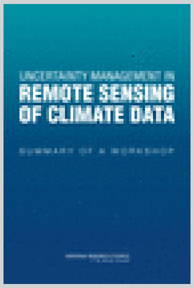

Uncertainty Management in Remote Sensing of Climate Data: Summary of a Workshop

ISBN
978-0-309-13958-8

64 pages
6 x 9
PAPERBACK (2009)

Martha McConnell and Scott Weidman, Rapporteurs; National Research Council

UNCERTAINTY MANAGEMENT IN REMOTE SENSING OF CLIMATE DATA

SUMMARY OF A WORKSHOP

Martha McConnell and Scott Weidman, *Rapporteurs*

Board on Atmospheric Sciences and Climate
Climate Research Committee

Board on Mathematical Sciences and Their Applications
Committee on Applied and Theoretical Statistics

Space Studies Board
Committee on Earth Studies

Division on Earth and Life Studies

Division on Engineering and Physical Sciences

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

This study was supported by the National Aeronautics and Space Administration contract number NNH07CC37B for the Climate Research Committee and NNH06CE15B for the Committee on Earth Studies, and National Science Foundation contract number DMS-0456571 for the Committee on Applied and Theoretical Statistics. Any opinions expressed in this material are those of the author(s) and do not necessarily reflect the views of NASA, NSF, or any of its sub agencies.

International Standard Book Number-13: 978-0-309-13958-8
International Standard Book Number-10: 0-309-13958-9

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, http://www.nap.edu.

Printed in the United States of America

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

PLANNING COMMITTEE FOR A WORKSHOP ON UNCERTAINTY MANGEMENT IN REMOTE SENSING OF CLIMATE DATA

AMY BRAVERMAN (*Chair*), Jet Propulsion Laboratory, Pasadena, California
PHILIP E. ARDANUY, Raytheon Information Solutions, Reston, Virginia
JOHN J. BATES, National Oceanic and Atmospheric Administration, Asheville, North Carolina
JAMES A. COAKLEY, JR., Oregon State University, Corvallis
KAREN KAFADAR, Indiana University, Bloomington
DOUGLAS NYCHKA, National Center for Atmospheric Research, Boulder, Colorado
JOYCE E. PENNER, University of Michigan, Ann Arbor
STEVEN PLATNICK, National Aeronautics and Space Administration, Greenbelt, Maryland

NRC Staff

MARTHA MCCONNELL, Associate Program Officer
SCOTT WEIDMAN, Study Director
IAN KRAUCUNAS, Senior Program Officer
ART CHARO, Senior Program Officer
LAUREN BROWN, Christine Mirzayan Fellow
KATIE WELLER, Research Associate
SHELLY-ANN FREELAND, Senior Program Assistant

CLIMATE RESEARCH COMMITTEE

GERALD MEEHL (*Chair*), National Center for Atmospheric Research, Boulder, Colorado
ANA P. BARROS, Duke University, Durham, North Carolina
CECILIA BITZ, University of Washington, Seattle
JAMES COAKLEY JR., Oregon State University, Corvallis
GABRIELE HEGERL, University of Edinburgh, Scotland
HENRY D. JACOBY, Massachusetts Institute of Technology, Cambridge
ANTHONY C. JANETOS, Pacific Northwest National Laboratory/University of Maryland, College Park
ROBERT LEMPERT, RAND Corporation, Santa Monica, California
ROGER B. LUKAS, University of Hawaii, Honolulu
LINDA MEARNS, National Center for Atmospheric Research, Boulder, Colorado
SUMANT NIGAM, University of Maryland, College Park
JOYCE E. PENNER, University of Michigan, Ann Arbor
RICHARD RICHELS, Electric Power Research Institute, Washington, D.C.
TARO TAKAHASHI, Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York
LONNIE G. THOMPSON, Ohio State University, Columbus

COMMITTEE ON APPLIED AND THEORETICAL STATISTICS

KAREN KAFADAR (*Chair*), Indiana University, Bloomington
AMY BRAVERMAN, Jet Propulsion Laboratory, Pasadena, California
CONSTANTINE GATSONIS, Brown University, Providence, Rhode Island
MICHAEL F. GOODCHILD, University of California, Santa Barbara
MICHAEL A. NEWTON, University of Wisconsin, Madison
MICHAEL STEIN, University of Chicago, Illinois

COMMITTEE ON EARTH STUDIES

BERRIEN MOORE III (*Chair*), Climate Central, Princeton, New Jersey
RUTH S. DeFRIES *(Vice Chair)*, Columbia University, New York
MARK R. ABBOTT, Oregon State University, Corvallis
RICHARD A. ANTHES, University Corporation for Atmospheric Research, Boulder, Colorado
PHILIP E. ARDANUY, Raytheon Information Solutions, Reston, Virginia
STEVEN J. BATTEL, Battel Engineering, Scottsdale, Arizona
ANTONIO J. BUSALACCHI, JR., University of Maryland, College Park
HEIDI M. DIERSSEN, University of Connecticut, Storrs
THOMAS H. VONDER HAAR, Colorado State University, Fort Collins
HUNG-LUNG ALLEN HUANG, University of Wisconsin, Madison
ANNE W. NOLIN, Oregon State University, Corvallis
JAY S. PEARLMAN, Boeing Company, Kent, Washington

Preface

The National Academies' Climate Research Committee (CRC), organized under the Board on Atmospheric Sciences and Climate, is charged to foster atmospheric, oceanic, and related research aimed at advancing knowledge and understanding of climate and climate change. The Committee on Applied and Theoretical Statistics (CATS), which is organized under the Board on Mathematical Sciences and Their Applications, is charged to provide a locus of activity and concern for the statistical sciences, statistical education, use of statistics, and issues affecting the field. The Committee on Earth Studies (CES), organized under the Space Studies Board, provides oversight on civil earth observation space activities in the general areas of earth sciences and other remote sensing applications, including both applicable technology and all earth science disciplines that can be addressed from space. All three committees have expertise and a longstanding interest in the use of statistics in climate research and applications.

Satellites provide a unique vantage point for studying the earth's climate and associated systems, but obtaining climate-relevant data from remotely sensed platforms is a demanding task requiring careful analysis and expertise from numerous disciplines. Many of the techniques currently used to process and analyze remotely sensed climate data could potentially be improved using modern statistical techniques. This is particularly true because the amount of data involved has increased so dramatically. To address these issues, and stimulate additional opportunities for beneficial collaboration between statisticians, climate scientists, and experts in remote sensing, a workshop was convened December 4, 2008, in Washington, D.C., under the auspices of the CRC, CATS, and CES to explore uncertainty management in remote sensing, with an emphasis on remotely sensed climate information. This workshop brought together the statistics and geoscience communities from academia, government, and industry.

Through invited presentations and discussion, participants investigated the sources of uncertainty throughout satellite and other remote data collection systems, described the statistical methods currently used to quantify sources of uncertainty, and discussed how modern statistical methods might be used to provide a more useful framework for characterizing and propagating these uncertainties. The primary objectives were to examine sources of uncertainty in remote sensing data collection systems that include, among other things, issues of sampling, scale, processing, and validation. Other topics covered at the workshop included the challenge of communicating uncertainties to the end-user and building institutional capacity to address problems that require expertise from both statisticians and earth scientists.

Under the National Academies' policy, workshops do not produce findings and recommendations. Thus, the goal of this workshop report is to summarize the major discussion items that arose and to synthesize key points from the presentations on uncertainty management in remote sensing. This report was written by *rapporteurs* and we hope that the broad topics covered in this report will stimulate additional opportunities for collaborations betweens statisticians and earth scientists. We thank the planning team, the presenters, and other participants at the workshop.

Martha McConnell
Scott Weidman
Rapporteurs

Acknowledgments

This workshop report was written by National Academies staff based on the presentations and discussion at the workshop. It does not necessarily represent the views of the workshop planning committee, which was not involved in its production. This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's (NRC's) Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this summary:

Amy Braverman, Jet Propulsion Laboratory, Pasadena, California
John J. Bates, National Oceanic and Atmospheric Administration, Asheville, North Carolina
Anna Michalak, University of Michigan, Ann Arbor
Ivanka Stajner, Noblis, Inc., Falls Church, Virginia

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions and recommendations from the speakers nor did they see the final draft of the summary before its release. **Lee Branscombe**, Climatological Consulting Corporation, oversaw the review of this summary. Appointed by the NRC, he was responsible for making certain that an independent examination of this summary was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this summary rests entirely with the authors and the institution.

Contents

1

Introduction

Great advances have been made in our understanding of the climate system over the past few decades, and remotely sensed data have played a key role in supporting many of these advances. Improvements in satellites and in computational and data-handling techniques have yielded high quality, readily accessible data. However, rapid increases in data volume have also led to large and complex datasets that pose significant challenges in data analysis (NRC, 2007). Uncertainty characterization is needed for every satellite mission and scientists continue to be challenged by the need to reduce the uncertainty in remotely sensed climate records and projections. The approaches currently used to quantify the uncertainty in remotely sensed data, including statistical methods used to calibrate and validate satellite instruments, lack an overall mathematically based framework. An additional challenge is characterizing uncertainty in ways that are useful to a broad spectrum of end-users.

In December 2008, three standing committees of the National Academies held a workshop to survey how statisticians, climate scientists, and remote sensing experts might address the challenges of uncertainty management in remote sensing of climate data. The emphasis of the workshop was on raising and discussing issues that could be studied more intently by individual researchers or teams of researchers, and on setting the stage for possible future collaborative activities. Issues and questions that were addressed at the workshop include the following:

1. What methods are currently used to compare time series at single points in space with instantaneous but sparsely sampled area averages to "validate" remotely sensed climate data? Are there more sophisticated or advanced methods that could be applied to improve validation tools or uncertainty estimates? Are there alternative means of measuring the same phenomena to confirm the accuracy of satellite observations?

2. How can fairly short-term, spatially dense remote sensing observations inform climate models operating at long time scales and relatively coarse spatial resolutions? Are there remotely sensed data that could, through the use of modern statistical methods, be useful for improving climate models or informing other types of climate research?

3. What are the practical and institutional barriers (e.g., lack of qualified statisticians working in the field) to making progress on developing and improving statistical techniques for processing, validating, and analyzing remotely sensed climate data?

In her introductory remarks at the workshop, planning team chair Amy Braverman from the Jet Propulsion Laboratory presented Table 1-1 to illustrate how statistical methods (rows) can help address three major challenges in the use of remotely sensed climate data (the columns). The first of these three major challenges is the validation of remote sensing retrievals. When a remote sensing instrument retrieves a measurement that is used to infer a geophysical value (e.g., atmospheric temperature), uncertainties exist both in the measured values and in the statistical model used to validate the remotely sensed parameter. The second challenge is improving the representation of physical processes within all types of climate models. Workshop participants stressed the need to better represent physical processes within global earth system models, a critical component to projecting future climate accurately, reducing uncertainty, and ultimately aiding policy decisions. The third major challenge in climate research where statistics plays an important role is aggregating the observed and modeled knowledge, each with their associated uncertainties, to develop a better understanding of the climate system that can lead to useful predictions.

Complex and multifaceted relationships in the physics of the climate system contribute uncertainty over and above that which is normally present in making inferences from massive, spatio-temporal data. Isolating and quantifying these uncertainties in the face of multiple scales of spatial and temporal resolution, nonlinear relationships, feedbacks, and varying levels of a priori knowledge poses major challenges to achieving the linkages shown in Table 1-1. A formal statistical model that articulates relationships among both known and unknown quantities of interest and observations can sharpen the picture and make the problem

TABLE 1-1 Three Major Challenges in the Use of Remotely Sensed Climate Data (Columns) and Three Roles Played by Statistical Methods (Rows)

	Challenge: Validation of remote sensing retrievals	Challenge: Improving physical representations and understanding	Challenge: Extrapolating to future climate predictions
Role for statistics: Clarify and characterize sources of uncertainty in remote sensing data	Characterize spatio-temporal mismatches, retrieval algorithm differences; address sparseness or absence of ground truth	Develop new statistical methods to make the most of new data types to address new science questions	Maximize value of limited data and hard-to-formalize assumptions about relationships among past, present, and future
Role for statistics: Develop statistical methods to quantify and reduce uncertainty	Develop formal statistical error measures for both bias and variance	Develop new methods to exploit massive datasets in an inferential setting	Develop formalisms for combining output from different models in light of available data
Role for statistics: Provide an overarching framework	Overcome mismatches by statistical modeling of relationships between observed and unobserved quantities.	Pose problems as formal questions of statistical inference	Combine physical and statistical models

SOURCE: Table courtesy of Amy Braverman, JPL.

more tractable. Random variables can represent uncertain quantities and describe relationships through joint and conditional distributions. Random variables can also be infused into systems of physical equations, to carry information about uncertainties along with information provided by physical knowledge.

Crafting such hybrid physical-statistical models to capture the essence of our understanding is not easy. The climate system is inherently nonlinear and includes feedback loops where variables directly and indirectly affect one another. Figure 1-1, presented at the workshop by William Rossow from the City College of New York, is a simple diagram of the energy and water cycles of the climate system that demonstrates how the system is interconnected. In order to gain a true understanding of climate feedbacks it is important to understand multiple variables in the climate system and their interactions. Clouds and precipitation, for example, play a crucial role in both the water cycle and the earth's energy balance, affecting the sources and sinks of heat in the climate system. The release of latent heat during precipitation events provides energy that drives atmospheric circulations, and, in turn, atmospheric circulation processes that affect the distribution of water vapor and the formation of clouds have a pronounced effect on the transfer of radiation through the atmosphere. Therefore, analyzing the interrelationships between multiple variables in the climate system is key to understanding processes of interest.

Large volumes of remote sensing data are available to assist in refining models of physical systems like that shown in Figure 1-1. Data provide information about physical mechanisms at work in the atmosphere, and also about the uncertainties or gaps in our understanding of how those mechanisms operate. To make use of data in this way, however, requires that inherent uncertainties and biases in the data themselves be known and quantified. Therefore, the problem requires a holistic approach to uncertainty management beginning with data collection and validation strategies that are cognizant of the uses of the data. These challenges can be addressed in two ways:

1. By identifying data collection and analysis methods that minimize the uncertainties; and
2. By identifying the contributions to uncertainty at the various steps in collection and analysis, thereby pointing out the most promising targets for improvement.

Uncertainty quantification, in the broadest sense, is to account for not only uncertainty in individual parameters within the models that are used, but also to account for the uncertainty inherent in the actual models themselves, which are only approximate representations of physical processes. Workshop participants emphasized that improving physical process representation is critical for both improving climate models and for better characterizing their uncertainties. Statistics can contribute to solving this problem by moving beyond linear analysis for individual parameters to capture more complex relationships that have a physical

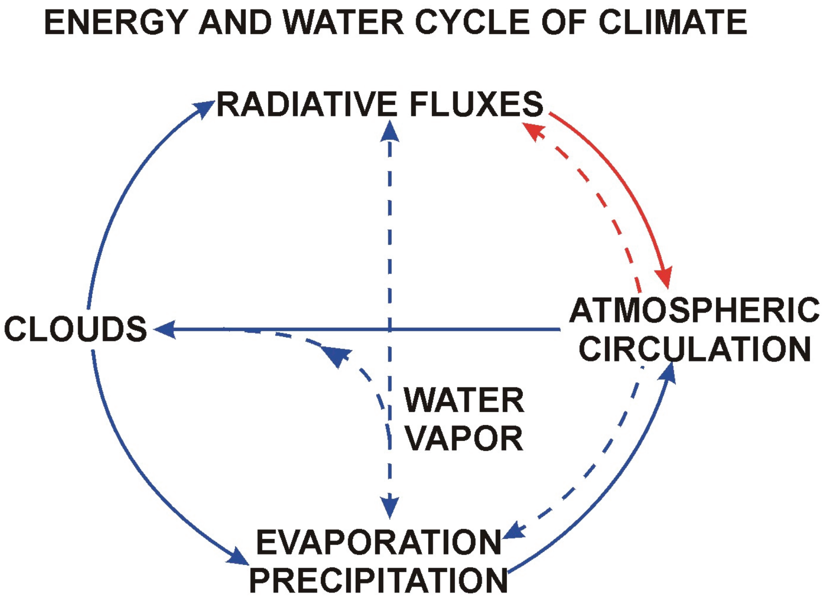

FIGURE 1-1 Schematic of energy and water cycles. Red represents transfers of energy while blue lines show transfers of water. Figure courtesy of William Rossow, City College of New York.

meaning. A good statistical model is built in a way that captures some of the physical processes that control elements of the climate system, or alternative hypotheses about those processes.

The classical method, described at the workshop, for characterizing uncertainty in earth science modeling is through sensitivity analysis. Simply, this method includes changing parameter values in a model to learn how much that parameter affects the model output. This method does not account for the possibility that more than one process represented in the model might rely on the parameter itself, which will affect the uncertainty estimate. In addition, the compounding effect of different sources of uncertainty on different parameters is difficult to quantify through sensitivity analyses.

Alternative statistical approaches define uncertainty through joint probability distributions of parameters. While it is difficult to use this approach to identify the correct parameters and distributions when datasets are small, advances in data collection, management, and processing technologies are increasingly resulting in large datasets. Statistical distributions and their parameters can be estimated accurately when large volumes of data are available. Scientists in the satellite era have this luxury, but are concomitantly faced with massive data volumes that create challenges for processing and analysis techniques. In principle, large, complex, and detailed datasets offer the promise of new knowledge from which to better understand the climate system. Statistical methods that are developed specifically for new data types can better exploit these large, complex datasets that traditional methods (i.e., sensitivity analysis) cannot.

Understanding the uncertainties of different processes in the climate system requires a variety of approaches. Collaborations involving climate scientists and statisticians were identified at the workshop as an effective way to promote the development of targeted new methods that would aid the science community to question all aspects of the data, and geophysical and statistical models. Workshop participants also remarked that modern statistical methods can be useful for fusing data from two different instruments, which is a more challenging problem than is generally appreciated. For example, data assimilation techniques are one approach to addressing the spatial and temporal mismatch between models and observations (Daley, 1991; Luo et al., 2007). As described by workshop participant Anna Michalak from the University of Michigan, such approaches need to account for the spatial and temporal structure of the dataset to allow a better understanding of the physical processes that make up the climate system.

WHY WORRY ABOUT STATISTICAL STRUCTURE: AN EXAMPLE FROM MODELING SNOW DEPTH

Anna Michalak at the University of Michigan described how the statistical properties in remote sensing datasets offer both a challenge and an opportunity. For example, understanding and accounting for statistical dependence, including spatial and temporal correlations, can improve the utility of observational datasets. The opportunity is that by skillfully handling these complexities, we can better take advantage of the full information content of the available data, and use this information to guide high-payout improvements in models of the Earth system.

In some cases, statisticians and earth scientists use similar techniques to evaluate, and take advantage of, the spatial and temporal structure of observations of environmental parameters. For example, spatial statistical techniques allow one to interpolate (the earth scientist's term) or predict (the statistician's term) the value of specific environmental parameters at unsampled locations. The vast majority of environmental parameters (e.g., clouds, precipitation, winds) exhibit spatial and/or temporal correlation, with associated characteristics of scales of variability. As stated by Tobler as the "first law of geography": "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). Both statisticians and earth sciences have used quantitative tools to assess the spatial autocorrelation exhibited by sampled data. Both use variograms and/or covariance functions to quantify the degree of spatial autocorrelation. An accurate assessment of the spatial variability of observed parameters can be used to better understand the underlying physical processes.

Figure 1-2 illustrates how understanding and exploiting the spatial and temporal structure of data can be useful. In this example, a limited number of measurements that are clustered in a non-ideal way are used to estimate the mean snow depth in a valley. Simply averaging the ten measurements of snow depth does not provide a good representation of mean snow depth. Instead, the clustered observations in the left portion of the valley clearly need to be weighted less relative to the isolated observation in the right region. However, how much weight should be assigned to each data point? Spatial statistics methods can be used to determine the degree of spatial variability in the snow-depth distribution based on an analysis of how similar nearby measurements are to one another, and how dissimilar far-away measurements are to one another. This information, in turn, can be used to quantify the optimal weights to be assigned to each measurement. This simple method in spatial statistics allows one to calculate an unbiased estimate of mean snow depth in the valley based on an uneven distribution of measurements.

In Figure 1-3, we look at a hypothetical dataset describing snow depth as a function of elevation, and assuming that the snow depth is also autocorrelated in space (top panel). These synthetic data were generated in such a way that, in reality, there is no overall trend of snow depth with elevation, and any observed trend is therefore the result of randomness introduced in generating the data. This hypothetical dataset is then used to test whether two competing approaches are able to correctly conclude that there is no relationship between snow depth and elevation (middle panel). In the first approach (red line), classical linear regression is used, which ignores the spatial correlation in the data. In the second approach (green line), the spatial correlation is accounted for in the estimation process. In the example shown in the figure, the classical approach incorrectly rejects the null hypothesis that there is no trend between snow depth and elevation, whereas the approach based on spatial statistics correctly does not reject this hypothesis at the 95 percent confidence level. As the experiment is repeated multiple times with new synthetic data (bottom panel), we observe that the linear regression approach incorrectly concludes that there is a trend between elevation and snow depth approximately 20 percent of the time, which is much too high given that the test was run

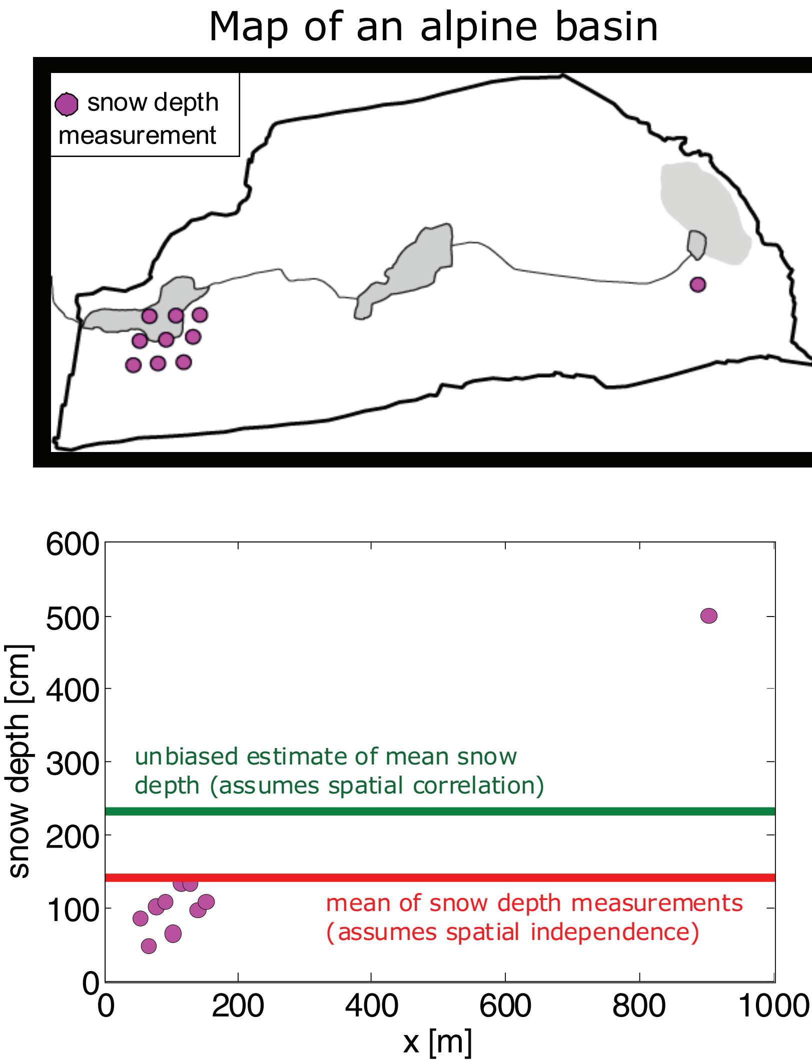

FIGURE 1-2 Example of sampling snow depth in a watershed. Top: aerial map of an alpine basin with sample locations (•). Bottom: snow depth at sampling location versus distance from the left edge of the valley. The red line represents the biased estimate of average snow depth obtained from a simple average of the available observations. The green line represents the unbiased estimate obtained by assigning weights to the observations based on an understanding of the scales of spatial variability of the snow depth in the valley. Figure courtesy of Anna Michalak, University of Michigan. Original figure by Tyler Erickson, Michigan Tech Research Institute.

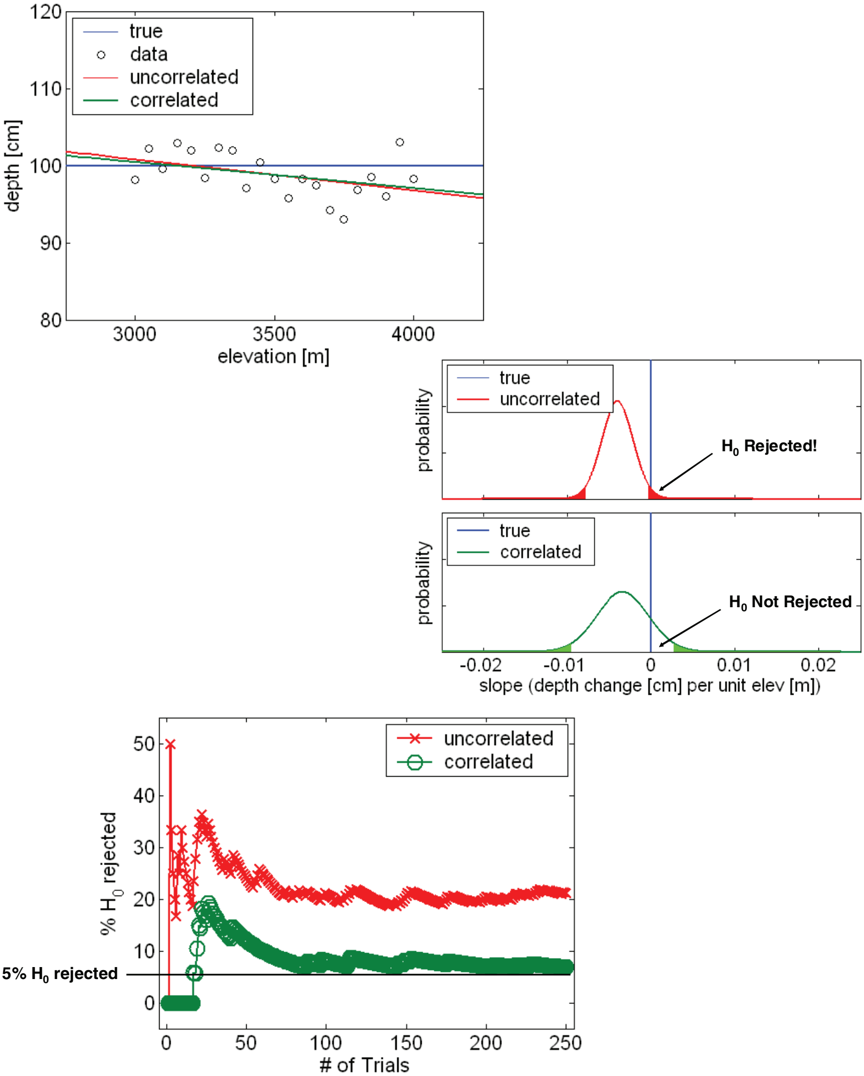

FIGURE 1-3 Hypothetical data on snow depth as a function of elevation. Top: illustrates one case of the generated data, and the estimated slope between snow depth and elevation, using simple linear regression (red line), and an approach that accounts for the spatial correlation of the data (green line). Middle: illustrates the probability distribution of the trend of snow depth with elevation using the two approaches. Bottom: demonstrates that if the experiment were repeated many times, one would erroneously conclude that there was a relationship between snow depth and elevation too often if using simple linear regression. Figure courtesy of Anna Michalak, University of Michigan. Original figure by Tyler Erickson, Michigan Tech Research Institute.

in a way that should have yielded only a 5 percent chance of incorrectly concluding that there was a trend. The approach that accounts for spatial correlation, concludes that there is a trend 5 percent of the time, as expected. Overall, this example illustrates that statistical approaches that ignore spatial and/or temporal correlation inherent in environmental data carry with them an increased risk of erroneously concluding that significant relationships exist between physical phenomena (snow depth and elevation, in this case), and, more generally, yield biased estimates due to their assumption of independent observations.

2

Cross-Cutting Issues

The December 2008 workshop on uncertainty management for remotely sensed climate data identified three major issues that cut across multiple areas of climate research, illustrating the need for a more sophisticated framework for data analysis in order to better understand climate processes:

1. The challenge of validating remotely sensed data.
2. The need to stay focused on the users of climate information, including policymakers. Even if an analysis is perfect and the statistical tools work, the results might not be interpretable and useful for the end-user.
3. The need for strengthening collaboration between the earth science and statistical communities, the importance of leveraging the strengths of each, and the challenges inherent in doing so.

VALIDATION OF REMOTELY SENSED CLIMATE DATA

Validation of parameters is an essential component of nearly all remote sensing-based studies and there are many considerations in performing validation. Errors in different validation techniques are complex and difficult to quantify. The workshop participants discussed many challenges of validating remotely sensed clouds, precipitation, winds, and aerosols, though the presentations did not go into great detail on methodologies. Some common questions include: Can the data meet the needs of remote sensing application? What is the accuracy of the data product, and retrieval product? The following section offers examples of validation techniques for remotely sensed climate data.

Techniques for validating remotely sensed data vary for different geophysical parameters (e g., winds, aerosols, precipitation, clouds). The most common technique is comparing ground measurements to remote sensing observations or modeled results. In many situations, a mismatch exists between the sensor's field of view and the scale at which *in situ* measurements are collected. Ground-based measurements cover small spatial scales while satellite retrievals cover an area of many kilometers. However, in the process of working through a validation, the real structure of the data can be revealed. This was nicely described at the workshop by Tom Bell from the National Aeronautics and Space Administration (NASA), who presented challenges in remote sensing of precipitation. The use of *in situ* measurements for model calibration and validation requires a robust method to spatially aggregate ground measurements to the scale at which the remotely sensed data are acquired (Box 2-1).

As previously mentioned, defining the uncertainty of model parameters is a continuing challenge, but there are multiple methodologies in validation studies that can be combined in an optimal way. Examples described at the workshop were studies to understand fluxes in atmospheric carbon dioxide (CO_2). Some studies will employ measurements of CO_2 concentration to infer sources and sinks, while other studies build biosphere models in an attempt to predict the fluxes. These can be combined, using the biospheric models as a first guess followed by a Bayesian framework to integrate the modeled outputs with atmospheric data to get a best estimate of carbon sources and sinks. In this approach, there is an opportunity to account for the uncertainty in the individual parameters as well as the modeling framework that is used to predict the processes of interest.

Inherent in many different techniques that are used in processing remotely sensed data is the issue of biases. A workshop participant described that bias in validation studies of some geophysical parameters occurs because of the uneven global distribution of surface cloud observations. The oceans tend to be cloudier on average than most of the land, but there are fewer surface observations over the oceans. For example, if a threshold is set for the number of surface observations present in a 2.5 degree grid box before accepting a data point, the global mean that is calculated will depend on that threshold. A threshold will therefore force parts of the earth (i.e., the southern oceans), which are known to be very cloudy, to be omitted from the averaged data. Furthermore, the samples do not stay constant; measurements in a grid box can change from month to month, which introduces a source of variability that is generated from

BOX 2-1
Challenges in Validation of Remote Sensing Data: Comparison of Error for Monthly Average Precipitation

A common challenge in remote sensing datasets is accounting for errors due to sparse sampling. Tom Bell, a NASA scientist, discussed the challenges for precipitation. In general, it does not rain very often and it is difficult to quantify rain amounts through rain gauges and compare this data with satellite measurements taken over many kilometers. The most common method for validating satellite rain estimates is to compare rain gauge data collected over a time interval during which the satellite passes over, with the satellite rain estimates. However, rain gauges do not actually measure what the satellite sees, and the average rain estimates from rain gauges and satellites differ over spatial and temporal scales. Therefore, it is problematic to account for the validation problem as a difference between the satellite average and rain gauge average. The best approach for validating satellite rain estimates is to create a model of errors in precipitation estimates. A spectral model can be used to help predict the best time interval over which to average the rain gauge data when comparing the gauge measurement with a single overflight of the satellite (Figure 2-1). In addition to understanding the time interval to average gauge data, the average area also needs to be determined. Figure 2-1 shows the various sampling intervals between different satellite visits occurring over one month and demonstrates how the error between satellite averages and (surface) rain gauge averages varies in a complex way with the sampling interval of different satellites and the choice of spatial area over which averages are taken.

For example, the error between the rain gauge and the satellite with one sample per day is minimized over larger areas, a swath width approximately 300 km, as shown by the blue curve, representing the Tropical Rainfall Measuring and Aqua missions. However, if satellite measurements were to be taken every three hours, as shown in the black curve, representing the planned Global Precipitation Measurement mission, the error would be minimized by spatially averaging over a swath of width about 70 km. This figure demonstrates not only that there is an optimal point for sampling but that temporal and spatial sampling are related. If satellite data are averaged over 100-km swaths, it is best to look at three-hour data, whereas with 500-km swaths, sampling every 24 hours is sufficient. The curves also show that the spatial phenomenon of rain events tend to occur on the 100-km scale and that there is a "sweet spot" of error for each spatial scale (i.e., three-hr sampling for 70-km swaths).

the sampling, rather than the cloud itself. Hence, not covering the domain of a phenomenon completely (in either space or time) leads to biases. Another bias can result in studies of cloud processes if spatial and temporal autocorrelation is ignored. William Rossow, City College of New York, described that the polar orbiting satellite samples low latitudes twice per day and has spatial scales of approximately 2,000 km. Physical processes that evolve rapidly and on scales smaller than 2,000 km are therefore dif-

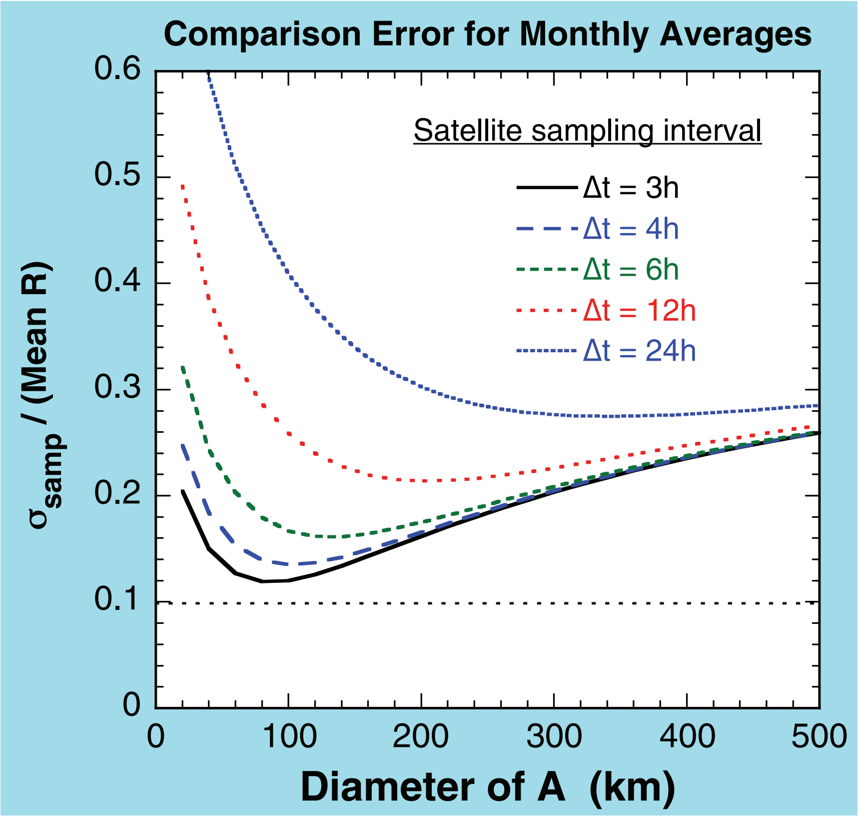

FIGURE 2-1 Relationship between spatial and temporal sampling of different satellites used for measuring precipitation (TRMM—blue curve; GPM—black curve). Figure courtesy of Tom Bell, NASA.

ficult to quantify. To study cloud processes, both the space and the time scales need to be measured appropriately to see the physical process unfold. It is also important not to ignore temporal autocorrelation in the tropics, where convection is rapid, or this will introduce bias into the results. A monthly mean is not always based on 30 independent samples. Caution should be taken in studies investigating interannual variability as this is often based on monthly means. The assumption that a monthly mean is based on 30 independent samples can lead to what looks like climate variation in the data, when it could actually be statistical noise. Box 2-2 gives an example by Tom Bell, NASA, of biases resulting in rainfall measurements.

At the 2008 workshop, Jay Mace, University of Utah, compared and

BOX 2-2
The Need for a Model of Biases:
An Example from Rainfall Measurements

Regression analyses between satellite and parameter measurements are commonly used for calibration studies and to evaluate the error level of the retrievals. However, validation with regression analysis is a common source of bias in precipitation studies. For example, comparing long-term averages of rain gauges to three different windows of time during satellite overpass reveals that the relationship between what the gauge sees and what the satellite sees improves as the averaging interval around the satellite overpass is reduced (Bowman, 2005). Conversely, the agreement is poor when the averaging interval is increased. This comparison can be problematic given that the common methodology (e.g., linear regression analysis) to understand the amount of agreement between the remote sensing estimate and the ground estimate is based on the assumption that the ground measurements are accurate. Rain gauges, however, are an imperfect measure of how much it actually rains, have their own set of biases and sampling issues, and thus, this methodology is not appropriate for precipitation studies given that the ground-based estimates are not true values of what the satellite sees. The regression analysis also does not take into account the minimum overlap in sampling. For example, a poor agreement between the satellite and ground-based estimates does not necessarily mean that the satellite is performing poorly. Rather, the regression technique may not be the proper tool for calibration studies. This is where a statistical model will be helpful in not only describing retrieval error, but also useful for validation exercises. An error model can help disentangle spurious biases, those generated from methodology, from the real biases of the instruments. Validation exercises can generate spurious biases, such as trends that look like biases in the remote sensing method which are not really present in the data, but are byproducts of the analysis method.

validated cloud microphysical properties, with data from multiple instruments, including satellite measurements, ground-based remote sensing measurements, and aircraft measurements. The premise for this approach is that clouds and precipitation influence the radiation and hydrology of the earth through an evolving vertical profile of microphysics. Therefore, scientists need to become more skillful at deriving the vertical profile of microphysics from remote sensing data in a statistically meaningful way.

An example is the comparison of Moderate Resolution Imaging Spectroradiometer (MODIS) derived ice water paths with ice water paths derived from a ground-based validated radar algorithm. One is a snapshot spatially averaged measurement and the other is a time average point measurement. For cirrus clouds, a spatial average is generated by averaging the MODIS measurements over a rectangle that is oriented along the mean wind at the cloud level, while the ground-based radar data is an average over a period of time when the cloud layer remained uniform (Mace, 2001). Comparing these measurements over a long period of time will aid in determining the error characteristics of the satellite data. However, the uncertainty in the data, combined with the uncertainty in the science, requires techniques to quantitatively assess these errors, and more sophisticated statistical approaches would be helpful in accomplishing that objective.

The validation problem for measurements of aerosols presents other issues. Lorraine Remer of NASA spoke at the workshop about validation of aerosol optical depth (AOD) measurements, which are measures of the column-integrated extinction, and the amount of light that is either scattered or absorbed as it passes through the aerosol layer, which gives an indication of the amount of aerosol. A satellite measures not just the radiation scattered from the aerosol layer, but it also collects some radiation that made it through the aerosol layer from the earth's surface. The surface effect needs to be removed from the satellite signal to estimate the extinction, which requires assumptions about the aerosol and the surface, thus leaving room for error.

A more direct way of measuring AOD is to use a sunphotometer on the ground to measure the transmitted light directly. By combining measurements of the sunlight at the top of the atmosphere with the amount of sunlight at the surface, the extinction can be determined. This approach uses fewer assumptions and under the best conditions MODIS can retrieve AOD to within ±0.03, and a well-calibrated sunphotometer can measure it within ±0.01. The widespread network of sunphotometers called AERONET retrieves data globally. The primary challenge with this technique is the mismatch between spatially varying MODIS data and the temporally varying sunphotometer data so there are only select areas with coincident coverage in measurements between MODIS and AERONET. Since AERONET is a land-based network, it is difficult to match an overpass with an aerosol observation, and, as addressed earlier, the location of the ground-based observation within the satellite grid square is a consideration in validation process.

There are many types of aerosols (e.g., sulfates, black carbon, sea salt). Some are natural, while others are anthropogenic. Uncertainties in aerosol models, as presented by Joyce Penner of the University of Michigan, result from uncertainties in the sources, types, and radiative properties of aerosols. Validation of these models cannot come solely from comparisons with ground-based data like that of AERONET, because the measured AOD is a composite of the effects of the different aerosol types the models are attempting to simulate. Aerosol models also cannot be validated with satellite data alone, but require a suite of observational data from surface stations of different aerosol species. Sources of uncertainty are also associated with cloud interaction, chemical production, vertical transport, and grid resolution. As noted by several workshop participants, the climate science community does not have a good understanding of several relevant processes, potentially creating biases in the models.

Today it is common to produce daily and monthly mean maps of the distribution of aerosols which can carry their own uncertainty. For example, few retrievals available in a particular grid square will generate a higher degree of uncertainty, ultimately affecting the mean distribution maps. Different methods of weighting and averaging the data result in different distribution maps and mean optical depths. Moreover, not all retrievals carry the same level of confidence. A number of options and methods need to be considered when analyzing satellite and sunphotometer data, all within the context of the application of the dataset.

THE NEED TO STAY FOCUSED ON THE END-USER

Several participants at the workshop noted that greater attention to uncertainties in climate data would help to address important questions in climate research and policy. Likewise, greater attention to uncertainty quantification, in part, can be driven by specific needs of researchers, policy makers, or other end-users of the remotely sensed product. This is in contrast, for example, to the situation in the nuclear weapons laboratories, which have already developed sophisticated methods of uncertainty quantification in order to address the policy question of whether aging warheads remain safe and functional in the absence of complete testing. With regard to testing of climate data and models, experiments are run to better understand specific physical processes, but there is no option for controlled experiments of complete systems.

As the uncertainties of remote sensing instruments become better characterized, the question arises of how to represent that information in a way that is useful to researchers who examine the data. How does one capture the uncertainty of an entire dataset so that the end user can use the information? This remains an open question and a difficult one. Researchers, both in the geoscience and statistics communities, need to become very familiar with a dataset, the data-collection process, and the applications of the data in order to understand all the issues and assumptions that are associated with the data and its use. Understanding the uncertainties of different processes in the climate system requires different approaches. For example, a challenge recognized by the community is reconciling initial states in models with the observational data. The initial states must reflect that uncertainty, thus, every dataset needs attention by the scientists to account for that uncertainty in the forecast model.

Collaborations between remote sensing climate scientists and statisticians can potentially result in products that are ultimately more useful to end-users. Participants pointed out the value of true collaborations, wherein statisticians who become familiar with an entire data-collection system will help characterize the uncertainty or the variability in the equations that model the physical processes, and not just the uncertainty in the data collection. The two communities working together may lead to different ways to look at the data for the science questions.

Jay Mace, University of Utah, presented an example of how uncertainty analysis can contribute to climate modeling. Sanderson et al. (2008) published results of a statistical study of climate models in which the investigators varied the parameterizations over some parameter space and looked at the sensitivity of the model results to the tuning. One of the tuning parameters examined was the ice crystal fall speed, and that paper concluded that it is an important parameter in a climate model, in part, because this process takes ice out of the upper troposphere where it shields the upward infrared radiation. Thus, ice fall speed turns out to be a powerful tuning knob for a climate modeler. In response to that paper, Deng and Mace (2008) published a study that looked at Doppler velocity data from the Atmospheric Radiation Measurement (ARM) sites and parameterized the ice crystal fall speed as a function of ice water path and temperature. This is an example of how information from various remote sensors can feed directly into model parameterizations. Because the many different climate models in existence use different parameters and generate different results, better knowledge of the uncertainties is critical for building the next generation of climate models.

Workshop participants noted that statisticians and earth scientists need to consider the end-user of the product because different end-users will require different applications of remotely sensed data, and this will determine how the dataset will be processed. For example, end-users, such as policy makers, want to know about uncertainty in climate projections, which includes model uncertainty, observational uncertainty, and the overall uncertainty of our knowledge. As the climate community focuses more on addressing the questions posed by policy makers and other end-users, the collaborations between earth scientists and statisticians will likely be encouraged. A remote sensing scientist produces a product, which has biases and uncertainties that are spatiotemporally correlated as a function of the statistical properties of the observed fields and the manner in which they were sampled. A modeler uses the products, for example as gridded, averaged fields, and introduces biases and uncertainties into the predictions, projections, and analyses. The remote sensing scientist and the modeler must collaborate to ensure the accuracy of the product. This implies an enterprise-wide need for collaboration between the data producers and data users.

CROSS-DISCIPLINARY COLLABORATIONS BETWEEN CLIMATE SCIENTISTS AND STATISTICIANS

The challenge of understanding a system as complex as climate requires a partnership between geoscientists and statisticians because neither community has all of the expertise that is required. Each community is attempting to understand climate processes interacting at multiple scales and the workshop participants called for more sophisticated techniques to study these interactions that can benefit both communities. The workshop demonstrated that the community is still at a fairly rudimentary stage of understanding the complexity of the climate system. Being at an early stage of understanding, it is not always obvious to outsiders, including those who might fund such investigations, that these lines of research and their necessary collaborations are essential to progress in climate science. The development of tools and methods to approach climate datasets for analyses is an ongoing endeavor. Research into better statistical methods enables us to, among other things, account for autocorrelation and provide ways to infer greater information from sparse data.

A primary motivation for this workshop, and a fundamental area for collaboration recognized by the workshop participants, lies in quantifying uncertainty in climate records, given that a better understanding of the climate system can be obtained with a more sophisticated approach to handling uncertainties. This includes accounting for the uncertainty in the individual model parameters and also the uncertainty that is inherent in the modeling framework. Statisticians can help formalize how geoscientists can use well-characterized uncertainties to ultimately understand the uncertainties in the forecast model.

Other areas were suggested as being ripe for collaboration between climate scientists and statisticians, including the monitoring of simple state variables in the atmosphere, and understanding the large cycles in the atmosphere, such as the water cycle and the carbon cycle and the interactions between them. It is difficult to examine these cycles due to their complexities, and statistical methods can be useful for teasing out information. Another area for collaboration is looking deeper into the interactions between variables, including a better understanding of forcing variables such as CO_2 concentrations and aerosols. We need a conceptual framework for applying different statistical techniques to these areas. Statisticians need to tackle the full state space at the full resolution, and physically quantify and validate the results of any models that are generated. The agencies that fund remote sensing benefit when statistical investigations explore data in a new way and find additional value in systems that are already aloft, as well as when they provide information that will guide resource decisions in the future.

Collaborations between geoscientists and statisticians open up the opportunity for tailoring methods to fit a specific geophysical situation, with specific improvements in accuracy, precision, and/or run time. Statistical textbooks, publications, and software do not necessarily provide the necessary technology transition, making collaborations critical. For example in a method like kriging, presented by Noel Cressie at Ohio State University,[1] it can be difficult to discern how to define a statistical model and probability distributions that will lead to optimal interpolation. The standard formulation assumes that parameters follow simple probability distributions (e.g., Gaussian distribution), but real data have many structures. Creating a prior distribution that incorporates an understanding, sometimes quite subtle, of the actual physical processes is a task best performed through collaborations. Workshop participants described that most productive collaborations are two-way, in that both perspectives, of the geoscientist and the statistician, are applied to understanding all aspects of the problem. Certainly, the geoscientist may evaluate the appropriateness of statistical steps and the assumptions embedded in them. Conversely, the statistician may also be intimately familiar with how the geoscientists developed the mathematical models, because some of the uncertainties and assumptions are embedded in the equations. As described earlier, a primary goal for climate scientists is to understand the physical processes that are directly relevant to climate model, and this can be addressed through the use of statistical models.

Many of the workshop participants recognized that interdisciplinary work is hard. It is difficult to get funding because research proposals need to convince two separate communities that often do not communicate. Additionally, if a study generates a new approach that is too complicated, that approach will often not be utilized. An advance that requires others to learn a new technique has to be of obvious value and explainable. Complex solutions can be dangerous if they hide assumptions that were made during the derivations of the method, but which are not appropriate for the remote sensing application. This is another reason why a statistician involved in collaboration must be intimately involved in the geosciences modeling to recognize such a situation. Climate research is ripe for additional statistical sophistication, even at the risk of adding complexity, because the climate model predictions are critical to society.

There are also impediments to cross-disciplinary collaborations. First,

[1] A detailed description of kriging is presented in Appendix B, in the talk summary for Noel Cressie.

communication across the disciplines is a necessity that was highlighted in the workshop discussion. Both communities, geosciences and statistics, will benefit from an ongoing interaction as well as continued effort to understand the literature outside their field to find what has been done in a certain area in order to make improvements. Second, if done well, cross-disciplinary collaboration can be a very productive, but requires commitment from both communities. And third, workshop participants expressed that the funding is not adequate to support enough cross-disciplinary collaboration.

What incentives exist for collaborations? For many years, statisticians have had little trouble finding interesting work. It takes a special kind of statistician to want to be in this cross-disciplinary area. Cross-disciplinary collaboration also needs support from the climate community; that is, geoscientists need to make known that they want statistical expertise to solve some of these complex problems. Participants also felt that the federal agencies that fund climate research need to be aware of this constraint.

Box 2-3 illustrates the nature of the problem as ensemble approaches to large datasets are becoming more common. A key point in the discussion is that a simple mean is not the complete answer. Rather, the statistics and earth science communities can come together to take advantage of the variability in the data. The structure of the dataset needs to be analyzed to better understand the multiple physical processes that make up the climate system.

BOX 2-3
The Nature of the Problem: Lessons from a Santa Claus

The use of ensembles, permutations of data, provides a sampling of the space over which the data can range and can be an effective way to begin thinking about complicated statistical problems. For example, Figure 2-2 presented by Doug Nychka, National Center for Atmospheric Research, includes 100 variants of a depiction of Santa Claus. Kriging, a common statistical technique, allows us to generate the best statistical estimate of the mean of the variants, and an ensemble around that mean can provide information about the uncertainty. However, a traditional method such as kriging may not be the optimal way to describe Santa's features, as this approach represents a point-wise "expected value" and does not preserve the spatial relationships present in individual sample images. The "mean Santa" shown in the large image does not capture specific information such as Santa's nose or other particulars. However, having the 100 realizations of Santa enables one to query what is known. In addition, the ability to infer details about the physical processes associated with Santa (e.g., his delivery of presents on Christmas Eve) is complicated by the fact that the only information available is through parameterizations that come from modeling efforts, much like the parameterizations in global climate models. Moreover, when a parameterization is needed, it might be better to rely on one or more of the 100 variants rather than basing the parameterization on the mean Santa, which does not directly represent any underlying model of processes as do the individual depictions. It is important to remember that the ensemble of variants is generated according to some assumptions about how to sample parameter space, and it might not be the best sampling for every purpose.

Workshop participants argued that climate is defined by characteristic variation, not by average values. Climate is the product of a complex fluid dynamical system, and the products of such systems are defined by the system's history, not its equilibrium. The atmosphere is constantly evolving, whereas the different versions of Santa in Figure 2-2 are not related in a fundamentally dynamical way. Rather than trying to understand the outcome of the averaging, it may be more beneficial to solve the problem of inaccurate model representation of climate. While kriging and generating an ensemble provide a valuable comprehensive view of a dataset, the ultimate goal of analysis is to discover the more subtle features of the structure of the distribution, which are often lost in simple analyses focusing on the mean, or average, or an observed process. For example, the average of a Mozart sonata is a single note, but that does not convey anything of importance about the piece of music.

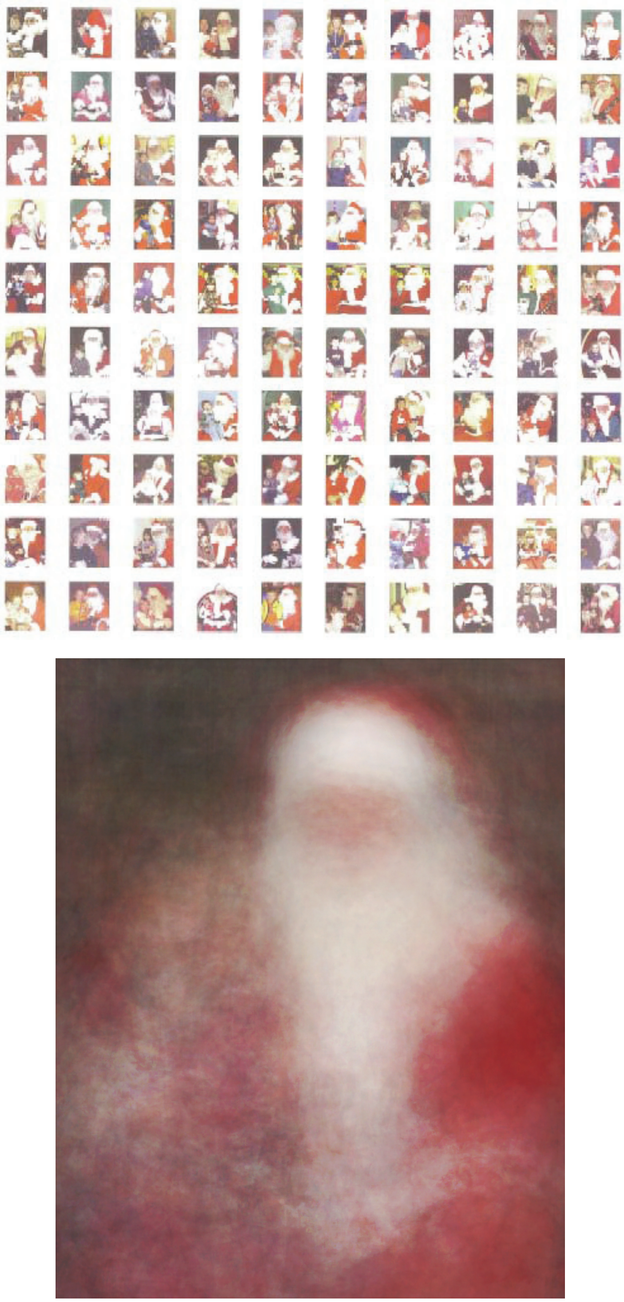

FIGURE 2-2 Top: 100 variants of Santa Claus; Bottom: the average Santa Claus based on the 100 variants. Figure courtesy of Jason Salavon.

3

Concluding Thoughts

As conveyed by many participants at the National Academies' workshop on uncertainty management for remotely sensed climate data, it would be helpful if the climate research community as a whole could settle on priority questions where collaboration with the statistics community would be most beneficial. The advancement of statistical techniques could then focus on these fundamental science questions. For example, the importance of climate models for policy making suggests that improved statistical techniques for improving their parameterizations and analyzing their output could have substantial benefits for both the scientific community and society as a whole. There are historical precedents for this type of progress; for example, in the 1970s and 1980s the community recognized that it was critical to gain a better understanding of how clouds affect the earth's radiation budget. Today, we have a much clearer picture of how clouds alter the transfer of radiation through the atmosphere on a variety of timescales, and progress on this topic has allowed scientists to develop a more comprehensive, although still incomplete, understanding of the feedbacks between clouds and other aspects of the climate system.

As described throughout the workshop, the components of the climate system are coupled and interact in multiple ways and at multiple scales, which makes it difficult to discern which components are contributing the most to the uncertainties in our knowledge. In addition, the challenges associated with quantifying uncertainties in a certain geophysical parameter (e.g., precipitation) are typically unique, and cannot necessarily be readily translated to other system components. However, a statistical framework for evaluating uncertainties throughout the system can help lead to a logical roadmap for a research enterprise.

As conveyed by many workshop participants, stronger collaboration between the earth science and statistics communities would likely result in many benefits. Workshop participants suggested that additional, more focused workshops might be organized as a way of spurring progress in understanding uncertainties associated with particular geophysical processes and with some particular data-collection challenges (e.g., measuring aerosols or cloud cover). These follow-on workshops could probe more deeply into particular models, research, and challenges. Examples might include the advancement of statistical techniques to address spatial and temporal autocorrelation in large datasets and methods to incorporate more physical knowledge and physical modeling into the statistical models that will help improve calibration and validation studies.

With the rich collection of remotely sensed data, the workshop participants discussed that considerable progress could be made by going beyond simple monthly and annual averages to describe the climate system, and that modern statistical methods had much to offer in the area of representing the physical processes that make up the climate system. There is a wealth of data to be processed, and analysis of this data requires both good physical models and modern statistical methods to fully understand the biases and residual errors. In the era of advanced earth science collection and processing techniques, the fusion of multiple datasets is a challenge in the remote sensing community and represents another rich area for collaboration, as does data assimilation.

The goal for a statistical framework is to account for uncertainty not only in individual parameters but also in the entire modeling framework used to predict the processes of interest. An overall statistical framework for accounting for uncertainty in remotely sensed climate data and in climate models might also assist in the development of an integrated strategy for communicating uncertainties. Such progress would not only aid research in the earth science and statistics communities, but will result in more useful information for the climate policy community.

References

Bowman, K.P. 2005. Comparison of TRMM precipitation retrievals with rain gauge data from ocean buoys. *Journal of Climate* 18:178-190.

Cressie, N., and G. Johannesson. 2006. Spatial prediction for massive datasets. In *Mastering the Data Explosion in the Earth and Environmental Sciences: Proceedings of the Australian Academy of Science Elizabeth and Frederick White Conference*. Canberra, Australia: Australian Academy of Science. 11 pp.

Cressie, N., and G. Johannesson. 2008. Fixed rank kriging for very large spatial data sets. *Journal of the Royal Statistical Society, Series B* 70:209-226.

Daley, R. 1991. *Atmospheric Data Analysis*. Cambridge, UK: Cambridge University Press. 457 pp.

Deng, M., and G.G. Mace. 2008. Cirrus cloud microphysical properties and air motion statistics using cloud radar Doppler moments: Water content, particle size, and sedimentation relationships. *Geophysical Research Letters* 35, L17808, doi:10.1029/2008GL035054.

Luo, M., C.P. Rinsland, C.D. Rodgers, J.A. Logan, H. Worden, S. Kulawik, A. Eldering, A. Goldman, M.W. Shephard, M. Gunson, and M. Lampel. 2007. Comparison of carbon monoxide measurements by TES and MOPITT: Influence of a priori data and instrument characteristics on nadir atmospheric species retrievals. *Journal of Geophysical Research* 112:D09303, doi:10.1029/2006JD007663.

Mace, G.G. 2001. Atmospheric Radiation Measurement Program Southern Great Plains Case Study, 6 March 2001. Available online at http://www.met.utah.edu/mace/homepages/research/archive/sgp/sgp.html, accessed July 17, 2009.

NRC (National Research Council). 2007. Environmental data management at NOAA: Archiving, stewardship, and access. Washington, DC: The National Academies Press.

Sanderson, B.M., C. Piani, W.J. Ingram, D.A. Stone, and M.R. Allen. 2008. Towards constraining climate sensitivity by linear analysis of feedback patterns in thousands of perturbed-physics GCM simulations. *Climate Dynamics* 30:175-190.

Shi, T., and N. Cressie. 2007. Global statistical analysis of MISR aerosol data: A massive data product from NASA's Terra satellite. *Environmetrics* 18:665-680.

Shi, T., B. Yu, and A.J. Braverman. 2002. MISR Cloud Detection over Ice/Snow Using Linear Correlation Matching. Technical Report 630. Berkeley, CA: University of California Berkeley Department of Statistics.

Shi, T., B. Yu, E. Clothiaux, and A. Braverman. 2004. Cloud detection over ice and snow using MISR data. Technical Report 663. Berkeley, CA: University of California Berkeley Department of Statistics.

Tobler, W.R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46:236.

Appendix A

Workshop Agenda

WORKSHOP ON UNCERTAINTY MANAGEMENT IN REMOTE SENSING OF CLIMATE DATA

December 4, 2008
The Doubletree Hotel
1515 Rhode Island Ave., NW
Washington, DC 20005

8:30 Welcoming remarks and overall workshop goals
Amy Braverman, Jet Propulsion Laboratory

Session A: Introduction

8:40 Differences in terminology, techniques, and approaches between statisticians and earth scientists
Anna Michalak, University of Michigan

9:00 Remote sensing of surface winds
Ralph Milliff, Northwest Research Associates, Inc., Colorado Research Associates

9:30 Remote sensing and precipitation
Tom Bell, NASA

10:00 Discussion
Moderator: Amy Braverman, Jet Propulsion Laboratory

10:15 Break

Session B: Clouds

10:30 Different types of uncertainties in cloud data sets
William Rossow, City College of New York, CUNY

11:00 Machine learning techniques for cloud classification
Bin Yu, University of California at Berkeley

11:30 Validation of cloud property measurements from multiple instruments
Jay Mace, University of Utah

12:00 Discussion
Moderator: Karen Kafadar, Indiana University

12:30 Working Lunch

Session C: Aerosols

1:30 Uncertainty issues associated with remotely sensed data sets for aerosols
Lorraine Remer, NASA

2:00 Spatial statistics with an emphasis on aerosol data
Noel Cressie, Ohio State University

2:30 Discussion
Moderator: Steve Platnick, NASA

3:00 Break

Session D: Integrating models and data

3:15 Aerosol and cloud representation in global models
Joyce Penner, University of Michigan

3:45 Data assimilation as a hierarchical statistical process, interacting dynamically with modeling
Christopher Wikle, University of Missouri

4:15 Discussion
Moderator: John Bates, NOAA

Session E: Making progress through practical and institutional barriers

4:30 The practical and institutional barriers for making progress on developing and improving statistical techniques for processing, validating, and analyzing remotely sensed climate data
Doug Nychka, NCAR

5:00 Discussion
Moderator: Amy Braverman, Jet Propulsion Laboratory

5:25 Wrap-up and final remarks

5:30 Adjourn

Appendix B

Summaries of Workshop Presentations

DIFFERENCES IN TERMINOLOGY, TECHNIQUES, AND APPROACHES BETWEEN STATISTICIANS AND EARTH SCIENTISTS

Anna M. Michalak, University of Michigan

Characterizing the complexity and quantifying the uncertainty in environmental systems will aid in better understanding the systems and improving model forecasts. However, the differences in terminology and approaches used by the statistics and earth science communities are just one of the impediments to successfully analyzing remote sensing climate datasets. This talk illustrates some applications of statistical methods for optimizing the use of *in situ* and remote sensing datasets of the climate system. In order to assess the predictions obtained using models that integrate such datasets, tools must be developed that quantify the full uncertainty associated with such models, rather than simply evaluating the sensitivity of model predictions to a set of model parameters. In many cases, the uncertainty associated with the conceptual framework of the models and the specific parameterizations included in the models, outweigh the uncertainty caused by incomplete knowledge of individual parameters.

When sparse spatial data are integrated in analyses, uncertainties that arise due to spatially and temporally non-uniform sampling can be accounted for using the principles of spatial statistics. Because classical statistics is based on the assumption of independent observations, these tools do not account for the spatial and/or temporal autocorrelation inherent to the majority of environmental phenomena. This can lead to biased estimates and erroneous identification of relationships between parameters. Developing statistical tools that explicitly account for spatial or temporal autocorrelation avoids such errors. In addition, using models that quantify and account for spatial and/or temporal correlation can decrease the uncertainty associated with model predictions because the spatial or temporal information footprint of available data can be assessed and used to inform the model.

Spatial statistics tools can be used to combine data collected from different instruments with differing resolutions, and to reduce uncertainties associated with data interpolation, among other things. This talk emphasized that principles of spatial statistics can address many of the challenges in geoscience. The simplest examples of spatial statistics are interpolating data and generating realizations (i.e., equally likely scenarios) of a given process given sparse data. These principles can be applied to data at any scale. On a global scale, for example, data from the Orbiting Carbon Observatory, a satellite designed to measure carbon dioxide (CO_2) from space to improve our understanding of global CO_2 concentrations, would have contained large gaps due to the satellite track and the presence of clouds and aerosols. Methods based on geostatistics are being developed to generate estimates of the global distribution of CO_2 based on such data, by first characterizing the degree of spatial variability in the CO_2 observations, and using this information to estimate CO_2 for portions of the globe that are not measured. On a local scale, similar principles have been applied to a project that assesses areas of low oxygen in Lake Erie. It is difficult to quantify the extent of the Lake Erie dead zone and how it varies from year to year because the *in situ* measurements are sparse. Therefore, new statistical techniques were developed to identify remote sensing variables that are good predictors of the dissolved oxygen concentration, and these variables are then merged in a geostatistical framework with available *in situ* data to estimate the spatial extent of hypoxia, and how it varies across years. Results show that fusing the *in situ* and remotely sensed data yields a more realistic distribution of the extent of hypoxia, with lower associated uncertainties, when compared with the results using only the *in situ* data.

Spatial aggregation, averaging, and linear interpolation are often using to merge data collected from multiple remote sensing instruments. Such approaches, however, do not yield at optimal estimate at the target scale of analysis, and estimated values can be influenced by samples in neighboring pixels. This problem is exacerbated when remote sensing data are "re-sampled" multiple times. Spatial statistical tools applied to measurements from one or more instruments that may have different resolutions, coupled with an understanding of the point spread function, and a quantification of the spatial and temporal covariance, can be used to optimally combine data from multiple sensors to yield estimates at any target resolution.

Three areas that are ripe for collaboration between statisticians and earth scientists are (i) the development of rigorous tools for quantifying the uncertainty associated with parameters using in climate models, (ii) the development of tools for comprehensively quantifying the uncertainty associated with model predictions, including errors caused by the model formulation, and (iii) developing tools for integrating data across spatial and temporal scales. Furthermore, these two groups can come together to build a better understanding of the physical processes that are relevant to climate models, ultimately leading to improved physics-based models and projections. This could be achieved by developing statistical models that emulate many of the underlying physical processes, and explicitly account for the uncertainty associated with all model components.

REMOTE SENSING OF SURFACE WINDS

Ralph Milliff, Northwest Research Associates

Bayesian hierarchical modeling (BHM) is a fundamental statistical approach for addressing problems in remote sensing climate datasets. Building blocks for BHM include the data-stage distribution (i.e., likelihood), which quantifies the uncertainty in observations, and the process-model-stage distribution (i.e., prior), which quantifies the uncertainty in the physics of the process. These stages introduce parameters, and estimates in the posterior of the parameters can be determined. One advantage of the BHM approach is that it takes multiple data sources into one model. The first two stages allow for satellite data and data from other platforms to be combined with the physics. Estimates in the posterior of the parameters introduced in a model may be influenced by both the data-distribution and process-model stages. BHM has been successful in characterizing surface-wind processes over the Labrador Sea, the tropical ocean and the Mediterranean. This talk describes how BHM may be used to characterize surface-vector winds for ensemble data assimilation and an ocean forecast model for the Mediterranean Sea. The research further demonstrates that surface-vector winds can be used to identify uncertainties in dependent variables in the forecast model.

The Mediterranean ocean forecast system uses realizations from the posterior of the wind BHM to generate physically realistic spreads in the forecast initial conditions from which ocean ensemble forecasts can be launched. The strategy is to exploit the precisely characterized uncertainty of satellite observations. The prediction grid shows clusters of wind vectors, and the uncertainty in the clusters is a function of space and time. This is an alternative way of thinking about uncertainty. Traditionally, geoscientists think of an error in the satellite as projected to the error in the field. However, the uncertainty in this case depends on whether or not the satellite was there and the strength of the surface winds. Thinking about uncertainty in this way can lead to important understandings of processes that depend on the dynamics of the system. Subtle shifts in the spread of wind vectors from the posterior generate different wind stress curls, the vorticity in the wind that drives the eddy field in the ocean. The vorticity-driven eddies occur on a scale (i.e., the ocean mesoscale) that is a primary source of uncertainty in the ocean forecast (i.e., at synoptic scales). Over the Mediterranean Sea, the eddies affect the general circulation of the basin including deep water formation. Mistral events, where cold dry air blows off the European continent in late winter, generate large wind stress curl, the uncertainty of which can be characterized. The forecast model generates physically realistic spreads in the forecast initial condition, and the spread is focused on uncertain scales of the general circulation in the Mediterranean. The uncertainty is temporally and regionally specific.

This methodology can also be used to better understand the physics from the posterior distribution of BHM parameters in the process model (i.e., the pressure gradient terms in the present case). The posterior distributions for those parameters on the geostrophic and ageostrophic terms, with and without the QuikSCAT data, are examined. The scatterometer is providing more ageostrophic information than the weather center fields. The example demonstrates that the prior can be more complex.

REMOTE SENSING AND PRECIPITATION

Tom Bell, National Aeronautics and Space Administration

Precipitation is highly erratic in space and time, and success in developing adequate statistical descriptions of precipitation has been mixed. The fact that precipitation rates are zero much of the time over much of an area is a special nuisance. In addition, remote sensing errors in this area are complex and difficult to quantify. Rain gauges, providing the most accurate ground-based measurements of precipitation, cover only a very small area, essentially the size of a bucket, and are used to validate time-specific satellite measurements taken over many kilometers. This poses many challenges in modeling the errors associated with these estimates to characterize the types of biases in the satellite estimates.

Rain gauges provide time-averaged rain rates, whereas satellites take measurements from a volume of space above the ground. It is, therefore, problematic to try to make validations and comparisons between these two very different sources of data. Analogous challenges can arise in comparing estimates collected from different satellites. For example, the Tropical Rainfall Measuring Mission (TRMM) satellite has a radar instrument that measures rainfall amounts at every level in the atmosphere, as well as a visible infrared instrument and a microwave instrument that detects raindrops. The observational swath of the microwave instrument is 800 km wide, and the satellite makes roughly one observation per day. This results in a very sparse dataset distribution in time that also becomes increasingly complex to describe once one considers the orbital variations of the satellite and its measurements. It is important for statisticians to appreciate the complexity of the dataset to gain an understanding of the issue prior to analysis.

The temporal and spatial discrepancies between the rain gauge data and the satellite measurements lead to several questions about the validation process that must be considered: what size area should be used in averaging satellite data, what is the optimal time period over which the rain gauge data should be averaged, and if a satellite passes several times in one month, how long before and after the pass should the rain gauge data be averaged? A spectral model can be very useful for spatial and temporal statistics. The model is designed to represent the larger areas that tend to evolve more slowly in time and have very long correlation times versus small areas with short correlation times. This is an important characteristic in rainfall and other geophysical fields. Small events evolve and move over an area more quickly than large events. An example of the uses for the spectral model is represented in Box 2-1, illustrating that there is an optimal time averaged sampling interval.

The complexity of rain distributions requires better statistical models that can account for the rain-rate distribution, the space-time behavior distribution, the multivariate distribution, and also the error. The underlying assumptions need to be carefully considered before the models are used for any specific application. In addition, it is highly beneficial to both the earth science and statistics communities when techniques are well understood and accessible.

DIFFERENT TYPES OF UNCERTAINTIES IN CLOUD DATASETS

William Rossow, City College of New York

Clouds pose multiple challenges for interpreting datasets to detect climate variability. Better statistical methods can aid the geoscience community by helping to better define important measurements related to clouds, such as cloud area and coverage, point-area comparisons of clouds, sampling of cloud variability, and monitoring cloud evolution. Remote sensing of clouds is very difficult in part because of resolution and the effects of finite sensitivity and detection. The latter effects had been largely ignored until the last couple of decades. Cloud size varies greatly, and this complicates the measurement of cloud fraction. A well-designed detection threshold that produces a good estimate of cloud fraction despite the problems associated with size and area measurement has been demonstrated by the International Satellite Cloud Climatology Project (ISCCP). In terms of point-area comparisons, it is important to be cautious when comparing ground measurements and satellite data. The two represent very different measurements and have different accuracies and errors associated with them.

Cloud sampling requires knowledge of the scales of the variability, the sampling scale, the precision of the measurements and the end goal to which the estimates will be applied. Clouds exhibit the most variability at larger spatial scales. The variability is dominated by the same scales as is the atmospheric circulation. In addition, temporal autocorrelation must be accounted for when considering inferences about monthly means. For example, monthly data may consist of only 10 truly independent samples rather than 30 because of day-to-day variations. If not properly calculated, this can lead to monthly means that, for instance, give misleading information about inter-annual variability, such as El Niño-Southern Oscillation (ENSO) patterns. Monthly means are commonly used in discussions of climate variation, but the mean can also mask small variations in the dataset. In addition, statistical methods need to be improved so we can remove the effects of autocorrelations.

Cloud interactions with the climate system should be considered in the context of multiple processes and the relationships between those processes. Atmospheric circulation, clouds, precipitation, and radiative fluxes are all related and often occur within the same cloud system. Clouds and precipitation, for example, are fluid systems that are coupled spatially and temporally and should be considered together. Studying just one parameter at a time does not provide a full picture of the dynamics that are actually occurring. Similarly, averaging can provide valuable information, but averaged values do not help us determine the structure of the physics. Thus we need more sophisticated ways to analyze the non-Gaussian, highly nonlinear, multi-variate, and multi-scale coupled data that are available. A partnership between the fields of geosciences and statistics is needed to develop techniques and models capable of rendering the available data into something interpretable and useful. There is a multitude of climate models that utilize different parameters and generate different results. The community needs better ways to determine which of these models best represent the climate system.

MACHINE LEARNING TECHNIQUES FOR CLOUD CLASSIFICATION

Bin Yu, University of California at Berkeley

The uncertainties in cloud radiation feedback on global climate remain a great obstacle in understanding and predicting future climate changes. This talk describes a case study of cloud detection over the Arctic region using machine learning methods and data generated from the Multi-angle Imaging SpectroRadiometer (MISR). The MISR's algorithm retrieves the cloud height and cloud movement by matching the same cloud from three angles. However, clouds above snow- and ice-covered surfaces are particularly difficult to detect because the temperature and reflectivity of the clouds are similar to the snow and ice surfaces. Retrieval is also particularly difficult when trying to detect thin clouds in polar regions. This talk describes a methodology for addressing these challenges. The starting point was to measure the ground rather than clouds, because the ground is fixed. It was found that correlations between angles are strong over snow- and ice-covered surfaces and weak in areas covered by high clouds. Exploiting the multiple angles in MISR, the linear-correlation matching clustering (LCMC) (Shi et al., 2002) technique was developed to distinguish between smooth surfaces and thin clouds. However, the LCMC in polar regions was insufficient for detecting smooth surfaces, such as frozen rivers, and areas of thin clouds, which led to the development of an enhanced LCMC (ELCMC) (Shi et al., 2004).

The question remains of how to quantify the clouds. This talk compares performance of the ELCMC against two other machine learning techniques, the quadratic discriminant analysis (ELCMC-QDA) and the ELCMC support vector machine. The latter was determined to be computationally slow. Expert labeling provides the highest accuracy, but that is too slow for most purposes; the ELCMC-QDA provides about 92 percent of the accuracy of expert labeling. The use of expert labels improve the accuracy rates, however, they are expensive and impossible to obtain for every block of data. Information from the Moderate Resolution Imaging Spectroradiometer (MODIS), was then utilized to inform accurate labels, and gives complementary information to MISR. The MISR and MODIS consensus pixels are more accurate than MISR or MODIS alone and fusing the data from MISR ELCMC and MODIS improves the average accuracy of polar cloud detection.

This is an example of truly interdisciplinary work where statistical machine learning merges statistics with computational sciences. This was an iterative process with feedbacks and inputs from MISR team at every step. The goal was to solve the scientific problem with the streaming data constraint, and an uncertainty measure was given based on posterior probability.

VALIDATION OF CLOUD PROPERTY MEASUREMENTS FROM MULTIPLE INSTRUMENTS

Jay Mace, University of Utah

The influence of clouds and precipitation on the radiation and hydrology of Earth depend fundamentally on the evolution of the vertical profile of microphysics. The understanding of this evolution is possible through observation-based models. The validity of the observations, however, is dependant on how well the vertical profile of microphysics is derived from remote sensing data in a statistically meaningful way. This talk demonstrates how to compare and validate microphysical cloud properties using multiple instruments including satellite measurements, ground-based remote sensing measurements, and aircraft measurements.

Cloud properties play an important role in the climate system. Reducing the uncertainties in climate models is necessary to improve our understanding of feedback mechanisms in the climate system, including feedback between the hydrologic cycle and cloud aerosol precipitation. The ability to infer cloud properties using remote sensing will aid in reducing these uncertainties in the climate models and allow us to gain a better understanding of the climate system and feedback mechanisms. There are multiple phases in the evolutions of clouds and cloud-size distributions. The interaction of clouds and their evolution occurs at the particle level, and therefore the particle size distribution is a critical component to the cloud problem. The evolution of clouds from aerosols to precipitation fundamentally occurs at the particle level and therefore, to understand feedback mechanisms, we must drill down to the particle size distribution (PSD). Statistical distributions are used to estimate PSD in climate models. Integrating across the PSD estimates how many aerosol particles are nucleating to create a cloud. This is one of the ill-defined quantities derived from remote sensing data.

The mass of a cloud is also of interest. Some clouds with very different properties (e.g., type, water content, ice content) can exhibit the same radar reflectivity. Remote sensing algorithms are required that can identify differences in extinction for different clouds with similar properties. A simple PSD characterization requires 3-4 parameters, which is the essence of the cloud problem. With satellite data, we seldom have more than one or two independent pieces of information to describe a given cloud volume or cloud column, which makes it fundamentally impossible to derive the details of the more complicated size distribution.

Many instruments are used in conjunction to measure clouds, including *in situ* aircraft, ground sites, and satellite sensor suites. *In situ* aircraft measurement within the cloud volume is the common tool used to provide "ground truth" against which satellite measurements are compared. Ground sites through the ARM program have upward-looking remote sensors that provide detailed profiles and have the advantage of working over long periods of time. However, these sites are constrained in coverage because they only measure one point. Satellite sensor suites move beyond single-instrument measurements, but the measurement profiles are limited, and they are very expensive.

Each of these techniques has associated uncertainties. For aircraft, uncertainties are associated with sample volume relevance and artifacts, such as those caused by the shattering of ice crystals on aircraft surfaces. Uncertainty is also associated with the algorithm used to determine cloud properties from remote sensing at ground sites. Like all algorithms, the algorithms used for cloud properties include key assumptions, such as the mass of ice crystal as a function of particle size. The retrieval technique also adds to the error. Algorithms incorporate assumptions about physical parameters, which contribute to the uncertainties in the algorithms in subtle ways.

An example of comparing ground-based and satellite data is seen in examining Moderate Resolution Imaging Spectroradiometer (MODIS) derived ice water paths. Ice water paths may be derived from a validated radar algorithm and compared with the MODIS data, showing reasonable agreement. However, one is a snapshot spatially averaged measurement and the other is a time average point measure. For cirrus clouds, a spatial average is generated by averaging the MODIS measurements over a rectangle that is oriented along the mean wind at the cloud level, and then averaging the ground-based radar data over a period of time when the cloud layer remains uniform. Comparing these measurements over a long period of time will aid in determining the error characteristics of the satellite. The uncertainty in the data, however, combined with the uncertainty in the science requires techniques to quantitatively assess these errors, and more sophisticated approaches would be helpful.

While techniques to quantitatively assess error in datasets exist, more sophisticated approaches would be helpful. A systematic approach to collecting and managing aircraft data to define empirical relationships is suggested. Combining multiple measurements in instrument suites and continuing and improving *in situ* measurements are critical for validation. Several elements are required to make progress in this area. Using single instruments is outdated, and a suite of sensors that provide multiple independent measurement profiles needs to be continued and improved upon. A long-term systematic global *in situ* measurement program is a critical adjunct to remote sensors.

UNCERTAINTY ISSUES ASSOCIATED WITH REMOTELY SENSED DATASETS FOR AEROSOLS

Lorraine Remer, National Aeronautics and Space Administration

A number of uncertainties are associated with deriving aerosols from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite. While many varieties of substances are considered aerosols, the diffuse and smooth aerosols are the most important when studying the climate system, as they reflect and absorb sunlight. It is difficult to distinguish between different types of aerosols from clouds in remote sensing retrievals. The satellite measures the aerosol optical depth (AOD), which are measures of the column-integrated extinction, the amount of light that is either scattered or absorbed as it passes through the aerosol layer giving an indication of the amount of aerosol. A satellite measures not just the radiation scattered from the aerosol layer, but it also collects some radiation that made it through the aerosol layer from the earth's surface. The surface effect needs to be removed from the satellite signal to estimate the extinction, which requires assumptions about the aerosol and the surface, leaving room for error. A more direct way of measuring AOD is to use a sunphotometer on the ground to measure the transmitted light directly. Combining measurements of the sunlight at the top of the atmosphere with the amount of sunlight at the surface, the extinction can be determined. This approach uses fewer assumptions and under the best conditions MODIS can retrieve AOD to within ±0.03, and a well-calibrated sunphotometer can measure it within ±0.01. The widespread network of sunphotometers called Aerosol Robotics Network (AERONET) retrieves data globally. The primary challenge with this technique is the mismatch between the spatially varying MODIS data and temporally varying sunphotometer data. There are select areas with coincident coverage in measurements between MODIS and AERONET. AERONET is a land-based network, and matching an overpass with an observation is rare. The location of the ground-based observation within the satellite grid square is a consideration in validation process.

Collocating AOD observations does not address uncertainty issues with other retrieved parameters like aerosol particle size. Collocating with sunphotometers takes advantage of the ground instruments' cloud-clearing algorithm. Comparing the spatial statistics with the temporal statistics increases the number of points that can be used from AERONET. With these results, information about the uncertainties of the remote measurements of AOD can be determined.

Another challenge is using the validated retrievals to generate daily and monthly mean maps of aerosol distributions that improve our understanding of climate. A complication arises when the weighting of the data is considered in the calculation of the monthly means. With few retrievals available for a particular grid square, there is a greater likelihood of high uncertainty. Different methods of weighting and averaging the data result in significantly different distribution maps and mean optical depths. Not all retrievals carry the same level of confidence. There are a number of options and methods to consider when analyzing the means from satellite and sunphotometer data, and finding a suitable solution is largely situational and dependent on the specific application under consideration.

SPATIAL STATISTICS WITH AN EMPHASIS ON AEROSOL DATA

Noel Cressie, Ohio State University

Statistical science is a paradigm that incorporates uncertainty into the description, explanation, and prediction of processes and parameters in the world. Spatio-temporal statistics incorporates space and time into statistical models of all aspects of the uncertainty. There are a multitude of statistical techniques that can be applied to climate datasets, and hierarchical statistical modeling is one such. The uncertainty can be modeled in a hierarchical manner where, at each level of uncertainty, a statistical model is used. There are at least two levels in the hierarchy, resulting in data models and process models, the latter being a representation of a physical process of interest. Fundamental to hierarchical statistical models is the use of conditional distributions. This results in a product form for the joint distribution; then Bayes' Theorem can be used to obtain the posterior distribution of unknown processes or parameters (e.g., aerosol optical depth in a modeling of data from the MISR instrument on NASA's Terra satellite).

Kriging is a spatial regression technique applied to datasets to filter out noise and fill gaps, and it can be shown to arise from a Gaussian-Gaussian hierarchical statistical model. Kriging methodology requires estimation of the spatial covariances in the dataset, from which a mathematical formula can be constructed to fill in the gaps in the dataset. This formula is simply the mean of the posterior distribution referred to earlier, with suitable substitution of estimates of unknown parameters. If we build a spatio-temporal model at the finest scale, even at a scale where there is a lack of data, and then aggregate up from that scale, the relationships will allow computation of any covariance at the scale of interest. This approach also allows one to determine the kriging standard error, which is a measure of the uncertainty in the smoothed, gap-filled product. Algorithms that can be used with the data, such as nearest-neighbor smoothing and inverse-distance weighting, do not allow the estimation of uncertainty that kriging allows.

One of the limitations of kriging is that it cannot be applied to large datasets because it does not scale well computationally. However, fixed rank kriging (FRK), a rather new statistical method (Cressie and Johannesson, 2006) using a spatial mixed-effects model, is a technique that can be applied to large, remote sensing datasets. This technique is computationally feasible and is able to deal with the dimension reduction that other spatial and temporal statistical methods have not been able to address. The "fixed" adjective refers to a fixed number of basis functions; as the number of bases grows, more information can be gleaned about the finer-scale variability. If the finer-scale variability is not present, then fewer basis functions can be used in this method. Importantly, the spatial covariances for FRK do not have to be stationary (Shi and Cressie, 2007; Cressie and Johannesson, 2008). The covariances determined by FRK can also be used in a cross-validation technique to estimate the uncertainty of the data that lie within a grid box.

AEROSOL AND CLOUD REPRESENTATION IN GLOBAL MODELS

Joyce Penner, University of Michigan

Global climate modelers continually work to improve climate models by analyzing observational data to gain better insights into the physics of atmospheric processes. The challenges associated with comparing climate models and data are similar to validation studies that compare satellite retrievals with ground-based measurements. Global climate models (GCMs) use a grid that is based on how long a period is to be represented by the simulation, typically 100 years, and how many times the model is run. Within each grid cell, the model physics and dynamics are represented and provide information about various attributes of the atmosphere, such as aerosol and cloud concentrations. The resolution, however, is too coarse to resolve clouds.

The coarse grid resolution requires parameterization of physical processes. Scientists must approximate many of the prognostic variables needed to draw conclusions about significant parts of the climate system, such as clouds and their effects on the radiation budget. Some of the processes that must be parameterized are convection, turbulence, radiation, and microphysics, all of which influence the complex cloud distributions that we observe, but which cannot be represented in detail in GCMs. The water mass of clouds, cloud fraction, and mixing processes, for example, have to be predicted and these predictions include uncertainties. There are many challenges associated with parameterization. In both the vertical and horizontal planes, the cloud cover can be thinner than what the resolution allows. When combining these two, it becomes necessary to make assumptions about the variation in the horizontal coverage in the vertical direction. Adding more complexity, clouds within one horizontal grid layer might have varying densities and vertical positions. The assumptions made about cloud distribution and radiation effects have consequences in the overall output of the model.

Climate models are validated by comparing the results to satellite data (as well as *in situ* data). In GCMs, clouds and the feedback from clouds, resulting from changes in the temperature, are one of the greatest sources of uncertainty. This is due to the lack of fine-scale grid models that can properly represent the sub-grid-scale features that are required to accurately predict climate changes.

Aerosol models represent various processes, such as emissions that form secondary aerosols, chemistry, aerosol microphysics, transport, and dry and wet deposition. The uncertainty associated with these models comes partly from variability in the sources of aerosols, given that there are multiple aerosol species. In addition, the representation of the other processes differs from model to model. Validation of the chemical composition of the aerosols in the models cannot come from comparisons with satellite data or ground-based remote sensing data like that of AERONET, because the measured aerosol optical depth (AOD) is a composite of the effects of the different aerosol types we attempt to model. So while measurements of the global average aerosol optical depth can improve, this will not completely resolve the differences between various models and observations. Convection also has large implications for the prediction of the vertical distribution of a particular aerosol. The impact of an aerosol such as black carbon varies depending on its vertical distribution. If the black carbon is located near the surface, it will tend to heat the surface whereas if it is primarily located in the free troposphere, it will act to cool the surface. The vertical distribution of the rest of the aerosols that make up the total AOD also affects their impact, since AOD can increase by aerosol water uptake and the uptake will generally be larger if the aerosol is located in the boundary layer where the relative humidity is generally higher than it is in the free troposphere. Thus the AOD and the effect of black carbon are dependent on how the vertical transport is treated. There is a significant difference in the vertical distribution of aerosols when aerosols from different GCMs are compared with each other. In addition there are significant differences in AOD between different versions of the same GCMs when different horizontal resolutions are compared.This is due to the changes in the predicted relative humidity in higher resolution GCMs.

The sources of uncertainty associated with aerosol models come from emissions, wet/dry removal, chemical production, and vertical transport. Scientists must address the biases and potential errors connected to grid resolution and also with the shortcomings in modeling some specific processes. Fortunately, it may be possible to use the extensive satellite data to address some of these issues and to make improvements in modeling the processes that are not well parameterized.

DATA ASSIMILATION AS A HIERARCHICAL STATISTICAL PROCESS, INTERACTING DYNAMICALLY WITH MODELING

Christopher Wikle, University of Missouri

A Bayesian approach to climate models has an advantage by addressing processes that are non-Gaussian and nonlinear (i.e., geophysical processes in the climate system). From a statistical perspective, data assimilation is combining data with prior knowledge from sources such as mathematical models, other datasets, expert opinion, and others to gain an understanding of the true state of the system. A Bayesian perspective to retrospective data assimilation, combining information to create datasets for use in climate model initial conditions, is explored in this talk. For example, a wind dataset that can be replicated over time and space has uncertainty associated with the satellite observations and these uncertainties influence the weighted combination of the prior mean and the mean data as well as the outcome of the posterior data distributions. In order to minimize these uncertainties, physical properties of the system need to be improved and applied to the statistical representation. The fundamental challenges in data assimilation techniques include model complexity, model uncertainty, state process dimensionality, and data volume, which are all interrelated.

The Bayesian hierarchical model methodology uses building blocks by separating the data and process variables (e.g., temperature, wind) in the model from the model parameters, which are quantities introduced in the model development, such as measurement error, to better understand the associated uncertainties. A critical question is quantifying the uncertainty in the process model. A common solution is to include an additive error term, but this does not represent all model uncertainty. Another option is to treat the model output as data to be used in a statistical model, however, in this approach the underlying dynamics of interacting state processes are ignored. A spectrum of models between deterministic and stochastic is the most useful and should be dependant on the goal of the analysis, the type of data, and the type of prior information that is available in the study. The advantage is that the models can accommodate more complex dynamics, however the interpretability of the parameters suffers.

Data volume is increasingly becoming problematic for better understanding the physical processes that are interacting with the climate system and for quantifying the uncertainty in remotely sensed datasets. Datasets are large and, at times, of poor quality and may include unknown biases and errors. This challenge requires the statistics and earth sciences communities to work together in this era of large datasets. Given that earth scientists and statisticians could learn a great deal from each other, it is unfortunate that they face poor communication across the two disciplines. Finally, funding is not adequate to support the type of collaboration that is needed to foster growth and improvement in these disciplines.

THE PRACTICAL AND INSTITUTIONAL BARRIERS FOR MAKING PROGRESS ON DEVELOPING AND IMPROVING STATISTICAL TECHNIQUES FOR PROCESSING, VALIDATING, AND ANALYZING REMOTELY SENSED CLIMATE DATA

Doug Nychka, National Center for Atmospheric Research

A few significant obstacles that are most commonly felt by statisticians and geoscientists who deal with processing, validating, and analyzing remotely sensed climate data include resistance to new ideas by members of the community, lack of funding for analysis following a satellite launch, and the need for more work on combining models with observations through data assimilation. This talk addresses these barriers from case studies in the community.

A colleague working in the field of GPS meteorology noted that there are opportunities for statisticians and geoscientists to work together toward a common objective, but that community resistance can be problematic. As an example, an atmospheric sounding project encountered strong community resistance which was not overcome until personal bridge-building was carried out between people who wanted to implement various outputs from the proposed data analysis and people that were able to build and launch the satellite. In this circumstance, the barrier was lifted and progress was made.

A second colleague, who works with the Measurements in Pollution of the Troposphere (MOPITT) instrument, raised a concern about the difficulty of obtaining sustained data-analysis efforts. The analysis work must be re-justified every few years after the initial funding period. Although there can be funding available to build and launch a new satellite-based instrument, there is less support to analyze the resultant data or to integrate one set of satellite retrevials with other sets of observations and data products. In addition, the expertise of those in the field may not be used to the maximum extent possible when it comes to analysis, perhaps because of the shortage of funding for an analysis study. One way to improve the chances of obtaining funding could be a virtual simulation facility that would allow scientists to test new instrument designs with the conditions represented by various global atmospheric/ocean models and with the kinds of statistical methodology that would be used for retrievals. The most ambitious evaluation would be using data assimilation to combine a numerical model with the synthetic satellite measurements and other more standard geophysical data to assess the added benefit of the new instrument. This would be an efficient way to verify the importance of new instruments because it would account for the data analysis of the measurements and their coincidence or uniqueness with other types of data.

A third scientist highlighted both opportunities and barriers that scientists face when they attempt to reconcile satellite data and model results. It is well known that climate is a long-term average of complicated geophysical processes. However, results from satellite data are on a much shorter time scale and are not suitable for directly determining climatological averages. When these data are compared to model results over a specific time period, the atmospheric component of a climate model is essentially being used to forecast weather. The use of weather forecasting techniques such as data assimilation to improve geophysical models is an emerging interdisciplinary approach that falls outside of traditional methods in climate science. This presents a conceptual challenge to scientists because the short-time-scale process information must be reconciled with the performance of the model in simulating long-term climate. Also they must learn how to use a new and more complicated statistical tool. The ensemble data assimilation methods used in this case study also allow for characterizing the statistical uncertainty of the analysis and so add another layer to the interpretation of the scientific results. This can also be seen as an opportunity to make significant improvements to the models and the physics behind them. Using data assimilation in this way creates an opportunity to use multiple instruments and parameters from remotely sensed data to improve upon model physics and dynamics. It essentially blurs the line between climate and weather models, which can be a beneficial way to improve both.

Appendix C

Planning Committee and Rapporteur Biographies

Amy Braverman (Chair) is a scientist at the Jet Propulsion Laboratory. Dr. Braverman received a PhD degree in statistics from the University of California at Los Angeles (UCLA) in 1999, the M.A. degree in mathematics from UCLA in 1992, and the B.A. degree in economics from Swarthmore College in 1982. She is a MISR co-investigator responsible for Level 3 algorithm development. Her research priorities are data reduction and analysis of massive datasets, data mining, and high-dimensional data visualization. She is also the Level 3 Scientist for MISR and the Atmospheric Infrared Sounder (AIRS). She is a member of the National Research Council's Committee on Applied and Theoretical Statistics.

Philip E. Ardanuy is chief scientist and director for Remote Sensing Applications at Raytheon Information Solutions. Dr. Ardanuy has 31 years of professional experience participating in NOAA, NASA, NSF, and DoD environmental applications programs. He specializes in developing integrated mission concepts through government-industry-academic partnerships. Dr. Ardanuy's research and development career extends across net-centric and system-of-systems concepts, telepresence-telescience-telerobotics, tropical meteorology, the Earth's radiation budget and climate, satellite instrument calibration and characterization, remote sensing applications and systems engineering, scientific applications research-to-operational transition, and validation of environmental observations. Dr. Ardanuy has authored over 50 publications on environmental and weather monitoring and modeling; sustainable exploration; utilization of operational environmental satellite data; environmental sensing; and applications development and refinement approaches. He participates in numerous public service activities, including serving as associate editor for the International Society for Optical Engineering (SPIE) *Journal of Applied Remote Sensing*, chair of the AMS Committee on Satellite Meteorology and Oceanography, and organizing and science planning committees for AMS and SPIE. Dr. Ardanuy is currently a member of the NRC Committee on a Strategy to Mitigate the Impact of Sensor De-scopes and Demanifests on the NPOESS and GOES-R Spacecraft. He previously served on the NRC Committee on Environmental Satellite Data Utilization and on two panels of the Committee on Earth Science and Applications from Space: Panel on Earth Science Applications and Societal Objectives and Panel on Options to Ensure the Climate Record from the NPOESS and GOES-R Spacecraft.

John J. Bates is Chief of the Remote Sensing Applications Division of the U.S. National Oceanic and Atmospheric Administration's National Climatic Data Center. Dr. Bates received a PhD in meteorology from the University of Wisconsin-Madison on the topic of satellite remote sensing of air-sea heat fluxes. He joined the NOAA Environmental Research Laboratories in Boulder, CO, in 1988 and there continued his work in applying remotely sensed data to climate applications. In 2002, Dr. Bates moved to the NOAA National Climatic Data Center in Asheville, NC. Dr. Bates's research interests are in the use of operational and research satellite data and weather radar data to study the global water cycle and studying interactions of the ocean and atmosphere. He has authored over 25 peer-reviewed journal articles on these subjects. He served on the American Meteorological Society's (AMS) Committee on Interaction of the Sea and Atmosphere (1987-1990) and the AMS Committee on Applied Radiation (1991-1994). As a member of the National Research Council's Global Energy and Water Cycle Experiment (GEWEX) panel (1993-1997), Dr. Bates reviewed U.S. agency participation and plans for observing the global water cycle. He has also been a contributing author and U.S. government reviewer of the Intergovernmental Panel on Climate Change assessment reports. He currently serves on the International GEWEX Radiation Panel, whose goal is to bring together theoretical and experimental insights into the radiative interactions and climate feedbacks associated with cloud processes, including the effects of water vapor within the atmosphere and at the Earth's surface.

James A. Coakley, Jr. is a professor in the Department of Atmospheric Sciences, College of Oceanic and Atmospheric Sciences at Oregon State University (OSU). His research concentrates on the Earth's energy budget and climate change, with particular interest in the role clouds play in the climate system. Prior to joining OSU, he led the National Center for Atmospheric Research's Cloud-Climate Interaction Group and Satellite Data Analysis Group. He is a member of the NOAA Council on Long-term Climate Monitoring, NASA's Global Aerosol Climatology Project Science Team, NASA's PICASSO-CENA Science Team, the Science Steering Committee of the Scripps Institution of Oceanography's Center for Clouds, Chemistry and Climate, and the NASA CERES Science Team. He formerly served on the NRC Committee on Meteorological Analysis, Prediction, and Research, and currently serves on the NRC's Climate Research Committee.

Karen Kafadar is a professor in the Department of Statistics at Indiana University. Her research focuses on robust methods, data analysis, and characterization of uncertainty in the physical, chemical, biological, and engineering sciences. Her previous appointments include National Institute of Standards and Technology (NIST), Hewlett-Packard, and the National Cancer Institute. Until 2008, she was a professor at the University of Colorado-Denver, where she directed the Statistical Consulting Service, collaborated with researchers in the School of Medicine, and taught courses in applied and theoretical statistics. She is currently serving as chair of the NRC's Committee on Applied and Theoretical Statistics and as a member of the NRC's Board on Mathematical Sciences and their Applications. She has served as Editor or Associate Editor on several editorial review boards and on the governing boards of the American Statistical Association (ASA), the Institute of Mathematical Statistics, and the International Statistical Institute. Dr. Kafadar is a fellow of the ASA and has authored over 80 journal articles and book chapters. She received her B.S. in mathematics and M.S. in statistics from Stanford University in 1975 and received her PhD in statistics from Princeton University in 1979.

Douglas Nychka is a senior scientist at the National Center for Atmospheric Research (NCAR). Before joining NCAR, he spent 14 years as a faculty member in the Statistics Department at North Carolina State University. In his current role, his primary challenge is interdisciplinary research and migrating statistical techniques to important scientific problems and using these problems to motivate novel statistical research. His personal research interests include nonparametric regression, statistical computing, spatial statistics, and spatial designs. Dr. Nychka is a former member of the NRC's Committee on Applied and Theoretical Statistics. He received his PhD from the University of Wisconsin.

Joyce E. Penner is a professor in the Department of Atmospheric, Oceanic, and Space Sciences, and director of the Laboratory for Atmospheric Science and Environmental Research at the University of Michigan. Dr. Penner's research focuses on improving climate models through the addition of interactive chemistry and the description of aerosols and their direct and indirect effects on the radiation balance in climate models. She is also interested in urban, regional, and global tropospheric chemistry and budgets, cloud and aerosol interactions, and cloud microphysics, climate and climate change, and model development and interpretation. Dr. Penner has been a member of numerous advisory committees related to atmospheric chemistry, global change, and Earth science. She was most recently a member of the NRC Committee on Metrics for Global Change Research. She formerly served on the NRC's Committee on Research Priorities for Airborne Particulate Matter (1997-2004), the Committee on Geophysical and Environmental Data (1995-2000), and the Committee for Review of the Science Implementation Plan of the NASA Office of Earth Science (2000). Dr. Penner is a fellow of the American Geophysical Union and currently serves on the NRC's Climate Research Committee.

Steven Platnick is a TC4 platform scientist for the ER-2 aircraft. His research includes theoretical and experimental studies of satellite, aircraft, and ground-based cloud remote sensing. He has been involved extensively with remote sensing field studies, including use of the MODIS Airborne Simulator instrument flown on the NASA ER-2. Dr. Platnick is the Deputy Project Scientist for NASA's Aqua spacecraft and a member of the MODIS Science Team working on operational cloud optical and microphysical products. He is also a member of the CloudSat Science Team. His collaborations with NASA's Goddard Space Flight Center began in 1993, and he is currently a member of the Climate and Radiation Branch in the Laboratory for Atmospheres. Previously, Dr. Platnick was a research associate professor in the Joint Center for Earth Systems Technology, University of Maryland Baltimore County. He received a PhD in atmospheric sciences from the University of Arizona.

Staff

Martha McConnell is a program officer for the Board on Atmospheric Science and Climate (BASC) and the Polar Research Board (PRB) at the National Research Council. Prior to joining the NRC in 2008, she was a congressional fellow working for Senator Lautenberg (NJ) working on ocean and climate issues with a focus on ocean acidification legislation. Martha has also worked for the Sea Education Association, Inc., and the

Support Office for Aerogeophysical Research (SOAR). She holds a BA in geology from Colgate University and a MS and PhD in paleoceanography/paleoclimatology from the University of South Carolina. Her research interests include calibration of climate proxies and the application of those proxies to the sediment record in order to identify the timing and magnitude of climate variability in the past.

Scott T. Weidman is the director of the National Research Council's Board on Mathematical Sciences and Their Applications (BMSA). He joined the NRC in 1989 with the Board on Mathematical Sciences and moved to the Board on Chemical Sciences and Technology in 1992. In 1996 he established a new board to conduct annual peer reviews of the Army Research Laboratory, which conducts a broad array of science, engineering, and human factors research and analysis, and he later directed a similar board that reviews the National Institute of Standards and Technology. He has been full-time with the BMSA since June of 2004. During his NRC career, he has staffed studies on a wide variety of topics related to mathematical, chemical, and materials sciences, laboratory assessment, and science and technology policy. His current focus is on building up the NRC's capabilities and portfolio related to all areas of analysis and computational science. He holds bachelor degrees in mathematics and materials science from Northwestern University and M.S. and PhD degrees in applied mathematics from the University of Virginia. Prior to joining the NRC, he had positions with General Electric, General Accident Insurance Company, Exxon Research and Engineering, and MRJ, Inc.

Lauren Brown is a Research Associate and former Christine Mirzayan Science and Technology Policy Fellow with the Polar Research Board and the Board on Atmospheric Sciences and Climate. She is currently completing her MS Degree in marine studies with a concentration in physical ocean science and engineering at the University of Delaware. Her research involves the analysis of tidal currents, velocity structure and ocean physics off the coast of northwestern Greenland to determine the influence on the larger regional dynamics. She holds a Bachelor of Arts from the University of Delaware in physics and astronomy.