which differences in phenology can be used to classify different LUC and end-member un-mixing is applied to determine per cent cover within 250m MODIS pixels (Geerken *et al.*, 2005; Evans & Geerken 2006). These advanced techniques are expected to be particularly important as the project scales up from well-known sites around Iquitos to the broader Peruvian Amazon region. Evaluation of land cover classifications is being performed through field survey with local experts as well as comparisons with selected high-resolution commercial imagery.

When integrated to PAFLDAS, these land cover classifications represent one of numerous satellite-derived products used to inform ecologically-based mosquito predictions. An LDAS is a numerical modelling scheme that integrates observations from various sources within such models, using data assimilation and other techniques to produce optimal maps of land surface states

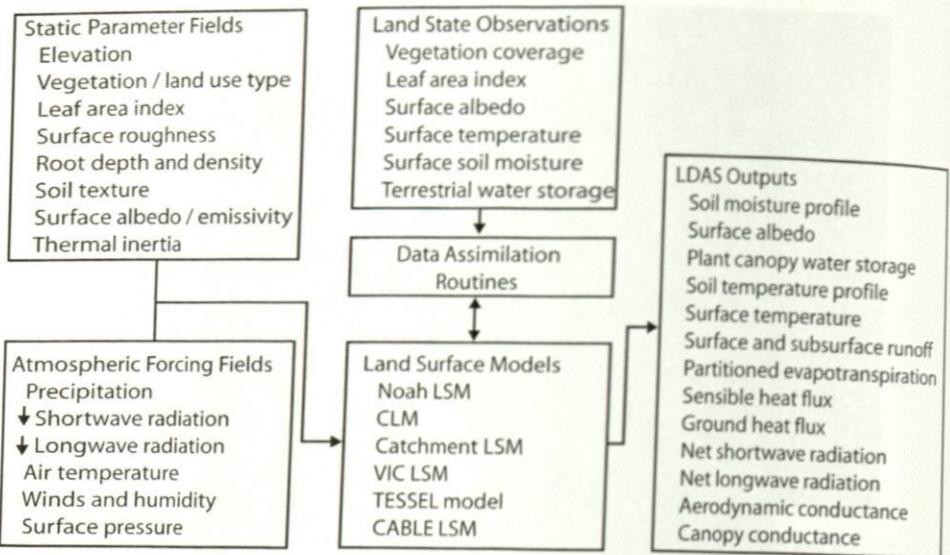


Figure 24. Components of an LDAS. Some parameters (e.g. surface albedo) appear in multiple boxes and can be updated as an observational input and subsequently predicted as an LDAS output.

and fluxes (Figure 24). Data assimilation is a computational process by which two independent estimates of a variable can be combined to determine one best estimate, which is less uncertain than either of the two inputs. In this case, the variable is a land surface state such as soil moisture. One estimate comes from the land surface model and the other from a satellite observation. The error characteristics of each are used to weight their contributions to the final estimate. Other projects have demonstrated that these techniques can be implemented at regional and global scales (Mitchell *et al.*, 2004; Rodell *et al.*, 2004).

Land data analysis system models use physical equations to simulate the storage and movement of water and energy on and within the land surface. Net surface radiation, for example, is typically solved as:

$$R_{net} = (1 - \alpha) \cdot S_{\downarrow} + \varepsilon \cdot L_{\downarrow} - \varepsilon \sigma T^4, \quad (1)$$

where  $\alpha$  and  $\varepsilon$  are surface properties albedo and emissivity, respectively,  $\sigma$  is the Stephan-Boltzmann constant,  $T$  is simulated surface temperature, and  $S_{\downarrow}$  and  $L_{\downarrow}$  are incoming shortwave and longwave radiation drawn from atmospheric data. Net radiation is then partitioned into surface energy fluxes and storages using additional physically equations.

PAFLDAS will include four initial land surface models in parallel analysis: 1) Noah (Chen *et al.*, 1996; Ek, *et al.*, 2003); 2) the variable infiltration capacity (VIC) model (Liang *et al.*, 1996); 3) Catchment (Koster *et al.*, 2000); and 4) the Common Land Model (CLM) (Dai *et al.*, 2003). Additional land surface models are being considered also. These include the tiled ECMWF scheme for surface exchanges over land (TESSEL) (Van den Hurk *et al.*, 2000) and CSIRO's atmosphere biosphere land exchange (CABLE) model (Kowalczyk *et al.*, 2006).

The basic implementation of PAFLDAS has been accomplished at one kilometre resolution (Figure 25), with plans to improve data inputs to achieve an ultimate spatial resolution of 500 metres, which is sufficient to capture most features of the Amazonian land use mosaic. LDAS techniques have been applied at sub-kilometre resolutions before, but modelling at this scale is not common. The low topographic relief of the study area allows for confidence in the application of one-dimensional land surface models, as the lateral transfer of moisture across grid cells is unlikely to be a significant influence on near-surface conditions. At the same time, land use transitions are quite dramatic, so there is reason to believe that high-resolution application of LDAS models

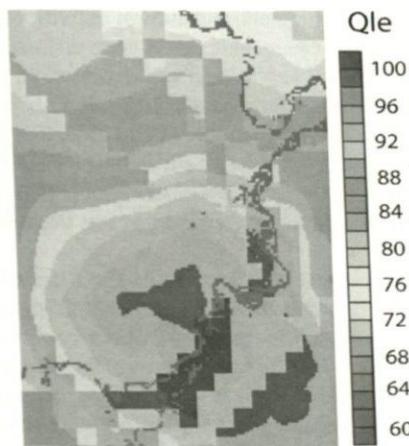


Figure 25. Monthly averaged latent heat flux ( $\text{W}/\text{m}^2$ ) for February 2009 over a portion of the PAFLDAS domain. Data cells have different sizes due to differing resolution of input parameters and forcing data, but can be improved as LDAS incorporates additional satellite data sets and downscaling techniques. (see colour plate 56)

and assimilation techniques will add value to the analysis of conditions relevant to malaria risk. In implementing PAFLDAS, the research team is focused on the representation of temporal and spatial variability in near surface soil moisture, surface temperature and humidity, and the impact of land use on hydrologic states. Data assimilation efforts are challenged by the fact that available passive microwave soil moisture observations from AMSR-E have coarse spatial resolution and very limited reliability in tropical forest environments. For this reason, the team is particularly focused on the potential of the future SMAP mission to provide soil moisture estimates for PAFLDAS, as well as on innovative assimilation options such as GRACE (Zaitchik *et al.* 2008), which will become more relevant when the feasibility study is scaled to larger areas, and thermally-based soil moisture estimates (Anderson *et al.*, 2007; Meng *et al.*, 2009).

All remote sensing and data assimilation work in this project is motivated by the value these analyses have for the development of spatially-distributed malaria risk monitoring and early warning systems. Land cover analysis and PAFLDAS fields will drive spatial-temporal models of *Anopheles* systems. These models, derived using mosquito databases collected across four areas around Iquitos over four different time periods, will use a multilevel Poisson modelling framework (Raudenbush & Bryk 2002). Since results will be used for development of risk maps, models will be fit using Bayesian analysis to allow more effective means of smoothing the map surface (Clayton & Kaldor 1987). This analysis will also involve examination of land fragmentation metrics to capture the importance of landscape structure and function on *Anopheles* distribution (Turner 1989). Once appropriate scales and fragmentation measures are identified, a multilevel simultaneous equation model will be used to estimate jointly the presence of larva and adults, particularly for the Iquitos-Nauta road where larva and adult data were collected concurrently over time. These models will be used to create predictive models of *Anopheles* presence for the entire study region.

Beyond ecological modelling, mosquito prediction models and environmental determinants need to be linked to malaria risk through accurate measures of population exposure. The research team will develop models to link environmental changes like forest conversions and changes in non-forest cover, to regions of permanent and temporary human occupation including urban expansion, agricultural expansion, logging, oil exploration activities, and mining. The goal will be to create plausible population risk sets that are defined not just where a person lives, but by their social and economic spaces that are directly and indirectly linked to land use and land use cover (LCLUC). The resulting output will be seasonal *Human Activity and Settlement Maps* that depict population

density based on permanent human settlement, and human activities that show regions of recent and long-term anthropogenic LCLUC. These include forest conversion due to logging and agriculture, urban expansion, road construction, oil exploration and production, mining, and other identifiable forms of LCLUC. Three types of data sources are being used for: 1) census data of the regional population (1981, 1993, 2005, 2007) and agriculture (1994 & 2010); 2) demographic and economic survey data available from demographic and health surveys (DHS: 1991–92, 1996, 2000, 2004–08, 2009) and the World Bank's living standards measurement surveys (LSMS: 1985, 1991, 1994); and 3), data related to development projects such as road paving between 1995 and 2001, and forest and oil exploration concessions.

Annual population density will be estimated for all of the State of Loreto as well as at smaller geographic subdivisions (provinces and districts) by first estimating annual intercensal population for the Loreto area by age-sex composition using the cohort-component method (Shryock *et al.*, 1976). This method will be adjusted accordingly based on fertility, mortality and migration data obtained from DHS and LSMS surveys as well as national mortality data when available. Second, population will be disaggregated by province and district by computing population percentages for each sub-region in each census year and conducting a constrained interpolation to produce annual estimates of population distribution. Finally, population will be further disaggregated within districts by assigning population values to known cities and towns like Iquitos, Nauta, and Mazan, and by using the LCLUC output to identify permanent areas of human settlement. The *Human Activity and Settlements Map* will identify at least crudely those areas undergoing or potentially undergoing change. Non-forested areas that have not experienced change for a minimum of five years will be considered urban or agricultural based on their particular land cover. Forested areas that have not undergone change for five years will be identified as potential areas of human activity, while areas that have undergone change in the previous five years will be identified as current areas of human activity. The Map will take advantage of auxiliary data on road construction, infrastructure development, and logging and oil concession areas. With the benefit of the Landsat data archive and retrospective PAFLDAS simulations, the research team will examine several spatial scales to identify areas as well as examine the sensitivity of the five year cut-off.

Results of the Human Settlements and Activity analysis and the spatially explicit *Anopheles* ecological models will provide inputs to malaria transmission models that can be used predictively for risk mapping and early warning. These applications are of considerable importance to Ministries of Health operating in Iquitos and in other malaria-prone regions of the Amazon.

#### 4.3 Linking environmental data and public health cohort data to enhance decision making

##### 4.3.1 Project objectives and research design

There are seven objectives in this application effort: 1) produce daily gridded estimates of PM<sub>2.5</sub> for the conterminous US for 2003 to 2008 using MODIS Aqua data; 2) produce daily gridded solar insolation maps for the same area and time frame using data from the NARR; 3) produce daily gridded LST maps from MODIS; 4) link the estimates of PM<sub>2.5</sub>, insolation and LST with data from more than 30,000 participants in REGARDS; 5) determine whether exposure to PM<sub>2.5</sub> or solar insolation are related to the rate of cognitive decline among the participants in the REGARDS study, independent of other known risk factors for cognitive decline; 6) examine the relationships between the estimated PM<sub>2.5</sub> and insolation and other health-related conditions among REGARDS participants, including diminished kidney function, hypercholesterolemia, hypertension, and inflammation measured by C-reactive protein; and 7) deliver daily gridded environmental data sets (PM<sub>2.5</sub>, SI, and LST) to CDC-WONDER for the six year period.

A summary of our research methodology is provided in Figure 26. It shows the different data elements, how they will be linked, and the outputs generated from the project. In addition, it provides a description of the decision support system and identifies who potential end-users may be. Scientists associated with NASA's Marshall Space Flight Center (MSFC) downloaded and processed daily AQUA MODIS AOD data for the conterminous US for 2003–2008. The data processing included extracting and mosaicking different swaths of data into one national dataset.

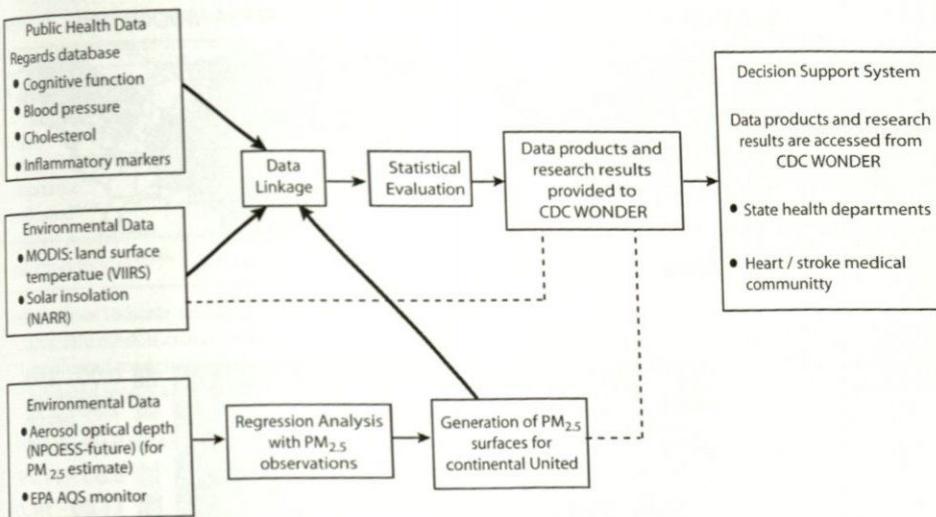


Figure 26. Schematic describing the research methodology (Courtesy McClure, UAB and Alhamdan USRA).

Part of the processing was also to estimate ground-level fine particulates (PM<sub>2.5</sub>) from MODIS AOD using regression equations per EPA region per season from (Zhang *et al.*, 2009). The MSFC team also obtained and processed EPA Air Quality System (AQS) PM<sub>2.5</sub> data for the whole conterminous US for 2003–2008. The regional surfacing algorithm of Al-Hamdan (*et al.*, 2009) was modified and used to generate continuous spatial surfaces of daily PM<sub>2.5</sub> for the conterminous US, which also involved a quality control procedure for the EPA AQS data and a bias adjustment procedure for the MODIS data (Figure 27).

PM<sub>2.5</sub> data processing is nearly complete. The surfacing algorithm has been modified and the MODIS and AQS data are processed; thus, preparation of daily national surfaces of PM<sub>2.5</sub> is underway. This procedure also includes a quality control element for examining the US/EPA data, as well as a bias adjustment for the MODIS data. Upon completion of these surfaces, the PM<sub>2.5</sub> data will be linked spatially and temporally to the REGARDS participants, and the data incorporated with the public health data.

In Figure 28 solar insolation (SI) and LST data sets have been linked to the REGARDS participants and analyses of the linked data have begun. The original aim of the project did not include an assessment of the association between SI, LST and stroke. However, the project assessed several periods of exposure to SI and LST to utilize the most robust measure in models describing the relationship between SI and stroke. The determination was that the one-year prior to baseline was most appropriate to include as the primary exposure. Table 14 provides the hazard ratios and ninety-five per cent confidence intervals for the associations between SI and stroke, and for maximum temperature averages and stroke, when exposure is defined as occurring the year prior to the in-home interview. The results for solar insolation indicate that for those participants exposed to SI at levels below the median, the risk of stroke is 1.44 times higher than for those exposed to SI at levels above the median (95% CI: 1.16, 1.79), but that after the addition of LST, the risk of death is actually higher for those exposed to lower levels of SI, compared to those exposed to higher (HR: 1.73, 95% CI: 1.25, 1.29). This association is not attenuated after multivariable adjustment for known stroke risk factors, such as demographics, behavioural and medical risk factors (HR: 1.61, 95% CI: 1.15, 2.26). Similarly, there is a U-shaped association between temperature and stroke, with those exposed to the lowest temperatures, and those exposed to the higher temperatures, at higher risk for stroke than those exposed to mid-temperatures. These patterns are detailed in Table 14, including after multivariable adjustment.

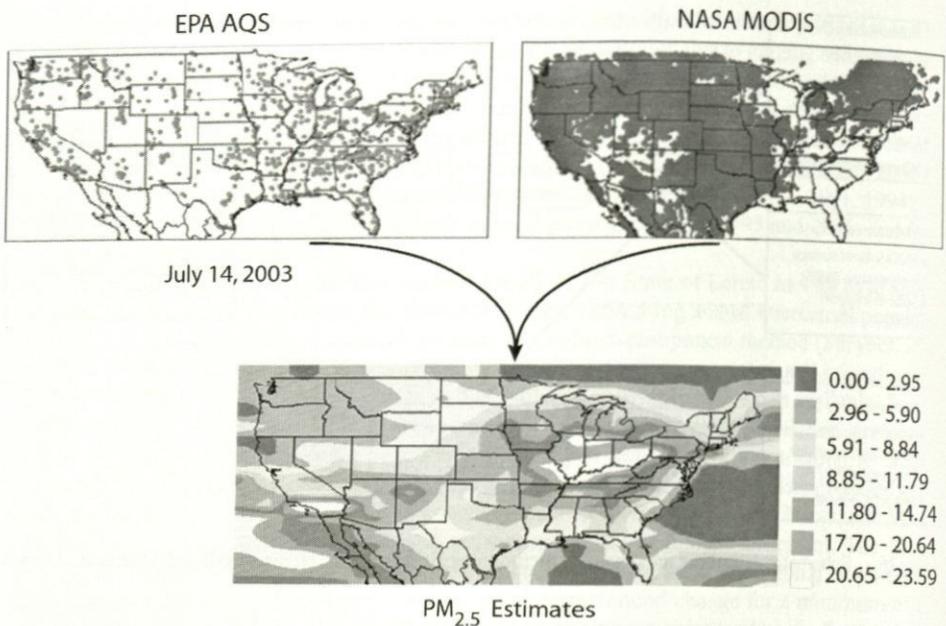


Figure 27. A schematic showing how EPA/AQS data are merged with NASA MODIS data to form a smoothed PM<sub>2.5</sub> surface (Courtesy McClure & Alhamdan, USRA). (see colour plate 57)



Figure 28. The REGARDS participants; those in red are whites, those in blue are African Americans. (see colour plate 58)

In addition to the assessment of stroke incidences, the project has assessed the relationship between SI and cognition. This assessment utilizes the *six-item screener* as the cognitive outcome, and includes LST as a potential confounder of the relationship between SI and cognition (Callahan et al., 2002). The analysis included an assessment of exposure periods to determine which were most robust to include. Again the 1-year prior to baseline was determined to be most suitable for modelling. The relationship between SI, LST and cognitive impairment was more complicated than for stroke. The interaction between SI and cognitive impairment exists but the relationship between them differs as a function of quartile of LST. Thus, it is not possible to provide results for each association separately. However after multivariable adjustment among those participants for whom LST exposures were in the 1st tertile; that is, those exposed to lower SI levels, were 1.26 times more likely to experience incident cognitive impairment than those exposed to higher levels (95%

Table 14. Hazards ratios and 95% confidence intervals for associations between insolation and maximum temperature averages for the year previous to in-home interview with stroke incidence (N = 16,529).\*

Model	Below median insolation	Quartile of maximum temperature			
		1st	2nd	3rd	4th
Univariable**	<i>1.44 (1.16, 1.79)</i>	<i>1.86 (1.39, 2.52)</i>	Ref	1.22 (0.88, 1.69)	1.36 (0.99, 1.89)
Unadjusted***	<i>1.73 (1.25, 2.29)</i>	<i>1.68 (1.25, 2.29)</i>	Ref	<i>1.51 (1.07, 2.14)</i>	<i>2.05 (1.36, 3.11)</i>
Adjusted model****	<i>1.59 (1.14, 2.23)</i>	<i>1.49 (1.05, 2.14)</i>	Ref	<i>1.73 (1.19, 2.51)</i>	<i>1.95 (1.29, 2.96)</i>
Mediation model*****	<i>1.61 (1.15, 2.26)</i>	1.41 (0.99, 2.03)	Ref	<i>1.69 (1.17, 2.46)</i>	<i>1.91 (1.27, 2.91)</i>

\*Italic values indicate variables with chi-square p-values <0.05

\*\*Univariate models contain only a measure of insolation or temperature with stroke of any subtype as outcome

\*\*\*Unadjusted model contains both measures of insolation or temperature with stroke of any subtype as outcome

\*\*\*\*Adjusted model adds all demographic, behavioural and medical confounders from Table 1 to the unadjusted model

\*\*\*\*\*Mediation model adds medical mediators to the adjusted model

CI: 0.94, 1.68); for those in the 2nd tertile of LST exposure, participants with lower SI exposure levels were 1.30 times more likely to experience incident cognitive decline than those with higher SI levels; and for those in the 3rd tertile of LST exposure, participants with lower SI exposure levels were 1.95 times more likely to experience incident cognitive impairment than those with higher SI levels. A similar analysis assessing the relationship between SI and other cognitive measures is underway for Word List Learning (assesses memory) and Animal Naming (assesses executive function).

#### 4.3.2 Acknowledgement

The authors wish to acknowledge D. Quattrochi & D. Rickman (NASA); W. Crosson, S. Estes, M. Estes, S. Hemmings, M. Alhamdan & G. Wade (USRA); and S. Kent & L. McClure (UAB) for their contributions and support of the REGARDS project.

## 5 CONCLUSION

The purpose of this chapter is to broaden veteran remote sensing data users' understanding, and to inspire new remote sensing data users in the latest uses of satellite data for public health research. Experts in the field have come together to describe their work in their particular areas of research. In doing so, they have also provided insight to the data sets they use, the tools they build, the research methods they create, and the relationships they discover. It is hoped by the authors that this chapter will entice further public health research using remote sensing data and information services, and to tease out new ideas that further advance public health research.

## REFERENCES

- Adams, R.M. & Croker, T.D. 1989. The agricultural economics of environmental chance: Some lessons from air pollution. *J. Environ. Manage.* 28: 295–307.
- Adimi, F.A., Soebiyanto, R.P., Safi, N. & Kiang, R. 2010. Towards malaria risk prediction in Afghanistan using remote sensing. *Malaria J.* 9: 125. doi:10.1186/1475-2875-9-125.
- Al-Hamdan, M., Crosson, W., Limaye, A., Rickman, D., Quattrochi, D., Estes, M., Qualters, J., Sinclair, A., Tolsma, D., Adeniyi, K. & Niskar, A. 2009. Methods for characterizing fine particulate matter using ground observations and satellite remote-sensing data: Potential use for environmental public health surveillance. *J. Air & Waste Manag. Assoc.* 59: 865–881.

- Anderson, M.C., Kustas, W.P. & Norman, J.M. 2007. Up-scaling flux observations from local to continental scales using thermal remote sensing. *Agron. J.* 99: 240–254.
- Anenberg, S.C., Horowitz, L.W., Tong, D.Q. & West, J.J. 2010. An estimate of the global burden of anthropogenic ozone and fine particulate matter on premature human mortality using atmospheric modelling. *Environ. Health Perspect.* 118: 1189–1195.
- Anyamba, A., Chretien, J-P., Small, J., Tucker, C.J., Formenty, P.B., Richardson, J.H., Britch, S.C., Schnabel, D.C., Erickson, R.L. & Linthicum, J. 2009. Prediction of a Rift Valley fever outbreak. *PNAS* 106(3): 955–959.
- Aumann, H.H. & Pagano, R.J. 1994. Atmospheric infrared sounder on the Earth observing system. *Optical Engin.* 33: 776–784.
- Avnery, S., Mauzerall, D.L., Liu, J. & Horowitz, L.W. 2011. Global crop yield reductions due to surface ozone exposure: 1. Year 2000 crop production losses & economic damage. *Atmos. Environ.* 45: 2284–2296.
- Balcan, D., Colizza, V., Gocals, B., Hu, H., Ramasco, J., Vespignani, A. 2009. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Nat. Acad. Sci.* 106(51): 21484–21489.
- Balk, D., Brickman, M., Anderson, B., Pozzi, F. & Yetman, G. 2005. Mapping global urban and rural population distributions: Estimates of future global population distribution to 2015. UNFAO & CIESIN. Available from: [http://sedac.ciesin.columbia.edu/gpw/docs/GISn.24\\_web\\_gpwAnnex.pdf](http://sedac.ciesin.columbia.edu/gpw/docs/GISn.24_web_gpwAnnex.pdf) [Accessed 19th January 2012].
- Barker, C.M., Kramer V.L. & Reisen W.K. 2010. Decision support system for mosquito and arbovirus control in California. *Earthzine*. September 24th. Available from: <http://www.earthzine.org/2010/09/24/decision-support-system-for-mosquito- and -arbovirus-control-in-california/> [Accessed 19th January 2012].
- Beck, L.R., Rodriguez, M.H., Dister, S.W., Rodriguez, A.D., Rejmankova, E., Ulloa, A., Meza, R.A., Roberts, D.R., Paris, J.F., Spanner, M.A., Washino, R.K., Hacker, C. & Legters, J. 1994. Remote sensing as a landscape epidemiological tool to identify villages at high risk for malaria transmission. *Amer. J. Trop. Med. & Hyg.* 51(3): 271–280.
- Beck, L.R., Lobitz, B.M. & Wood, B.L. 2000. Remote sensing and human health: New sensors and new opportunities. *Emerg. Infect. Dis.* 6(3): 217–226.
- Beer, R., Glavich, T.A. & Rider, D.M. 2001. Tropospheric emission spectrometer for the Earth observing system's Aura satellite. *Appl. Optics* 40: 2356–2367.
- Bell, M.L. 2006. The use of ambient air quality modelling to estimate individual and population exposure for human health research: A case study of ozone in the Northern Georgia region of the United States. *Environ. Int.* 32(5): 586–593.
- Bell, M.L., McDermott, A., Seger, S.L., Samet, J.M. & Dominici, F. 2004. Ozone and short-term mortality in 95 US urban communities, 1987–2000. *JAMA* 292(19): 2372–2378.
- Bell, M.L., Dominici, F. & Samet, J.M. 2005. A meta-analysis of time-series studies of ozone and mortality with comparison to the national morbidity, mortality, and air pollution study. *Epidemiol.* 16 (4): 436–445.
- Benedict, K., Yang, P., Huang, Q. 2010. Project final report & feasibility report. NASA Applied Sciences Program cooperative agreement number NNX09AN53G.
- Bernath, P.F., McElroy, C.T., Abrams, M.C., Boone, C.D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P.F., Colin, R., DeCola, P., DeMazière, M., Drummond, J.R., Dufour, D., Evans, W.F.J., Fast, H., Fussen, D., Gilbert, K., Jennings, D.E., Llewellyn, E.J., Lowe, R.P., Mahieu, E., McConnell, J.C., McHugh, M., McLeod, S.D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F., Nowlan, C., Rinsland, C.P., Rochon, Y.J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J.J., Soucy, M.-A., Strong, K., Tremblay, P., Turnbull, D., Walker, K.A., Walkty, I., Wardle, D.A., Wehrle, V., Zander, R. & Zou, J. 2005. Atmospheric Chemistry Experiment (ACE): Mission overview. *Geophys. Res. Lett.* 32, L15S01. doi:10.1029/2005GL022386.
- Bishop, T.D., Miller III, H.L., Walker, R.L., Hurley, D.H., Menken, T. & Tilburg, C.E. 2010. Blue crab (*Callinectes sapidus* Rathbun, 1896) settlement at three Georgia (USA) estuarine sites. *Estuaries & Coasts* 33: 688–698.
- Borden, K.A. & Cutter, S.L. 2008. Spatial patterns of natural hazards mortality in the United States. *Int. J. Health Geogra.* 7:64. doi:10.1186/1476-072X-7-64.
- Bovensmann, H., Burrows, J.P., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V.V., Chance, K.V., & Goede, A.P.H. 1999. SCIAMACHY: Mission objectives and measurement modes. *J. Atmos. Sci.* 56: 127–150.
- Brandon, K. 2010. Maps: How mankind remade nature. Available from: <http://www.wired.com/wiredscience/2010/08/new-anthrome-maps/> [Accessed 13th April 2012].
- Brooker, S., Leslie, T., Koaczinski, K., Mohsen, E., Mehboob, N., Saleheen, S., Khudonazarov, J., Freeman, T., Clements, A., Rowland, M. & Kolaczinski, J. 2006. Spatial epidemiology of *Plasmodium vivax*, Afghanistan. *Emerg. Infect. Dis.* 12(10): 1600–1602.

- Burrows, J.P., Weber, M., Buchwitz, M., Rozanov, V., Ladstatter-Weissenmayer, A., Richter, A., DeBeek, R., Hoogen, R., Bramstedt, K., Eichmann, K.U. & Eisinger, M. 1999. The global ozone monitoring experiment (GOME): Mission concept and first scientific results. *J. Atmos. Sci.* 56: 151–175.
- Callahan, C.M., Unverzagt, F.W., Hui, S.L., Perkins, A.J. & Hendrie, H.C. 2002. Six-item screener to identify cognitive impairment among potential subjects for clinical research. *Med. Care* 40(9): 771–781.
- Caminade, C., Ndione, J.A., Kebe, C.M.F., Jones, A.E., Danuor, S., Tay, S., Tourre, Y.M., Lacaux, J.P., Vignalles, C., Duchemin, J.B., Jeanne, I. & Morse, A.P. 2011. Mapping Rift Valley fever and malaria risk over West Africa using climatic indicators. *Atmos. Sci. Lett.* 12: 96–103. doi:10.1002/asl.296.
- Ceccato, P., Gobron, N., Flasse, S., Pinty, B. & Tarantola, S. 2002. Designing a spectral index to estimate vegetation water content from remote sensing data (Part 1: theoretical approach). *Rem. Sens. Environ.* 82 (2–3): 188–197.
- Charland, K.M., Buckeridge, D.L., Sturtevant, J.L., Melton, F., Reis, B.Y., Mandl, K.D. & Brownstein, J.S., 2009. Effect of environmental factors on the spatio-temporal patterns of influenza spread. *Epidemiol. Infect.* 137(10): 1377–1387.
- Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H.L., Koren, V., Duan, Q.Y., Ek, M. & Betts, A. 1996. Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *J. Geophys. Res. Atmos.* 101(D3): 7251–7268.
- Christophers, S.R. 1911. Malaria in the Punjab. *Sci. Mem. Med. & Sanit. Deps. India*. New Series 46: 197.
- Clayton, D. & Kaldor J. 1987. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 43(3): 671–681.
- ClearLead. 2011. Available from: <http://www.clearleadinc.com/site/remote-sensing.html> [Accessed 18th January 2012].
- Clerbaux, C., George, M., Turquety, S., Walker, K.A., Barret, B., Bernath, P., Boone, C., Borsdorff, T., Cammas, J.P., Catoire, V., Coffey, M., Coheur, P.F., Deeter, M., De Maziére, M., Drummond, J., Duchatelet, P., Dupuy, E., de Zafra, R., Eddounia, F., Edwards, D.P., Emmons, L., Funke, B., Gille, J., Griffith, D.W.T., Hannigan, J., Hase, F., Höpfner, M., Jones, N., Kagawa, A., Kasai, Y., Kramer, I., Le Flochmoën, E., Livesey, N.J., López-Puertas, M., Luo, M., Mahieu, E., Murtagh, D., Nédélec, P., Pazmino, A., Pumphrey, H., Ricaud, P., Rinsland, C.P., Robert, C., Schneider, M., Senten, C., Stiller, G., Strandberg, A., Strong, K., Sussmann, R., Thouret, V., Urban, J. & Wiacek, A. 2008. CO measurements from the ACE-FTS satellite instrument: Data analysis and validation using ground-based, airborne and spaceborne observations. *Atmos. Chem. & Physics* 8: 2569–2594.
- CNES. 2011. Available from: <http://www.cnes.fr/web/CNES-en/5073-monitoring-and-predicting-epidemics-with-satellites.php> [Accessed 18th January 2012].
- Cohen, A.J., Anderson, H.R., Ostro, B., Pandey, K.D., Krzyzanowski, M., Kuenzli, N., Gutschmidt K., Pope C.A., Romieu I., Samet J.M. & Smith, K.R. 2004. Mortality impacts of urban air pollution. In M. Ezzati, A.D. Lopez, A. Rodgers & C.J.L. Murray (eds.), *Comparative quantification of health risks: Global and regional burden of disease due to selected major risk factors*, Vol. 2: 1353–1433. Geneva: WHO.
- Craig, M.H., Snow, R.W. & Sueur, D. 1999. A climate-based distribution model of malaria transmission in sub-saharan Africa. *Parasit. Today* 15(3): 105–111.
- Curriero, F.C., Patz, J.A., Rose, J.B. & Lele, S. 2001. The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948–1994. *Amer. J. Pub. Health* 91: 1194–1199.
- Dai, Y.J., Zeng, X.B., Dickinson, R.E., Baker, I., Bonan, G.B., Bosilovich, M.G., Denning, A.S., Dirmeyer, P.A., Houser, P.R., Niu, G., Oleson, K.W., Schlosser, C.A. & Yang, Z.L. 2003. The common land model. *Bull. Amer. Meteorol. Soc.* 84(8): 1013–1023.
- de la Beaujardiere, J. 2006 OpenGIS® web map server implementation specification. OGC Inc, Available from: <http://www.opengeospatial.org/> [Accessed 19th January 2012].
- Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S.J. & Ropelewski, C.F. 2007 Validation and inter-comparison of satellite rainfall products over East African complex topography. *Int. J. Rem. Sens.* 28: 1503 1526. doi:10.1080/01431160600954688.
- Dister, S.W., Beck, L.R., Wood, B.L., Falco, R. & Fish, D. 1993. The use of GIS and remote sensing technologies in a landscape approach to the study of Lyme disease transmission risk. In: *Proceedings of GIS '93: Geographic information systems in forestry, environmental and natural resource management*. Vancouver, B.C., Canada.
- Dister, S.W., Fish, D., Bros, S., Frank, D.H. & Wood, B.L. 1997. Landscape characterization of peridomestic risk for Lyme disease using satellite imagery. *Amer. J. Trop. Med. & Hyg.* 57: 687–92.
- Dousset, B., Gourmelon, F., Laaidi, K., Zeghnoun, A., Giraudet, E., Bretin, P. & Vandentorren, S. 2009. Satellite monitoring of summer time heat waves in the Paris metropolitan area. 7th International Conference on Urban Climate, Yokohama, Japan.

- Drummond, J.R. & Mand, G.S. 1996. The measurements of pollution in the troposphere (MOPITT) instrument: Overall performance and calibration requirements. *J. Atmos. & Oceanic Tech.* 13: 314–320.
- EASTWeb. 2012. Available from: <http://globalmonitoring.sdsstate.edu/projects/eastweb/index.php> [Accessed 19th January 2012].
- Ebi, K.L. & Meehl, G.A. 2007. The heat is on: Climate change & heatwaves in the Midwest. Excerpted from *Regional impacts of climate change: Four case studies in the United States*. Arlington, VA: Pew Center on Global Climate Change.
- Ebi, K.L. & Semenza, J.C. 2008. Community-Based Adaptation to the Health Impacts of Climate Change. *Amer. J. Prevent. Med.* 35(5): 501–507.
- ECMWF. 2007. Available from: [http://www.ecmwf.int/publications/newsletters/pdf/110\\_rev.pdf](http://www.ecmwf.int/publications/newsletters/pdf/110_rev.pdf) [Accessed 22nd March 2012].
- ECOCAST. Available from: <http://ecocast.arc.nasa.gov/topwp/> [Accessed 19th January 2012].
- Ek, M.B., Mitchell, K.E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G. & Tarpley, J.D. 2003. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res. Atmos.* 108(D22). doi:10.1029/2002JD003296.
- Ellis, E.C. & Ramankutty, N. 2008. Putting people in the map: Anthropogenic biomes of the world. *Front. Ecol. & Environ.* 6(8): 439–447.
- English, P.B., Sinclair, A.H., Ross, Z., Anderson, H., Boothe, V., Davis, C., Ebi, K., Kagey, B., Malecki, K., Shultz, R. & Simms, E. 2009. Environmental Health Indicators of Climate Change for the United States: Findings from the State Environmental Health Indicator Collaborative. *Environ. Health Perspect.* 117(11): 1673–1681.
- ESA. 2011 Available from: [http://www.esa.int/esaMI/Space\\_for\\_health/SEMNVMB474F\\_0.html](http://www.esa.int/esaMI/Space_for_health/SEMNVMB474F_0.html) [Accessed 18th January 2012].
- Evans, J. & Geerken R. 2006. Classifying rangeland vegetation type and coverage using a Fourier component based similarity measure. *Rem. Sens. of Env.* 105(1): 1–8.
- Fairlie, T.D., Jacob, D.J. & Park, R.J. 2007. The impact of transpacific transport of mineral dust in the United States. *Atmos. Environ.* 41: 1251–1266.
- Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D. & Alsdorf, D. 2007. The shuttle radar topography mission. *Rev. Geophys.* 45: RG2004. doi:10.1029/2005RG000183.
- Fayer, R., Dubey, J.P. & Lindsay, D.S. 2004. Zoonotic protozoa: From land to sea. *Trends in Parasitol.* 20: 531–536.
- Felzer, B., Kicklighter, D., Melillo, J., Wang, C., Zhuang, Q. & Prinn, R. 2004. Effects of ozone on net primary production and carbon sequestration in the conterminous United States using a biogeochemistry model. *Tellus B* 56(3): 230–248.
- Fielding, R. 2000. Architectural styles and the design of network-based software architectures. Ph.D. Dissertation. *Information and computer science* Univ. California, Irvine: Irvine, CA.
- Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Leach, P. & Berners-Lee, T. 1999. Hypertext Transfer Protocol HTTP/1.1 (RFC 2616). Available from: <http://www.rfc-editor.org/rfc/rfc2068.txt> [Accessed 10th April 2012].
- Finkelman, B.S., Viboud, C., Koelle, K., Ferrari, M.J., Bharti, N. & Grenfell, B.T. 2007. Global patterns in seasonal activity of influenza A/H3N2, A/H1N1, and B from 1997 to 2005: Viral coexistence and latitudinal gradients. *PLoS One* 2(12): e1296.
- Fischer, H., Birk, M., Blom, C., Carli, B., Carlotti, M., von Clarmann, T., Delbouille, L., Dudhia, A., Ehhalt, D., Endemann, M., Flaud, J.M., Gessner, R., Kleinert, A., Koopmann, R., Langen, J., Lopez-Puertas, M., Mosner, P., Nett, H., Oelhaf, H., Perron, G., Remedios, J., Ridolfi, M., Stiller, G. & Zander, R. 2008. MIPAS: An instrument for atmospheric and climate research. *Atmos. Chem. Phys.* 8: 2151–2188.
- Fishman, J., Creilson, J.K., Parker, P.A., Ainsworth, E.A., Vining, G.G., Szarka, J., Booker, F.L. & Xu, X. 2010. An investigation of widespread ozone damage to the soybean crop in the upper Midwest determined from ground-based and satellite measurements. *Atmos. Environ.* 44: 2248–2256.
- Fox, R.M. 1957. *Anopheles gambiae* in relation to malaria and filariasis in coastal Liberia. *Amer. J. Trop. Med. & Hyg.* 6: 598–620.
- Fraser, R.S. 1976. Satellite measurement of mass of Sahara dust in the atmosphere. *Appl. Opt.* 15: 2471–2479.
- Fraser, R.S., Kaufman, Y.J. & Mahoney, R.L. 1984. Satellite measurements of aerosol mass and transport. *Atmos. Environ.* 18: 2577–2584. doi:10.1016/0004-6981(84)90322-6.
- Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A. & Huang, X. 2010. MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Rem. Sens. Environ.* 114(1): 168–182.

- Geerten, R., Zaitchik, B. & Evans, P. 2005. Classifying rangeland vegetation type and coverage from NDVI time series using Fourier Filtered Cycle Similarity. *Int. J. Rem. Sens.* 26(24): 5535–5554.
- GEO. 2011a. Available from: [http://www.earthobservations.org/about\\_geo.shtml](http://www.earthobservations.org/about_geo.shtml) [Accessed 18th January 2012].
- GEO. 2011b. Available from: [http://www.earthobservations.org/documents/work%20plan/GEO%202012-2015%20Work%20Plan\\_Rev1.pdf](http://www.earthobservations.org/documents/work%20plan/GEO%202012-2015%20Work%20Plan_Rev1.pdf) [Accessed 18th January 2012].
- Gilles, H.M. & Warrell, D.A. (eds.). 1993. *Bruce-Chwatt's Essential Malariaology*. London: Arnold Publishing.
- Glantz, M. & Jamieson, D. 2000. Societal response to Hurricane Mitch and intra- versus intergenerational equity issues: Whose norms should apply? *Risk Anal.* 20: 869–882.
- Golden, M.L., Downs, R.R. & Davis-Packard, K. 2005. Confidentiality issues and policies related to the utilization and dissemination of geospatial data for public health applications. Available from: [http://www.ciesin.columbia.edu/pdf/SEDAC\\_ConfidentialityReport.pdf](http://www.ciesin.columbia.edu/pdf/SEDAC_ConfidentialityReport.pdf) [Accessed 18th January 2012].
- Golden, M.L., Yetman, G. & Chai-Onn, T. 2008. Assessment of populations in proximity to Superfund National Priorities List sites. Available from: <http://sedac.ciesin.columbia.edu/eh/sfpop.html> [Accessed 19th January 2012].
- Gomez-Elipe, A., Otero, A., van Herp, M. & Aguirre-Jaime, A. 2007. Forecasting malaria incidence based on monthly case reports and environmental factors in Karuzi, Burundi, 1997–2003. *Malaria J.* 6: 129. doi:10.1186/1475-2875-6-129.
- Goto, K., Nmor, J.C., Kurahashi, R., Minematsu, K., Yoda, T., Rakue, Y., Mizota, T. & Gotoh, K. 2010. Relationship between influx of yellow dust and bronchial asthma mortality using satellite data. *Scient. Res. & Essays* 5(24): 4044–4052.
- Grover-Kopec, E., Kawano, M., Klaver, R.W., Blumenthal, B., Ceccato, P. & Connor, S.J. 2005. An online operational rainfall-monitoring resource for epidemic malaria early warning systems in Africa. *Malaria J.* 4: 1–6.
- Hales, S., de Wet, N., Maindonald, J. & Woodward, A. 2002. Potential effect of population and climate changes on global distribution of dengue fever: An empirical model. *Lancet* 360: 830–834.
- Haque, U., Hashizume, M., Glass, G.E., Dewan, A.M., Dewan, A.M., Overgaard, H.J. & Yamamoto, T. 2010. The role of climate variability in the spread of malaria in Bangladeshi highlands. *PLoS ONE* 5(12): e14341. doi:10.1371/journal.pone.0014341.
- Harrington, L.C., Edman, J.D. & Scott, T.W. 2001. Why do female *Aedes aegypti* (Diptera: Culicidae) feed preferentially and frequently on human blood? *J. Med. Entomol.* 38(3): 411–422.
- Hay, S.I., Packer, M.J. & Rogers, D.J. 1997. The impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease. *Int. J. Rem. Sens.* 18(14): 2899–2930.
- Hay, S.I., Okiro, E.A., Gething, P.W., Patil, A.P., Tatem, A.J., Guerra, C.A. & Snow, R.W. 2010. Estimating the global clinical burden of *Plasmodium falciparum* malaria in 2007. *PLoS Med* 7(6): e1000290.
- Herman, J., Bhartia, P., Torres, O., Hsu, C., Seftor, C. & Celarier, E. 1997. Global distribution of UV-absorbing aerosols from Nimbus-7/TOMS data. *J. Geophys. Res.* 102: 16,911–16,922.
- Hertstein, U., Grunhage, L. & Jager, H.J. 1995. Assessment of past, present and future impacts of ozone and carbon dioxide on crop yields. *Atmos. Environ.* 29: 231–239.
- Higurashi, A. & Nakajima, T. 2002. Detection of aerosol types over the East China Sea near Japan from four-channel satellite data. *Geophys. Res. Lett.*, 29(17): 1836. doi:10.1029/2002GL015357.
- Holzer-Popp, T., Schroeder-Homscheidt, M., Breitkreuz, H., Martynenko, D. & Kluser, L. 2008. Improvements of synergistic aerosol retrieval for ENVISAT. *Atmos. Chem. Phys.* 8: 7651–7672.
- Hoshen, M.B. & Morse, A.P. 2004. A weather-driven model of malaria transmission. *Malaria J.* 3: 32. doi:10.1186/1475-2875-2875.
- Hu, R.M., Sokhi, R.S. & Fisher, B.E.A. 2009. New algorithms and their application for satellite remote sensing of surface PM2.5 and aerosol absorption. *J. Aerosol Sci.* 40: 394–402.
- Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X. & Ferreira, L.G. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Rem. Sens. Environ.* 83: 195–213.
- Huffman, G.J., Adler, R.F., Morrissey, M., Bolvin, D.T., Curtis, S., Joyce, R., McGavock, B. & Susskind, J. 2001. Global precipitation at one-degree daily resolution from multi-satellite observations. *J. Hydrometeorol.* 2(1): 36–50.
- Ichii, K., Maruyama, M. & Yamaguchi, Y. 2003. Multi-temporal analysis of deforestation in Rondonia state in Brazil using Landsat MSS, TM, ETM+ and NOAA AVHRR imagery and its relationship to changes in the local hydrological environment. *Int. J. Rem. Sens.* 24(22): 4467–4479.
- Igarashi, T. 2010. JAXA's concept on space initiatives for health. GEO Health and Environment Community Practice Workshop. Centre National d'Études Spatiales (CNES) Paris, France 27–28 July. Available from: [http://www.earthobservations.org/documents/cop/he\\_henv/20100727\\_France/16\\_JAXA.pdf](http://www.earthobservations.org/documents/cop/he_henv/20100727_France/16_JAXA.pdf) [Accessed 18th January 2012].

- IPCC. 2007. Impacts, adaptation and vulnerability. Working Group II, Fourth Assessment Report. Cambridge: Cambridge UP.
- Ito, K., De Leon, S.F. & Lippman, M. 2005. Associations between ozone and daily mortality: Analysis and meta-analysis. *Epidemiol.* 16: 446–457.
- Jerrett, M., Burnett, R.T., Ma, R., Pope III, C.A., Krewski, D., Newbold, K.B., Thurston, G., Shi, Y., Finkelstein, N., Calle, E.E. & Thun, M.J. 2005. Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiol.* 16(6) 727–736.
- Jerrett, M., Burnett, RT, Pope, C.A.-III, Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E., Thun, M. 2009. Long-term ozone exposure and mortality. *New Engl. J. Med.* 360(26): 1085–1095.
- Jones, K.E., Patel, N.G., Levy, M.A., Storeygard, A., Balk, D., Gittleman, J.L. & Daszak, P. 2008. Global trends in emerging infectious diseases. *Nature* 451(7181): 990–993.
- Joyce, R.J., Janowiak, J.E., Arkin, P.A. & Xie, P. 2004. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydromet.* 5: 487–503.
- Kahru, M. 1997. Using satellites to monitor large-scale environmental change in the Baltic Sea. In M. Kahru & C.W. Brown (eds.), *Monitoring algal blooms: New techniques for detecting large-scale environmental change*: 43–61. Berlin: Springer-Verlag.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R. & Joseph, D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.* 77(3): 437–471.
- Kaufman, Y.J., Tanré, D. & Boucher, O. 2002. A satellite view of aerosols in the climate system. *Nature* 419: 215–223.
- Kay, B.H., Fanning, I.D., Mottram, P. 1989. Rearing temperature influences flavivirus vector competence of mosquitoes. *Med. Vet. Entomol.* 3: 415–422.
- Kelly-Hope, L.A., Hemingway J. & McKenzie, F.E. 2009. Environmental factors associated with the malaria vectors *Anopheles gambiae* and *Anopheles funestus* in Kenya. *Malaria J.* 8: 268.
- Kiang, R.K., Adimi, F. & Soika, V. 2006. Meteorological, environmental remote sensing and neural network analysis of the epidemiology of malaria transmission in Thailand. *Geospat Health* 1: 71–84.
- Kiang, R.K., Adimi, F. & Soebiyanto, R.P. 2011. Remote sensing-based modelling of infectious disease transmission. In T. Kass-Hout & X. Zhang (eds.), *Biosurveillance: Methods and case studies*. Boca Raton: CRC Press.
- Knabb, R.D., Rhome, J.R & Brown, D.P. 2006. Tropical cyclone report: Hurricane Katrina: 23–30 August 2005. Available from: [http://www.nhc.noaa.gov/pdf/TCR-AL122005\\_Katrina.pdf](http://www.nhc.noaa.gov/pdf/TCR-AL122005_Katrina.pdf) [Accessed 19th January 2012.]
- Koster, R.D., Suarez, M.J., Ducharne, A., Stieglitz, M. & Kumar, P. 2000. A catchment-based approach to modeling land surface processes in a general circulation model 1. Model structure. *J. Geophys. Res. Atmos.* 105(D20): 24,809–24,822.
- Kovats, R.S. & Hajat, S. 2008. Heat stress and public health: A critical review. *Annu. Rev. Pub. Health* 29: 41–55.
- Kovats, R.S., Bouma, M.J., Hajat, S., Worrall, E. & Haines, A. 2003. El Niño and health. *Lancet* 362: 1481–1489.
- Kowalczyk, E.Z., Wang, Y.P., Law, R.M., Davies, H.L., McGregor, J.L. & Abramowitz, G. 2006. The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model. *CSIRO Marine & Atmos. Res. Paper* 013.
- Krueger, A.J. & Jaross, G. 1999. TOMS ADEOS instrument characterization. *IEEE Trans. Geosci. & Rem. Sens.* 37: 1543–1549.
- Kutser, T. 2004. Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. *Limnol. & Oceanogr.* 49: 2179–2189.
- LBL. 2011. Available from: <http://www.lbl.gov/Education/ELSI/Frames/pollution-health-effects-f.html> [Accessed 18th January 2012].
- Leptoukh, G., Csizsar, I., Romanov, P., Shen, S., Loboda, T. & Gerasimov, I. 2007. NASA NEESPI data center for satellite remote sensing data and services. *Environ. Res. Lett.*, 2. doi:10.1088/1748-9326/2/4/045009.
- Liang, X., Wood, E.F. & Lettenmaier, P. 1996. Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global & Planetary Change* 13(1–4): 195–206.
- Linthicum, K.J., Anyamba, A., Tucker, C.J., Kelley, P.W., Myers, M.F. & Peters, C.J., 1999. Climate and satellite indicators to forecast Rift Valley Fever epidemics in Kenya. *Science* 285: 397–400.

- Liu, Y., Sarnat, J.A., Coull, B.A., Koutrakis, P. & Jacob, D.J. 2004. Validation of Multiangle Imaging Spectroradiometer (MISR) aerosol optical thickness measurements using Aerosol Robotic Network (AERONET) observations over the contiguous United States. *J. Geophys. Res.*, 109, D06205. doi:10.1029/2003JD003981
- Liu, Y., Paciorek, C.J. & Koutrakis, P. 2009. Estimating regional spatial and temporal variability of PM<sub>2.5</sub> concentrations using satellite data, meteorology, and land use information. *Environ. Health Perspec.* 117(6): 886–892.
- Lynnes, C., Strub, R., Seiler, E., Joshi, T. & MacHarrie, P. 2009. Mirador: A simple, fast search interface for global remote sensing data sets. *IEEE Trans. Geosci. Rem. Sens.* 47(1): 92–96.
- Lynnes, C., Olsen, E., Fox, P., Vollmer, B., Wolfe, R.E. & Samadi, S. 2010. A quality screening service for remote sensing data In: H. Salim & K. Keahay (general chairs). Proc. 19th ACM. *International symposium on high performance distributed computing (HPDC)*: 554–559. New York: ACM.
- Lyons, W.A., Dooley, J.C. Jr. & Whitby, K.T. 1978. Satellite detection of long-range pollution transport and sulfate aerosol hazes. *Atmos. Environ.* 12: 621–631. doi:10.1016/0004-6981(78)90242-1.
- Matricardi, E.A.T., Skole, D.L., Cochrane, M.A., Qi, J. & Chomentowski, W. 2005. Monitoring selective logging in tropical evergreen forests using Landsat: Multitemporal regional analyses in Mato Grosso, Brazil. *Earth Interactions* 9(24): 1–24. doi:10.1175/EI142.1.
- McGranahan, G., Balk, D. & Anderson, B. 2007. The rising tide: Assessing the risks of climate change and human settlements in low elevation coastal zones. *Environ. & Urbaniza.* 19(1): 17–37.
- Meehl, G.A. & Tebaldi, C. 2004 More intense, more frequent and longer lasting heat waves in the 21st Century. *Science*. 305(5686): 994–997.
- Meng, C.J., Li, Z.L., Zhan, X., Xhi, J.C. & Liu, C.Y. 2009. Land surface temperature data assimilation and its impact on evapotranspiration estimates from the Common Land Model. *Water Res. Res.* 45(W02421). doi:10.1029/2008WR006971.
- Mitchell, K.E., Lohmann, D., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Cosgrove, B.A., Sheffield, J., Duan, Q., Luo, L., Higgins, R.W., Pinker, R.T., Tarpley, J.D., Lettenmaier, D.P., Marshall, C.H., Entin, J.K., Pan, M., Shi, W., Koren, V., Meng, J., Ramsay, B.H. & Bailey, A.A. 2004. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modelling system. *J. Geophys. Res.* 109: D07S90. doi:10.1029/2003JD003823.
- Morain, S.A. & Budge, A.M. 2010. Suggested practices for forecasting dust storms and intervening their health effects. In O. Altan, R. Backhaus, P. Boccardo & S. Zlatanova (eds.), *Geoinformation for disaster and risk management*: 45–50. Copenhagen: Joint Board Geospatial Information Sciences & United Nation Office of Outer Space Affairs.
- NASA. 2012. Available from: <http://appliedsciences.nasa.gov/health-air.html> [Accessed 24th January 2012].
- Nemani, R., Hashimoto, H., Votava, P., Melton, F., Wang, W., Michaelis, A., Mutch, L., Milesi, C., Hiatt, S. & White, M. 2009. Monitoring and forecasting ecosystem dynamics using the Terrestrial Observation and Prediction Systems (TOPS). *Rem. Sens. Environ.* 113: 1497–1509.
- Nickovic, S., Kallos, G., Papadopoulos, A. & Kakaliagou, O. 2001. A model for prediction of desert dust cycle in the atmosphere. *J. Geophys. Res.* 106(D16): 18,113–18,130.
- Njoku, E.G., Jackson, T.J., Lakshmi, V., Chan, T.K. & Nghiem, S.V. 2003. Soil moisture retrieval from AMSR-E. *IEEE Trans. Geosci. & Rem. Sens.* 41: 215–229.
- Noor, A.M., Mutheu, J.J., Tatem, A.J., Hay, S.I. & Snow, R.W. 2009. Insecticide-treated net coverage in Africa: Mapping progress in 2000–07. *Lancet* 373(9657): 58–67.
- NRC. 2001. Under the weather: Climate, ecosystems and infectious disease: Washington DC: National Academy Press.
- NWS. 2012. Available from: <http://www.nws.noaa.gov/os/heat/index.shtml> [Accessed 22nd March 2012].
- Oliveira, P.J., Asner, G.P., Knapp, D.E., Almeyda, A., Galvan-Gildemeister, R., Keene, S., Raybin, R.F. & Smith, R.C. 2007. Land-use allocation protects the Peruvian Amazon. *Science* 317(5842): 1233–1236.
- O'Neill, M.S., Zanobetti, A. & Schwartz, J. 2003. Modifiers of the temperature and mortality association in seven US cities. *Amer. J. Epidemiol.* 157(12): 1074–1082.
- Pampana, E. 1969. *A textbook of malaria eradication*. London: Oxford UP.
- Patz, J. 2005. Satellite remote sensing can improve chances of achieving sustainable health. *Environ. Health Perspect.* 113(2): 84–85.
- Patz, J.A., Daszak, P., Tabor, G.M., Aguirre, A.A., Pearl, M., Epstein, J., Wolfe, N.D., Kilpatrick, A.M., Foufopoulos, J., Molyneux, D. & Bradley, D.J. 2004. Unhealthy landscapes: Policy recommendations on land use change and infectious disease emergence. *Environ. Health Perspect.* 112: 1092–1098.

- Pope, C.A. III, Burnett, R.T., Thun, M.J., Calle, E.E., Krewski, D., Ito, K. & Thurston, G.D. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *J. Amer. Med. Assoc.* 287: 1132–1141.
- Prospero, J.M., Ginoux, P., Torres, O., Nicholson, S. & Gill, T. 2002. Environmental characterization of global sources of atmospheric soil dust identified with the NIMBUS 7 Total Ozone Mapping Spectrometer (TOMS) absorbing aerosol product. *Rev. of Geophys.* 41(1): 1–31.
- Rahman, A., Kogan, F. & Roytman, L. 2006. Analysis of malaria cases in Bangladesh with remote sensing data. *Amer. J. Trop. Med. & Hyg.* 74(1): 17–19.
- Raudenbush, S.W. & Bryk, A.S. 2002. *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks: Sage.
- Rinner, C., Patychuk, D., Bassil, K., Nasr, S., Gower, S. & Campbell, M. 2010. The role of maps in neighbourhood-level heat vulnerability assessment for the City of Toronto. *Cart. & GIS* 37(1): 31–44.
- Robine, J., Cheung, S., Le Roy, S., Van Oyen, H., Griffiths, C., Michel, J.P. & Herrmann, F. 2008. Death toll exceeded 70,000 in Europe during the summer of 2003. *C.R. Biol.* 331: U171–U175.
- Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J.K., Walker, J.P., Lohmann, D. & Toll, D. 2004. The global land data assimilation system. *Bull. Amer. Meteor. Soc.* 85(3): 381–394.
- Roy, D.P., Ju, J., Kline, K., Scaramuzza, P.L., Kovalsky, V. & Hansen, M. 2010. Web-enabled Landsat data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Rem. Sens. Environ.* 114(1): 35–49.
- Sanchez, G.M. 2007. The application and assimilation of Shuttle Radar Topography Mission Version 1 data for high resolution dust modelling. MA Thesis, University of New Mexico, Department of Geography.
- Schifano, P., Cappai, G., DeSario, M., Michelozi, P., Marino, C., Bargagli, A.M. & Perucci, C.A. 2009. Susceptibility to heat wave-related mortality: A follow-up study of a cohort of elderly in Rome. *Environ. Health* 8:50. doi:10.1186/1476-069X-8-50.
- Schlussel, P., Hultberg, T.H., Phillips, P.L., August, T. & Calbet, X. 2005. The operational IASI level 2 processor. In L. Burrows & J.P. Eichmann (eds.), *Atmospheric remote sensing: Earth's surface, troposphere, stratosphere and mesosphere I*: 982–988. Orlando: Elsevier.
- Schwartz, J. 2005. Who is sensitive to extremes of temperature? A case-only analysis. *Epidemiol.* 16(1): 67–72.
- SEDAC. 2007. National aggregates of geospatial data collection: Population, landscape and climate estimates (PLACE) Version II. Available from: <http://sedac.ciesin.columbia.edu/place/methods.jsp> [Accessed 19th January 2012].
- Semenza, J.C., Rubin, C.H., Falter, K.H., Selanikio, J.D., Flanders, W.D., Howe, H.L. & Wilhelm, J.L. 1996. Heat-related deaths during the July 1995 heat wave in Chicago. *N. Engl. J. Med.* 335: 84–90.
- Shaman, J., Pitzer, V.E., Viboud, C., Grenfell, B.T. & Lipsitch, M. 2010. Absolute humidity and the seasonal onset of influenza in the continental United States. *PLoS Biol* 8(2): e1000316.
- Shryock, H.S., Siegel, J.S. & Stockwell, E.G. 1976. *The methods and materials of demography*. New York: Academic Press.
- Soebiyanto, R.P., Adimi, F. & Kiang, R.K. 2010 Modelling and predicting seasonal influenza transmission in warm regions using climatological parameters. *PLoS ONE* 5(3): e9450.
- Spinhirne, J.D., Palm, S.P., Hart, W.D., Hlavka, D.L. & Welton, E.J. 2005. Cloud and aerosol measurements from the GLAS space borne lidar: Initial results. *Geophys. Res. Lett.* 32: L22S03. doi:10.1029/2005GL023507.
- Steininger, M.K., Tucker, C.J., Townshend, J.R.G., Killeen, T.J., Desch, A., Bell, V. & Ersts, P. 2001. Tropical deforestation in the Bolivian Amazon. *Environ. Conserv.* 28(2): 127–134.
- Storeygard, A., Balk, D., Levy, M. & Deane, G. 2008. The global distribution of infant mortality: A subnational spatial view. *Popul. Space & Place* 14(3): 209–229.
- Tanré, D., Bréon, F.M., Deuzé, J.L., Herman, M., Goloub, P., Nadal, F. & Marchand, A. 2001. Global observation of anthropogenic aerosols from satellite. *Geophys. Res. Lett.* 28(24): 4555–4558. doi:10.1029/2001GL013036.
- Tatem, A.J., Smith, D.L., Gething, P.W., Kabaria, C.W., Snow, R.W. & Hay, S.I. 2010. Ranking of elimination feasibility between malaria-endemic countries. *Lancet* 376(9752): 1579–1591.
- Teklehaimanot, A., McCord, G.C. & Sachs, J.D. 2007. Scaling up malaria control in Africa: An economic and epidemiological assessment. *Amer. J. Trop. Med. & Hyg.* 77(Supp6): 138–144.
- Thompson, D.F., Malone, J.B., Harb, M., Faris, R., Huh, O.K., Buck, A.A. & Cline, B.L. 1996. Bancroftian filariasis distribution and diurnal temperature differences in the southern Nile delta. *Emerg. Infect. Dis.* 2: 234–235.

- Thomson, M.C., Doblas-Reyes, F.J., Mason, S.J., Hagedorn, R., Connor, S.J., Phindela, T., Morse, A.P. & Palmer, T.N. 2006. Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. *Nature* 439: 576–579.
- Tong, D.Q. & Mauzerall, D.L. 2008. Summertime state-level source-receptor relationships between NO<sub>x</sub> emissions and downwind surface ozone concentrations over the continental United States. *Environ. Sci. & Tech.* 42(21): 7976–7984.
- Tong, D., Mathur, R., Schere, K., Kang D. & Yu, S. 2007. The use of air quality forecasts to assess impacts of air pollution on crops: Methodology and case study. *Atmos. Environ.* 41(38): 8772–8794.
- Tong, D.Q., Kan, H. & Yu, S. 2009. Ozone exposure and mortality. *New Engl. J. Med.* 360(26): 2787–2787.
- Tucker, C.J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Rem. Sens. Environ.* 8: 127–150.
- Uppala, S., Dee, D., Kobayashi, S., Berrisford, P. & Simmons, A. 2008. Towards a climate data assimilation system: Status update of ERA-Interim. *ECMWF Newslett.* 115: 12–18.
- Vandentorren, S., Bretin, P., Zeghnoun, A., Mandereau-Bruno, L., Croisier, A. & Cochet, C. 2006. August 2003. Heat wave in France: Risk factors for death of elderly people living at home. *Eur. J. Pub. Health* 16: 583–591.
- Van de Water, P., Main, C.E., Keever, T. & Levetin, E. 2003. An assessment of predictive forecasting *Juniperus ashei* pollen movement in the southern Great Plains. *Int. J. Biomet.* 48: 74–82.
- van den Hurk, B.J.M.M., Viterbo, P., Beljaars, A.C.M. & Betts, A.K. 2000. Offline validation of the ERA40 surface scheme. *ECMWF Tech. Memo.* 295: 1–42.
- van Donkelaar, A., Martin, R.V., Brauer, M., Kahn, R., Levy, R., Verduzco, C. & Villeneuve, P.J. 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. *Environ. Health Perspect.* 118(6): 847–855.
- Vittor, A.Y., Pan, W.K., Gilman, R.H., Tielsch, J., Glass, G., Shields, T., Sanchez-Lozano, W., Pinedo, V.V., Salas-Cobos, E., Flores, S. & Patz, J.A. 2009. Linking deforestation to malaria in the Amazon: Characterization of the breeding habitat of the principal malaria vector, *Anopheles darlingi*. *Amer. J. Trop. Med. & Hyg.* 81(1): 5–12.
- Whiteside, A. & Evans, J.D. (eds.). 2006. Web Coverage Service (WCS) Implementation Specification, Version 1.1.0. *Open Geospat. Consort.* 06-083r8: 129.
- Wilhelmi, O.V., Purvis, K.L. & Harriss, R.C. 2004. Designing a geospatial information infrastructure for mitigation of heat wave hazards in urban areas. *Nat. Haz. Rev.* 5: 147–158.
- Witt, C.J., Richards, A.L., Masuoka, P.M., Foley, D.H., Buczak, A.L., Musila, L.A., Richardson, J.H., Colacicco-Mayhugh, M.G., Rueda, L.M., Klein, T.A., Anyamba, A., Small, J., Pavlin, J.A., Fukuda, M.M., Gaydos, J. & Russell, K.L. 2011. The AFHSC-Division of GEIS operations predictive surveillance program: A multidisciplinary approach for the early detection and response to disease outbreaks. *BMC Pub. Health (Suppl2)*: S10.
- WHO 2009. Available from: <http://www.ciesin.columbia.edu/docs/001-007/001-007.html> [Accessed 18th January 2012].
- WHO 2012. HealthMapper. Available from: [http://www.who.int/health\\_mapping/tools/healthmapper/en/](http://www.who.int/health_mapping/tools/healthmapper/en/) [Accessed 23rd March 2012].
- Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Warner, R.A., Tester, P.A., Dyble, J. & Fahnenstiel, G.L. 2008. Relating spectral shape to cyanobacterial blooms in the Laurentian Great Lakes. *Int. J. of Rem. Sens.* 29: 3665–3672.
- Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Schwab, D.J., Watabayashi, G.Y. & Christensen, J.D. 2011. Estimating cyanobacterial bloom transport by coupling remotely sensed imagery and a hydrodynamic model. *Ecol. Appl.* 21: 2709–2721. doi:<http://dx.doi.org/10.1890/10-1454.1>.
- Xiao X, Gilbert, M., Slingenbergh, J., Lei, F. & Boles, S. 2007. Remote sensing, ecological variables, and wild bird migration related to outbreaks of highly pathogenic H5N1 avian influenza. *J. Wildife. Dis.* 43: 540–546.
- YCELP, CIESIN, World Economic Forum & the European Commission, Joint Research Centre. 2010. Available from: [http://www.ciesin.columbia.edu/repository/epi/data/EPI\\_2010\\_report.pdf](http://www.ciesin.columbia.edu/repository/epi/data/EPI_2010_report.pdf) [Accessed 18th January 2012].
- Yu, H., Remer, L.A., Chin, M., Bian, H., Kleidman, R.G. & Diehl, T. 2008. A satellite-based assessment of transPacific transport of pollution aerosol. *J. Geophys. Res.* 113: D14S12. doi:10.1029/2007JD009349.
- Zaitchik, B.F., Rodell, M., Reichle, R. 2008. Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River basin. *J. Hydrometeorol.* 9(3): 535–548.
- Zhang, H., Hoff, R.M. & Engel-Cox, J.A. 2009. The relation between moderate resolution imaging spectroradiometer (MODIS) aerosol optical depth and PM<sub>2.5</sub> over the United States: A geographical comparison by US Environmental Protection Agency regions. *J. Air & Waste Manag. Assoc.* 59: 358–1369.