# CLIMATE DATA RECORDS
## FROM ENVIRONMENTAL SATELLITES

Committee on Climate Data Records from NOAA Operational Satellites
Board on Atmospheric Sciences and Climate
Division on Earth and Life Studies

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 Contract No. 50-DGNA-1-90024 between the National Academy of Sciences and the National Oceanic and Atmospheric Administration. The views expressed herein are those of the author(s) and do not necessarily reflect the views of NOAA or any of its subagencies.

International Standard Book Number 0-309-09168-3 (Book)
International Standard Book Number 0-309-53080-6 (PDF)

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.

*Cover:* “Cirrus at Sunset” by G. Stephens, 2003. The original is an oil on canvas and measures 20 in × 24 in. This painting is part of the “Noble Clouds Under Variable Light” series.

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. Bruce M. Alberts 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. Wm. A. Wulf 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. Bruce M. Alberts and Dr. Wm. A. Wulf are chair and vice chair, respectively, of the National Research Council.

**www.national-academies.org**

## COMMITTEE ON CLIMATE DATA RECORDS
## FROM NOAA OPERATIONAL SATELLITES

**DAVID ROBINSON** *(Chair),* Rutgers University, Piscataway, New Jersey
**ROGER BARRY**, World Data Center for Glaciology, National Snow & Ice Data Center, Boulder, Colorado
**JANET CAMPBELL**, University of New Hampshire, Durham
**RUTH DEFRIES**, University of Maryland, College Park
**WILLIAM J. EMERY**, University of Colorado, Boulder
**MILTON HALEM**, NASA Goddard Space Flight Center (ret.), Greenbelt, Maryland
**JAMES HURRELL**, National Center for Atmospheric Research, Boulder, Colorado
**ARLENE LAING**, University of South Florida, Tampa
**ROBERTA BALSTAD MILLER**, Columbia University/CIESIN, Palisades, New York
**RANGA MYNENI**, Boston University, Massachusetts
**RICHARD SOMERVILLE**, Scripps Institution of Oceanography, La Jolla, California
**PAUL D. TRY**, Science and Technology Corporation, Silver Spring, Maryland
**THOMAS VONDER HAAR**, Colorado State University, Fort Collins

### NRC Staff

**SHELDON DROBOT**, Study Director
**ROB GREENWAY**, Project Assistant

## BOARD ON ATMOSPHERIC SCIENCES AND CLIMATE

### Ex Officio Members

### NRC Staff

# Preface

For the past four decades, data flowing from satellite-borne sensors have provided environmental information at spatial scales only dreamed of before the advent of these powerful observation tools. Data from satellites provided remarkable insights into Earth's land, atmosphere, oceans, and cryosphere systems. We have progressed in understanding Earth's internal dynamics and kinematics, along with important interrelationships between Earth systems. Time series data of elements within these systems have been scrutinized in attempts to better understand climate variability and to identify critical trends that may signal changes in the climate system. From these studies has emerged a growing appreciation of the importance of satellite climate data records (CDRs) that possess the accuracy, longevity, and stability to facilitate credible climate monitoring. These satellite CDRs provide abundant information to assist those making decisions regarding the status and fate of our environment.

The National Oceanic and Atmospheric Administration (NOAA) is to be commended for accepting the challenge to better understand climate variability and change. By requesting the formation of this ad hoc National Research Council (NRC) committee, it recognized the importance of generating and maintaining satellite climate data records in order to meet this mandate. This committee was tasked with assisting NOAA as it designs a plan to establish this agency as the chief steward of satellite CDRs. This task involves two phases. In phase one, NOAA requested an interim report on a range of different approaches and strategies for generating CDRs and identified key attributes common to successful CDR generation programs. NOAA will use this information as a guide in developing a plan to create CDRs from polar-orbiting satellites. In phase 2 (expected in late 2004), the committee will provide specific comments on the plan.

The NRC's Committee on Climate Data Records from Operational Satellites took a number of steps to conduct its analysis. Following a series of committee teleconferences, an information gathering workshop was held in August 2003, with several dozen scientists providing valuable input (see Appendix A for a list of participants). A questionnaire was also distributed to conference participants and others, followed by a busy autumn of teleconferences, e-mails, and face-to-face meetings in Washington, D.C., and Boulder, Colorado. It is a credit to the committee and those assisting us at the NRC that by mid-December this report was ready to go out for review.

Our report is divided into six chapters. In Chapter 1 we present a definition of a CDR and introduce the concepts of "fundamental climate data records" (FCDRs) and "thematic climate data records" (TCDRs), distinctions that are of utmost importance when designing and implementing a satellite CDR program. In Chapter 2 we discuss lessons learned from a sampling of past and present efforts to create satellite CDRs. This chapter benefits tremendously from the thematic expertise of all committee members. Elements of a successful satellite CDR generation program are outlined in Chapter 3, beginning with an organizational structure, continuing with suggested steps for creating CDRs, and finishing with suggestions on sustaining the program. A critical element to any CDR program is data management. In Chapter 4 we discuss data storage, archiving, and dissemination issues, emphasizing that the success of the satellite CDR program requires facilitating the straightforward and open access of subsets of satellite and ancillary data of interest to an investigator. NOAA is well suited to assume key stewardship of satellite CDRs, but it cannot and should not go it alone. In Chapter 5 the importance of partnering with other federal agencies, the international community, academia, and other sectors is discussed. Chapter 6 presents an overarching recommendation, along with a series of supporting recommendations.

Many individuals provided important information and insights that helped the committee as we prepared this report. Thanks go to Greg Withee, Tom Karl, Mitch Goldberg, John Bates, and George Ohring for their interest in and leadership of satellite CDR development efforts at NOAA and for presenting us with such an exciting and challenging task. We are grateful to all who took time from their busy summer schedules to participate in the August 2003 workshop, particularly those who made presentations: Eugenia Kalnay, Kevin Trenberth, Graeme Stephens, and Bill Rossow. We also appreciate all those who contributed to earlier NRC reports that illustrate and justify the importance of climate data records.

On behalf of the entire committee I want to express gratitude to those associated with the NRC Board on Atmospheric Sciences and Climate who provided keen insights, able direction, and tremendous support to our endeavor. This includes board director, Chris Elfring; project assistant, Rob Greenway; and especially our erudite study director, Sheldon Drobot.

A word of thanks to members of the committee; they are wonderful, talented individuals who volunteered countless hours to this effort. This reflects their dedication to the science community, and illustrates their belief that by having the opportunity to help guide NOAA in the detailed development of an end-to-end CDR program, they can make a difference.

Our committee's work is not yet done. Part of the attraction of serving on this committee is that we have two opportunities to produce advice: this interim report and a chance to comment in detail later this year on the satellite CDR plan that NOAA will now formulate. We look forward to getting back together in the middle of 2004, at which time we anticipate producing a second report in response to NOAA's draft plan.

David A. Robinson, *Chair*
Committee on Climate Data Records from
NOAA Operational Satellites

# Acknowledgments

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 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 report:

Lee E. Branscome, Climatological Consulting Corporation
James A. Coakley, Jr., Oregon State University
John Christy, University of Alabama, Huntsville
Jennifer Francis, Rutgers University
Dorothy Hall, National Aeronautics and Space Administration
Michael King, National Aeronautics and Space Administration
Kevin Trenberth, National Center for Atmospheric Research
Steve Worley, National Center for Atmospheric Research

Although the reviewers listed above have provided constructive comments and suggestions, they were not asked to endorse the report's conclusions or recommendations, nor did they see the final draft of the report before its release. The review of this report was overseen by Christopher Justice of the University of Maryland. Appointed by the National Research Council, he was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

# Contents

APPENDIXES

# Executive Summary

At the dawn of the twenty-first century, NOAA's mission includes a bold new mandate to "understand climate variability and change to enhance society's ability to plan and respond." An integral component of NOAA's emphasis on climate involves creating a stewardship plan to generate, analyze, and archive long-term satellite climate data records (CDRs) for assessing the state of the environment. Although the concept of a "climate data record" has surfaced numerous times in recent literature (e.g., NRC, 2000c,e), the climate community has yet to settle on a consistent definition. In this report the committee defines a climate data record as **a time series of measurements of sufficient length, consistency, and continuity to determine climate variability and change.** We further segment satellite-based CDRs into fundamental CDRs (FCDRs), which are calibrated and quality-controlled sensor data that have been improved over time, and thematic CDRs (TCDRs), which are geophysical variables derived from the FCDRs, such as sea surface temperature and cloud fraction.

To generate the best possible plan for creating satellite CDRs, NOAA asked the National Academies to conduct a two-phase study to provide advice on creating CDRs. In phase 1, the committee is providing an interim report with advice on the key elements of a satellite-based CDR program, including lessons learned from previous attempts, important considerations for identifying an appropriate organizational framework for long-term success and sustainability, suggested steps for generating and archiving CDRs, and the importance of partnerships. The objective of the interim report is to provide NOAA with general guidance about what needs to be included in its plan. More specific comments will be provided once NOAA writes the plan, expected to be completed in late summer of 2004.

NOAA's new climate mandate is fundamentally different from its traditional weather forecasting mandate and raises a new set of challenges owing to the varied uses of climate data, the complexities of data generation, and the difficulties in sustaining the program indefinitely. The task and structures being proposed for NOAA in this report are considerably more complex, costly, and demanding than those currently in place. A high level of commitment and a number of changes at multiple levels within the agency will be needed to institute and fund the various components of CDR stewardship. NOAA will not, however, be the first entity to generate climate-quality data and NOAA can learn many lessons from previous efforts; looking back on historical programs, some commonalities for success include science advisory panels, regular calibration and validation of data, adequate resources for reprocessing, user workshops to solicit advice on the future of the program, clear data storage and dissemination policies, and a willingness to form partnerships. Based on these historical lessons, community surveys, a workshop, and committee expertise, the committee identified 14 key elements for creating a climate data record program based mainly on satellites (Box ES-1). Adherence to these elements would help NOAA to create CDRs that are accepted as community standards, while ensuring that they remain responsive to user needs.

Underlying many of these elements of success is early attention to data stewardship, management, access and dissemination policies, and the actual practices implemented. Because a successful CDR program will ultimately require reprocessing, datasets and information used in their creation, such as metadata, should be preserved indefinitely in formats that promote easy access. The ultimate legacy of long-term CDR programs is the data left to the next generation, and the cost of data management and archiving must be considered as an integral part of every CDR program.

The new emphasis and importance of climate within NOAA's mission requires an increased focus on partnerships and new approaches as it relates to supporting extramural research. Many agencies and groups are interested and involved in creating, analyzing, and storing CDRs. By partnering with other government agencies, academia, and the private sector in development, analysis, and reprocessing of CDRs, NOAA can create and sustain a successful CDR effort; a high degree of interagency coordination on the requirements, definition, and implementation of CDRs is essential for satisfying the broad user communities of today and providing climate data stewardship for future generations.

**OVERARCHING RECOMMENDATION: NOAA should embrace its new mandate to understand climate variability and change by asserting national leadership for satellite-based climate data record generation,**

## BOX ES-1 KEY ELEMENTS OF SUCCESSFUL CLIMATE DATA RECORD GENERATION PROGRAMS

### *CDR Organizational Elements*

1. A high-level leadership council within NOAA is needed to oversee the process of creating climate data records from satellite data.
2. An advisory council is needed to provide input to the process on behalf of the climate research community and other stakeholders.
3. Each fundamental CDR (FCDR) should be created by a specifically appointed team of CDR experts.
4. Science teams should be formed within broad disciplinary theme areas to prescribe algorithms for the thematic CDRs (TCDRs) and oversee their generation.

### *CDR Generation Elements*

5. FCDRs must be generated with the highest possible accuracy and stability.
6. Sensors must be thoroughly characterized before and after launch, and their performance should be continuously monitored throughout their lifetime.
7. Sensors should be thoroughly calibrated, including nominal calibration of sensors in orbit, vicarious calibration with in situ data, and satellite-to-satellite cross-calibration.
8. TCDRs should be selected based on well-defined criteria established by the Advisory Council.
9. A mechanism should be established whereby scientists, decision makers, and other stakeholders can propose TCDRs and provide feedback that is considered in the selection of TCDRs.
10. Validated TCDRs must have well-defined levels of uncertainty.
11. An ongoing program of correlative in situ measurements is required to validate TCDRs.

### *Sustaining CDR Elements*

12. Resources should be made available for reprocessing the CDRs as new information and improved algorithms are available, while also maintaining the forward processing of data in near real time.
13. Provisions should be included to receive feedback from the scientific community.
14. A long-term commitment of resources should be made to the generation and archival of CDRs and associated documentation and metadata.

**applying new approaches to generate and manage satellite climate data records, developing new community relationships, and ensuring long-term consistency and continuity for a satellite climate data record generation program.**

NOAA is recognized as a national leader in weather information, including the management of a weather satellite program and creation of weather products. However, success in establishing and sustaining a CDR program requires a long-term commitment and a level of effort that goes beyond NOAA's weather program. A particularly key component of NOAA's success will be defining steps for creating FCDRs and TCDRs. NOAA's plan also needs to account for all of the data and metadata that must be stored in easily accessible, self-describing formats. Fortunately, NOAA should not feel obligated to generate all of the nation's CDRs, and by enhancing and expanding community involvement in the CDR program, NOAA can help to ensure community acceptance and creation of high-quality CDRs.

**Supporting Recommendation 1: NOAA should utilize an organizational structure where a high-level leadership council within NOAA receives advice from an advisory council that provides input to the process on behalf of the climate research community and other stakeholders. The advisory council should be supported by instrument and science teams responsible for overseeing the generation of climate data records.**

An important step in maintaining a successful program is developing or utilizing an appropriate organizational framework that incorporates feedback and advice from user communities. The committee believes that NOAA will help to ensure success if it includes scientists interested in CDRs, assigns committed people to generate the CDRs, develops technical and science support for users, and creates science teams that are renewed regularly. In particular, NOAA should utilize an advisory council of internationally recognized climate experts to:

1. Recommend and prioritize the variables that are developed into TCDRs;
2. Oversee the calibration of FCDRs and validation of TCDRs;
3. Evaluate proposed new TCDRs as measurement capabilities improve or scientific insights change over time;
4. Review the utility and acceptance of TCDRs and recommend the elimination of those that are not successful; and
5. Review and oversee NOAA's stewardship of the CDR program.

The actual creation of FCDRs should be carried out by a team of engineers and scientists, who should monitor satellite characteristics and document their work extensively so that future generations can assess and understand their work. Additionally, TCDR science teams with broad interdisciplinary representation should define algorithms for TCDR development and oversee TCDR generation. These teams should include research scientists funded by or employed by NOAA and scientists from other agencies, academia, or private industry who use the data, and they should be competitively selected, with limited (but renewable) terms.

> **Supporting Recommendation 2: NOAA should base its satellite-based climate data record generation program on lessons learned from previous attempts, which point out several unique characteristics of satellite climate data records, including the need for continuing calibration, validation, and algorithm refinements, all leading to periodic reprocessing and reanalysis to improve error quantification and reduce uncertainties.**

Because most of NOAA's operational satellites were created as weather rather than climate platforms, the committee stresses that NOAA should include nominal calibration, vicarious calibration monitoring, and satellite-to-satellite cross-calibration as part of the operational satellite system; this is important because orbital drift, sensor degradation, and instrument biases will affect the creation of consistent CDRs. Nominal calibration involves determining the calibration of a single sensor on a single platform, and while this is standard prelaunch practice, it is important to calibrate the sensor in orbit as well. Vicarious calibration monitoring involves measuring a known target or comparing the satellite signal with simultaneous in situ, balloon, radiosonde, or aircraft measurements; all instruments should undergo vicarious calibration monitoring at regular intervals, regardless of on-board nominal calibration, to prevent drifting of the data over time due to orbital drift and drift in the observation time, which aliases the diurnal cycle onto the record. Satellite-to-satellite cross-calibration involves adjusting several same-generation instruments to a common baseline, and this is particularly important for long-term studies, as each sensor will have slightly different baselines even if they are built to the same specifications.

An ongoing program of validation also should be carried out to determine the uncertainty associated with TCDRs. This is based on establishing rigorously derived uncertainties for the TCDR using independent correlative measurements conducted throughout the data record and over global scales, which in turn determines whether a trend can be detected. NOAA should establish a two-track generation program, including an upgradeable baseline CDR track and a second (mostly extramural) funded research program to validate, analyze, assess, and reduce uncertainties in future base versions. The two-track approach encourages a culture where scientists and users know that future improvements will be available over time.

**Supporting Recommendation 3: NOAA should define satellite climate data record stewardship policies and procedures to ensure that data records and documentation are inexpensive and easily accessible for the current generation and permanently preserved for future generations.**

History reveals that programs are more successful when the data management system provides free and open access to data, facilitates the reprocessing of CDRs, allows for new satellite TCDRs to be created, and has an easy problem-reporting procedure. A clear data policy can ensure continuity in the data record, including the ancillary data used to reprocess CDRs, project and dataset documentation, and the science production software. NOAA also should ensure that the data management infrastructure can accommodate user requests and provide different data formats, given the large satellite data volumes that a CDR program will create. This system should include the capability for temporal searches and subsetting. NOAA also can ensure a more robust program if the data are available in self-describing formats appropriate for a variety of uses, including geospatial and socio-economic applications. NOAA should establish a process for scientifically assessing the long-term potential of data and data products. Scientific assessments of the data can help NOAA to organize its archive so that data dissemination is efficient and cost-effective.

**Supporting Recommendation 4: NOAA should develop new community relationships by engaging a broader academic community, other government agencies, and the private sector in the development and continuing stewardship of satellite climate data records.**

One of the best methods NOAA can institute for gathering community input is to convene regular open science meetings where users share their research and discuss limitations and recommendations for improving the CDRs. It is important to hold these meetings regularly because research will improve data quality over time and the meetings will help to foster community support. These meetings could be held in conjunction with conferences held by such organizations as the American Meteorological Society or the American Geophysical Union, with benefits being cost savings and broader attendance. NOAA should actively encourage other agencies and user communities to assist in development, analysis, and reprocessing of CDRs because expertise for CDRs lies within many sectors. NOAA can create a more successful CDR program by developing these partnerships.

**Supporting Recommendation 5: NOAA should consider existing U.S. multi-agency organizations for implementation of the climate data record program, rather than devising a new structure. The most appropriate organization is the Climate Change Science Program.**

Stewardship of CDRs is complex, costly, and demanding, and NOAA should aggressively seek partnerships to help to ensure a successful program. The committee does not believe that NOAA needs to invent and implement a new management structure for generating, analyzing, and archiving CDRs; for instance, the goals and management structure of the Climate Change Science Program (CCSP) are similar to NOAA's climate goals, and NOAA may therefore be able to implement part of the CDR program under the CCSP. If NOAA were to volunteer to be the lead or executive agency (or delegate leadership to a partner) responsible for satellite CDRs under the CCSP umbrella, NOAA could advance its climate mandate and assert national leadership. Because the CCSP structure already has built-in interagency interactions, NOAA could also leverage them for the CDR program.

**Supporting Recommendation 6: NOAA should pursue appropriate financial and human resources to sustain a multidecadal program focused on satellite climate data records.**

Developing a CDR program is fundamentally important to the nation, and it is imperative that the effort not be inhibited by a lack of human or financial resources. Even if NOAA leverages funds and personnel from other agencies, academia, and private industry, and even if it integrates the CDR program into CCSP, it will still have to be aggressive in seeking additional funds. This program will require a long-term vision and commitment, and it will be important to account for inflationary increases when outlining the human and infrastructure needs for successfully generating, analyzing, reprocessing, storing, and disseminating CDRs.

# 1

# Introduction

In the 1950s, while still a postdoctoral fellow, Charles David Keeling designed and built the first highly accurate instrument to measure atmospheric carbon dioxide ($CO_2$) concentrations. In 1958, as part of the International Geophysical Year, he began measurements at the Mauna Loa Observatory in Hawaii, a site where $CO_2$ levels were thought to be characteristic of the unpolluted global atmosphere. These measurements are continued today under international agreements at several locations and by many investigators.

The graph of rising $CO_2$ concentration from 1958 to the present (Figure 1-1) is now known as the "Keeling curve." This time series illustrates the qualities of an outstanding climate data record (CDR). Keeling insisted from the start on impeccable quality control. In addition to revealing the increase in carbon dioxide caused by human activities, the exceptionally high accuracy of the measurements has made possible many investigations into the carbon cycle. The importance of these fundamental observations was not always widely recognized (Keeling, 1998), especially early on, and so the ultimate success of this early CDR illustrated the value of carefully planned, long-term commitments to data collection and analysis.

The sustained effort to maintain the atmospheric carbon dioxide record is valuable for its implications as a paradigm for CDR development. Today the Keeling curve, documenting the power of human beings to alter the chemical composition of the entire atmosphere, has iconic status as the single discovery most responsible for motivating research on anthropogenic climate change.

Unfortunately, the $CO_2$ record is an atypical CDR: many climate records are deficient in length, stability, or accuracy. The ability to understand, predict, and adapt to climate change and variability, however, necessitates high-quality, long-term, and stable measurements of Earth's environment. As noted by the National Research Council (NRC, 2001),

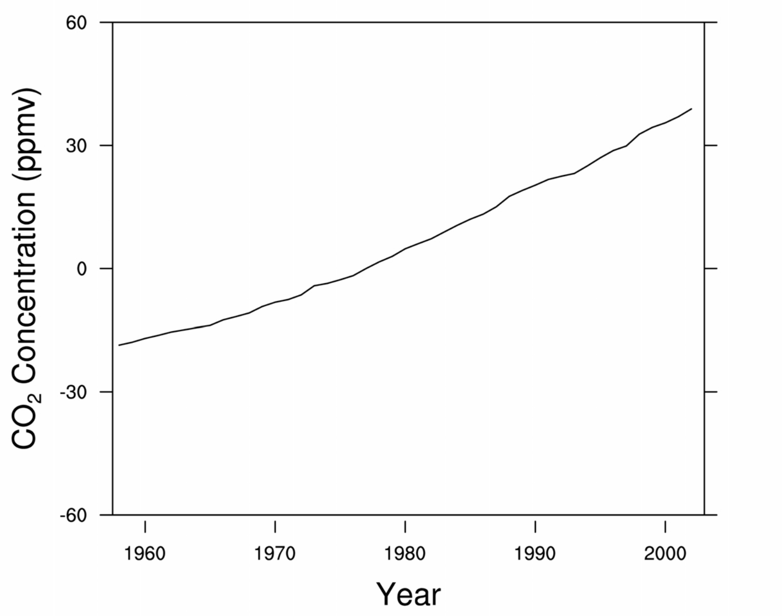


FIGURE 1-1 “The Keeling Curve.” Time series of annual departures from the 1961 to 1990 base period mean of 334 ppmv using direct measurements from Mauna Loa. SOURCE: Compiled by J. Hurrell, National Center for Atmospheric Research.

> . . . [T]he observing system available today is a composite of observations that neither provide the information nor the continuity in the data needed to support measurements of climate variables. Therefore, above all, it is essential to ensure the existence of a long-term observing system that provides a more definitive observational foundation to evaluate decadal- to century-scale variability and change.

## A NEED FOR CLIMATE OBSERVATIONS

Long-term observations sustained over decades are a critical first step in providing the climate data necessary for scientists, decision makers, and stakeholders to make adaptive choices that could improve resiliency to climate change and vulnerability, and maintain economic vitality. Many international and national activities and reports concur on the need for long-term climate observations. Internationally the 2002 Johannesburg World Summit on Sustainable Development called for strengthened cooperation and coordination among global observing systems and research programs, and the 2003 G8 Summit in Evian, France, appealed for strengthened international cooperation on global observations of Earth's environment. The UN Framework Convention on Climate Change (UNFCC), through the Subsidiary Board on Scientific and Technical Assessment (SBSTA), has adopted the Second Adequacy Report on Global Climate Observations (GCOS, 2003), which outlines the needed observations, networks, and climate variables.

Nationally, the U.S. Global Change Research Act of 1990 specifically highlighted the climate data needs for ". . . global measurements, establishing worldwide observations necessary to understand the physical, chemical, and biological processes responsible for changes in the Earth system on all relevant spatial and time scales." The NRC report on global environmental change (NRC, 1999a) also emphasized the critical nature of high-quality, long-term observations of the Earth system from both a public policy and a scientific perspective. More recently, the Climate Change Science Plan (CCSP, 2003), which integrates activities from the U.S. Global Change Research Program (USGCRP) and the Climate Change Research Initiative (CCRI), continues to emphasize the need for long-term, high-quality observations. A specific component of the CCSP plan addresses the following question:

> How can we provide active stewardship for an observation system that will document the evolving state of the climate system, allow for improved understanding of its changes, and contribute to improved predictive capability for society?

NOAA's mission for the next century includes a bold new mandate to "understand climate variability and change to enhance society's ability to plan and respond," and NOAA plans to create a global observing and data management system to help to achieve this goal (Box 1-1). With climate science now a high priority, NOAA is creating a CDR program to help to fulfill the climate mandate. The functions of the CDR program include:

- monitoring observing performance for long-term applications;
- generating authoritative long-term records from multiple observing platforms;
- assessing the state of atmospheric, oceanic, land, cryospheric, and space environments; and
- properly archiving and providing timely access to data and metadata.

### BOX 1-1 NOAA'S NEW PRIORITIES FOR THE TWENTY-FIRST CENTURY

NOAA's new mission for the twenty-first century is "to understand and predict changes in the Earth's environment and conserve and manage coastal and marine resources to meet our Nation's economic, social, and environmental needs." To fulfill its mission, NOAA has defined four interrelated goals.

- Protect, restore, and manage the use of coastal and ocean resources through ecosystem-based management.
- Understand climate variability and change to enhance society's ability to plan and respond.
- Serve society's needs for weather and water information.
- Support the nation's commerce with information for safe, efficient, and environmentally sound transportation.

Six cross-cutting priorities are

- integrated global environmental observation and data management system;
- environmental literacy, outreach, and education;
- sound, reliable state-of-the-art research;
- international cooperation and collaboration;
- homeland security; and
- organizational excellence: leadership, human capital, facilities, information technology, and administrative products and services.

## THE PURPOSE OF THIS STUDY

The creation of high-quality, long-term data of global atmospheric, oceanic, and terrestrial satellite observations is a key component of NOAA's strategy for achieving its new climate mandate (Box 1-1). Generation of these data also could be a pivotal aspect of the CCSP goal for observations and monitoring. Observations from both operational and research satellite programs are a primary information source for studying climate variability and change, in part because they uniquely provide global or near-global data. Yet, as noted by the NRC (2000a),

> The development of high-quality, long-term satellite-based time series suitable for detection of climate change as well as for characterization of climate-related processes poses numerous challenges . . . Long-term, con-

> sistent data sets require careful calibration, reprocessing, and analysis that may not be necessary to meet the needs of short-term forecasting . . . Such conflicts are difficult to resolve and are complicated by differences in agency cultures, charters and financial resources.

To generate the best possible plan for creating climate-quality data, NOAA asked the National Academies to assist in developing a plan for creating CDRs using satellites that monitor environmental conditions (see Box 1-2). The National Academies formed the Committee on Creating Climate Data Records from NOAA Operational Satellites and charged it with providing a comprehensive and practical evaluation of the NOAA CDR plan, including conclusions and recommendations. This interim report is the first phase of a two-phase process. It provides NOAA with general advice on the elements needed in a successful CDR generation process. NOAA will then use these recommendations to develop a plan to guide generation of satellite-based CDRs. In the second phase, the committee will review the NOAA CDR plan and make specific comments.

The committee's 13 members are experts in the creation, use, and maintenance of CDRs (see Appendix B); they met four times in generating this interim report. The first meeting was a large community workshop, attended by over 40 scientists familiar with CDRs (see Appendix A for the workshop agenda and participant list). The committee also solicited community input by distributing a short questionnaire. In creating this report the committee relied upon the expertise of its members and the opinions of the community as discussed in the workshop and the surveys.

To assist NOAA in its planning process, the committee first had to agree on a definition of a climate data record. The idea of a "climate data record" has surfaced numerous times in recent literature, yet comments from workshop participants and community surveys indicated that the climate community has yet to settle on a consistent definition. For this report the following definition is used:

> **A climate data record is a time series of measurements of sufficient length, consistency, and continuity to determine climate variability and change.**

This report focuses on CDRs that are derived from satellite observations, which combined with ancillary data, potentially resolve the time and space scales of climate change and variability. In general, production of CDRs involves long time series of data from a variety of sensors, with quantified error characteristics. Data life cycles are long in relation to a human life span and are definitely longer than any single mission or set of missions.

## BOX 1-2 STATEMENT OF TASK

The ad hoc committee charged to conduct this study will assist the National Oceanic and Atmospheric Administration-National Environmental Satellite, Data, and Information Service (NOAA-NESDIS) as it designs a plan to create climate data records (CDRs) from existing and new instruments aboard NOAA satellites for understanding, monitoring, and predicting climate variations and changes. The committee will provide input to the plan by summarizing major needs for and uses of climate data records, examining different approaches and strategies for generating climate data records, and identifying key attributes of CDRs that have proven useful. NOAA would then use this information as guidance to develop its plan for producing CDRs from operational satellites. Once the plan is drafted, the committee will review the draft Climate Data Records Plan to ensure that it is sound, comprehensive, and includes mechanisms for continued user involvement, and it will recommend improvements to ensure that CDRs are processed according to established scientific methods and packaged in forms that are useful for real-time assessments and predictions of climate as well as retrospective analyses, re-analyses, and reprocessing efforts.

In phase I, the committee will organize and host a workshop to facilitate discussion of an NOAA white paper that will outline its preliminary ideas on satellite data utilization for climate applications, and it will write an interim report that:

- Summarizes major needs for and uses of climate data records,
- Examines different approaches and strategies for generating CDRs, and
- Identifies key attributes of examples of successful attempts to create high quality CDRs from satellite data.

Questions to be addressed in the workshop and by the committee include:

- How does a CDR become a community standard (i.e., established as legitimate)?
- How can NOAA ensure that the CDRs are responsive to user needs?
- What are the key attributes of successful CDR generation programs?
- What are the advantages and disadvantages of different models or strategies for producing CDRs, such as using partnerships among government, academia, and the private sector, different blends of space-based and in situ data (e.g., all space-based versus some balance), or other approaches?
- How can NOAA learn from present and past efforts such as the NOAA/NASA Pathfinders, Earth Observing System Data and Information System, etc.? What are the successes and failures, and how do we emulate the successes or avoid the pitfalls?

Phase 2 will begin when NOAA provides the committee with a draft of its Climate Data Records Plan (estimated to be approximately three months after delivery of the interim report).

These data will need to be reprocessed from the beginning of the series as new information is obtained; the quality, spatial resolution, and temporal resolution of the product may also improve through the time span covered by the data. Ultimately the CDRs should be consistent, continuous records of a climate system variable that do (or will) span at least a multidecadal period.

Not all time series of climate-related variables are designated as CDRs. Variables chosen for CDR development should address key questions about the climate system and lead to clear improvements in (1) scientific understanding of the climate system; (2) projections for future climate states; (3) regional, national and international climate assessments; and (4) the nation's ability to respond to climate variations. The CDRs should be based on the best scientific research and measurement capability available, and they should represent a consensus within the scientific community regarding what is to be monitored and measured over time.

The committee further defines a hierarchy of CDRs (see Figure 1-2). Fundamental CDRs (FCDRs) are sensor data (e.g., calibrated radiances, brightness temperatures, radar backscatter) that have been improved and quality controlled over time, together with the ancillary data used to calibrate them. Thematic CDRs (TCDRs) are geophysical variables derived from the FCDRs, specific to various disciplines, and often generated by blending satellite observations, in situ data, and model output.

Plans for the National Polar-Orbiting Operational Satellite System (NPOESS) call for the generation of sensor data records (SDRs) and environmental data records (EDRs). The SDRs are time tagged, geolocated, and calibrated antenna signals, but they will not be created for long-term stability and reliability, and they will therefore not be suitable for climate purposes without reprocessing into FCDRs. Algorithms for TCDRs change over time as new scientific discoveries prompt changes; however, the FCDRs will eventually become fixed as our ability to improve calibrations of past satellite sensors will diminish over time. No one can know which theories, processes, or applied products will emerge as critical to scientists, decision makers, or stakeholders in future decades. Therefore, the generation, preservation, and maintenance of the FCDRs is vitally important for ensuring the success of NOAA's program. The FCDRs will be the ultimate legacy that the long-term satellite programs leave to the next generation. The EDRs are time tagged and geolocated parameters produced from the SDRs, but they also are not calibrated and validated for long-term studies, unlike the CDRs.

Although CDRs can be created with multiple satellite platforms and in situ data, this committee was asked to focus mainly on the steps necessary

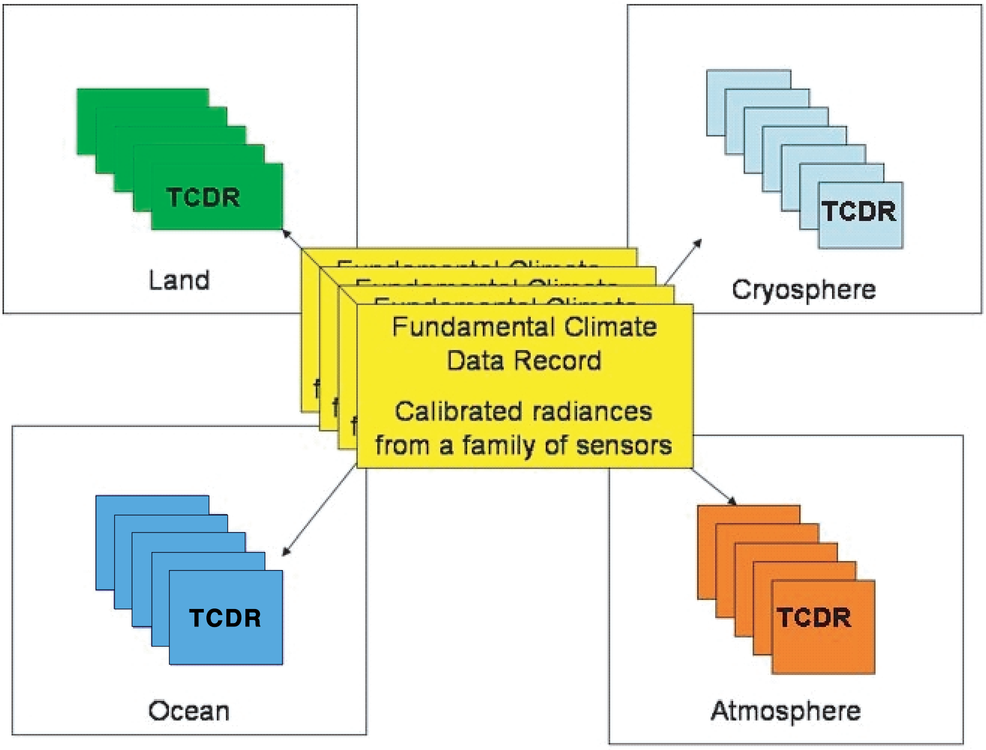


FIGURE 1-2 Thematic CDRs (TCDRs) related to different themes will be generated from the fundamental climate data records (FCDRs); for example, the calibrated antenna signals from a series of satellites (e.g., AVHRR, MODIS, VIIRS) will be used to generate a variety of TCDRs. A major effort should focus on creating and managing the FCDRs. The process of calibrating the FDCR generally involves the use of in situ measurements and critical feedback resulting from assessments of the TCDRs. Arrows might be shown in two directions. SOURCE: J. Campbell, University of New Hampshire.

to create and maintain state-of-the-art CDRs with polar-orbiting satellites. As a result the committee did not refer specifically to generation of CDRs primarily with geostationary platforms or in situ data, such as the $CO_2$ record. In comparison with CDRs generated solely with in situ data, satellite-based CDRs present some unique challenges:

- the need to manage extremely large volumes of data;
- restrictions of spatial sampling and resolution;

- accounting for orbit drift and sensor degradation over time;
- temporal sampling (aliasing);
- difficulty of calibrating after launch (e.g., vicarious or onboard calibration); and
- the need for significant computational resources for reprocessing.

## PAST STUDIES OF NOTE

Concern over the future availability of satellite-based climate-quality data led the National Research Council and several other bodies to issue reports on ensuring the climate record from satellites (e.g., NRC, 1999a,b, 2000a,b,c,d,e; NOAA, 2001; GCOS, 2003). In addition to highlighting the need for climate data records, many of the reports recommend steps for the long-term creation and preservation of climate data from the NPOESS and the NPOESS Preparatory Project (NPP). Although NPOESS and NPP were originally envisioned for serving civilian and defense needs for environmental data, the climate community quickly realized that these platforms also would be the primary information sources for any satellite CDRs in the coming decades. The committee viewed these reports as stepping-stones for this project; NRC (1999a) outlined the state of the observing system relative to the USGCRP and discussed several elements of a climate observing system, while NRC (1999b) assessed the adequacy of the climate observing system and endorsed the now well-known 10 climate monitoring principles. NRC (2000b) provided a short overview for creating and maintaining climate data specifically for NPP and NPOESS, and NRC (2000a,c) outlined in greater detail the science, design, and implementation of a potential program for creating climate-quality data for NPOESS. NRC (2000d) examined atmospheric temperature trends near the surface and in the lower to middle troposphere to reconcile disagreements in the observed trends. NRC (2000e) built upon the recommendations from NRC (2000d) and discussed strategies for NOAA to develop long-term monitoring capability of upper air temperature CDRs; as such, NRC (2000e) is particularly relevant for NOAA to refer to in addition to this report. NOAA (2001) was written by the NOAA Science Advisory Board, and this report suggested the creation of a new program for climate monitoring within NOAA, including but not limited to satellites. The GCOS (2003) report examined the state of the global climate observing system and suggested various methods to address inadequacies.

This report builds on the wealth of information available, giving specific attention to creating CDRs useful to NOAA's new climate mandate. It also provides practical advice to help NOAA to create CDRs from operational satellite data that respond to the needs of the climate science community as well as policy makers and other stakeholders, utilize the best scientific practices in the creation of CDRs, and are properly archived and disseminated to the user community.

## REPORT ORGANIZATION

This report is organized into six chapters. The following chapter (Chapter 2) discusses lessons learned from previous attempts at creating climate-quality data that NOAA should consider in developing its plan. Based on the historical lessons, committee expertise, community surveys, and the workshop, Chapter 3 outlines the key elements needed for a successful CDR generation program, beginning with identification of an appropriate organizational framework, continuing with suggested steps for creating the CDRs, and ending with comments on sustaining the program. Since data management is an integral component of the CDR legacy left to the next generation, Chapter 4 provides comments on data storage, archiving, and dissemination. Finally, with a realization that creating effective CDRs for every possible variable is a task that NOAA could never hope to achieve alone, Chapter 5 discusses the importance of partnerships. Chapter 6 summarizes the committee's recommendations, beginning with an overarching recommendation and six supporting recommendations.

# 2

# Lessons Learned from Previous Programs

NOAA will not be the first agency or group to generate climate-quality data, and NOAA can learn many lessons from previous efforts. This chapter reviews variables that have been observed over long periods of time and have evolved in their use by the wider science community to become de facto Climate Data Records (CDRs). The intention of this review is to gather lessons learned from these past experiences and to give guidance for future activities in creating CDRs. This list of variables is illustrative and not comprehensive. Each brief summary contains a list of findings for that variable. After the presentation of these examples, the chapter concludes with lessons culled from the individual examples. The emphasis in these examples is on satellite data analyses and their antecedent measurements as they apply to long-term climate problems.

## ATMOSPHERIC TEMPERATURE PROFILES

In 1969 Nimbus 3 carried the first of a new class of remote-sounding sensors, the Space Infra-Red Sounder (SIRS A), which demonstrated that satellite sensing can provide vertical temperature profiles extending from the stratosphere to the surface with global coverage and useful spatial resolution and accuracy. Shortly thereafter, simulation studies with General Circulation Models (GCMs) validated the "Charney conjecture" that given the continuous historical record of global atmospheric temperature profiles over a sufficiently long integration period, one should be able to infer, with the assistance of a GCM, the complete state of the atmosphere. National Environmental Satellite Data and Information Service (NESDIS) subsequently launched the first operational sounder system in 1972, the Vertical Temperature Profile Radiometer (VTPR), aboard the NOAA 2 satellite. In 1978 Television Ifrared Operational Satellite—Next generation (TIROS N) was launched with an improved 20-channel High Infrared Sounder (HIRS)

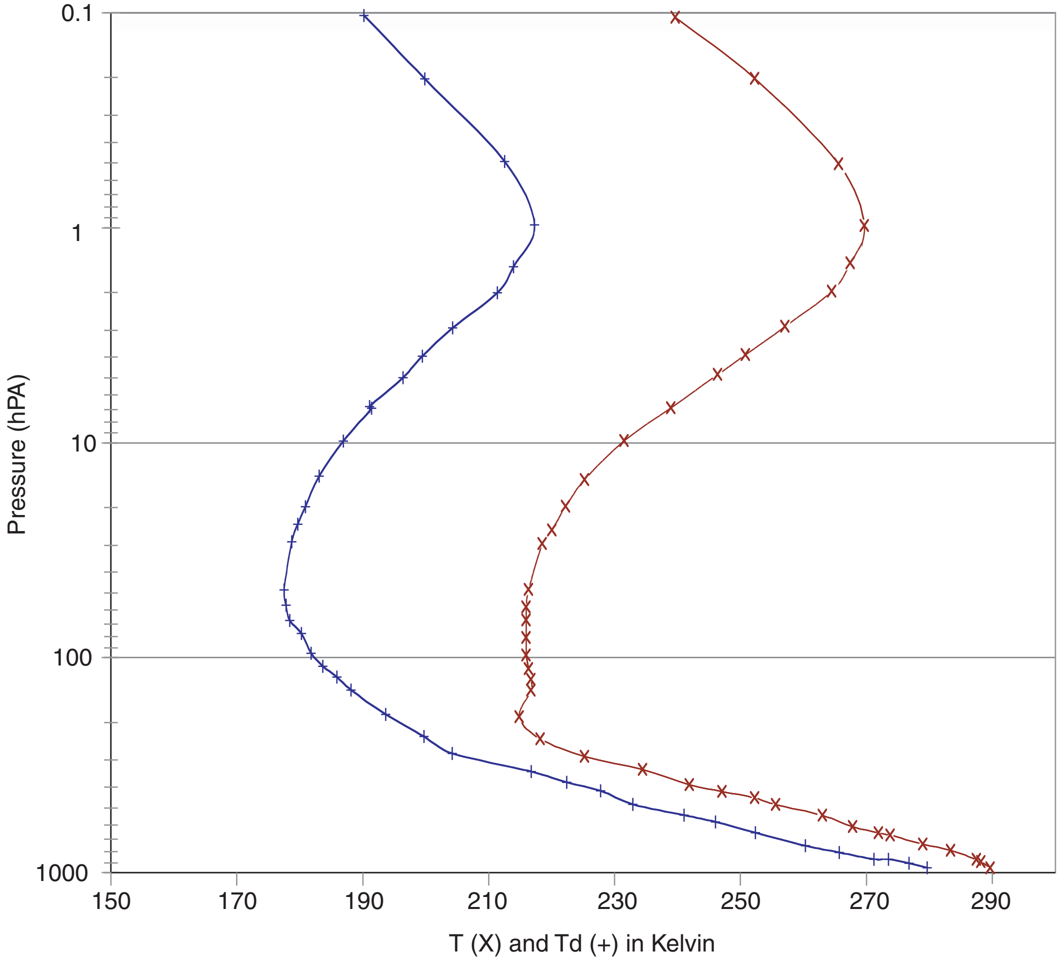


FIGURE 2-1 TOVS data from the NOAA 15 satellite using MSU and AMSU respectively for the European area.

accompanied by the Microwave Sounder Unit (MSU) and Stratospheric Sounder Unit (SSU) forming the TIROS Operation Vertical Sounder (TOVS) (Figure 2-1).

The NOAA polar-orbiting operational satellite series, starting with NOAA 9 and extending through NOAA 16, has carried essentially the same TOVS instrument package (an improved microwave sounder Advanced

Microwave Sounding Unit [AMSU] was added on NOAA 14 and subsequent satellites), providing a continuous record of global atmospheric temperature profiles for almost 25 years. These data have become indispensable inputs to operational forecast centers. In 2002 NASA launched Aqua carrying the Atmospheric Infrared Sounder (AIRS)/AMSU/Humidity Sounder for Brazil (HSB) sounder system with hyper IR spectral resolution as well as considerably higher spatial resolution. This represented the first significant sounder system upgrade to the TOVS system in 25 years, and the NPOESS Preparatory Project/National Polar-orbiting Operational Environmental Satellite System (NPP/NPOESS) system will carry advanced operational sounder systems through the year 2025.

To address problems related to understanding global climate NOAA and NASA initiated a Pathfinder program in the mid-1980s to carefully reprocess satellite products into climate data records extending over a common fixed period from April 1, 1987 to December 31, 1988. A TOVS science working group identified three conceptually distinct algorithms in 1991 for consideration, and NASA/NOAA supported TOVS Pathfinder climate data studies spanning the different algorithmic approaches for this suite of instruments.

The algorithm employed in the TOVS A dataset is a physically based algorithm using a GCM model analysis first guess. This dataset produces column radiances both clear and cloud contaminated, atmospheric temperature profiles, humidity profiles, sea and land surface air and ground temperatures, cloud cover, cloud heights, precipitation, surface emissivity, outgoing longwave radiation, and albedo and ozone profiles. These products have been subsetted into a variety of monthly, seasonal, and interannual means, variances, and climate anomalies. These geophysical parameters are mapped to a $1° \times 1°$ latitude-longitude grid at three different temporal resolutions: daily, 5 day, and monthly, separately for the AM and PM satellites. The second algorithm employed in the TOVS B dataset is a statistical-physical approach using a neural-net-matching approach from a select set of 1800 radiosondes to obtain a first guess, then followed by a Bayesiasn statistical inversion. These datasets include 1° grids derived from Channel 2 of the MSUs for daily and monthly lower and upper tropospheric temperatures as well as lower stratospheric temperatures for the AM and PM satellites. The third dataset TOVS C1 includes 1° grids derived from Channel 2 of the MSUs for daily and monthly lower and upper tropospheric temperatures, as well as lower stratospheric temperatures for the AM and PM satellites. MSU Channel 2 (53.74 GHz) is sensitive to deep layer average tropospheric temperatures with a weighting function peaking near 500 hPa and is very slightly affected by variations in tropospheric humidity, but is contaminated by precipitation-size ice in deep convective clouds and high elevation terrain. A TOVS C2 dataset for the 18-month period employed a second algorithm version for the Channel 2 MSU data. The Channel 2 MSU data were averaged along the scanline to produce a 2 × 2 degree gridded monthly dataset, except for one product that used adjacent field of views. The data were also screened for precipitation and high terrain.

### Lessons Learned

- Preparing long-term scientifically credible CDRs spanning two decades and longer requires frequent reprocessing to deal with unanticipated problems arising from a variety of factors that are often not revealed until several years of data are processed and analyzed.
- An issue faced by all the TOVS Pathfinder teams was acquiring the long-term resources needed to keep a team of scientists, instrument engineers, and computer programmers together to make longer CDRs and to utilize evolving technology.
- The TOVS Pathfinder project was conceived for a limited data period with relatively homogeneous instrument sensors and so did not address many of the problems faced in producing a climate dataset of more than two decades with instrument changes and long-term satellite stability.
- Significant additional computing resources are needed to revise algorithms and to rerun processing systems to remove any spurious interannual drifts and jumps.
- Calibrating and tuning even the same instruments for CDRs requires at least annual overlaps.
- Introducing new instruments with increased spectral functionality, resolution, and coverage into the system raises the possibility of CDRs branching into multiple versions. New sounder instruments significantly increase the data volume, requiring greater computing and archiving resources.
- Multiple algorithmic approaches with different science teams for the same data products should be supported in order to evaluate product accuracy as well as the strengths and weaknesses of varying algorithms.
- Trends are difficult to evaluate owing to remaining calibration issues.
- A collocated raob-radiance database should be maintained.

## TROPOSPHERIC TEMPERATURES

The possibility of inadvertent climate change prompted many analyses of historical weather data. Globally averaged surface temperature (combined land surface air temperature and sea surface temperature), for instance, increased 0.4-0.8ºC since the late nineteenth century, with the most rapid warming (0.20 ± 0.06ºC decade$^{-1}$) over the past 25 years (Figure 2-2). Upper air data, although available for only the last few decades, have also received increased scrutiny, including the radiances provided by the MSU instruments on NOAA polar-orbiting satellites since 1979. Global lower to middle tropospheric temperatures inferred from the MSU record exhibit considerably less warming than the surface record (Figure 2-2).

This situation, in which multiple trends with nonoverlapping error ranges for supposedly identical products were published in reputable journals, has resulted in considerable confusion to those not closely related to this area of research. This apparent discrepancy therefore motivated much recent research and debate (NRC, 2000d,e), and it is now clear that a number of factors likely contribute to the different trends. Among them are (1) the influence of real changes in the vertical structure of the atmosphere associated with both natural variability and human-induced climate change;

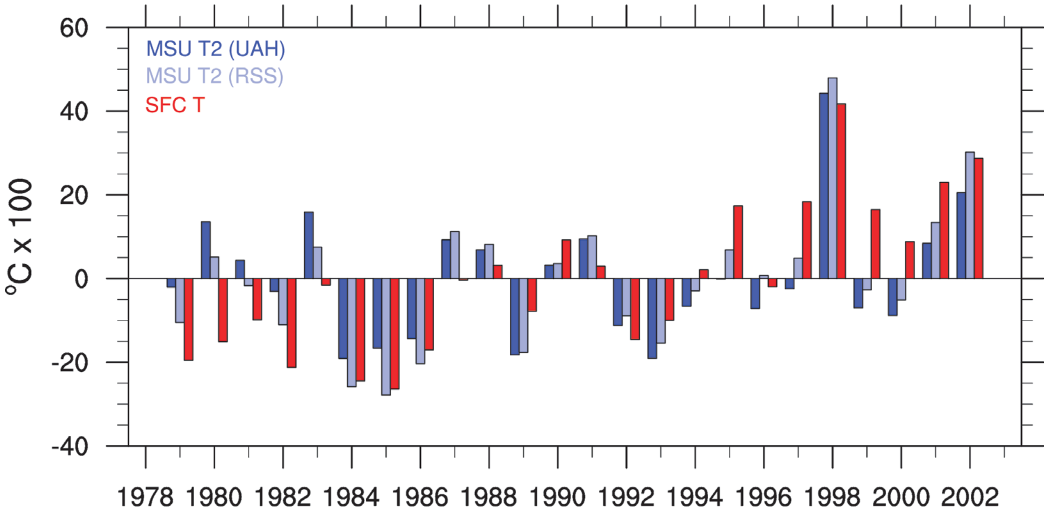


FIGURE 2-2 Annual mean anomalies of global average temperature (1979-2002) for the lower troposphere from satellites (T2) and for the surface. The MSU T2 products are University of Alabama Huntsville (UAH) 5.1 (Christy et al., 2003) and Remote Sensing System (RSS) (Mears et al., 2003). The surface temperature trend is +0.20 ± 0.06°C decade$^{-1}$. The linear trend through 2002 for the UAH (RSS) T2 product is 0.03 ± 0.09°C decade$^{-1}$ (0.11 ± 0.09°C decade$^{-1}$). The estimated 95 percent confidence intervals on the trends due strictly to measurement error are +/- 0.05°C for UAH and +/- 0.02°C for RSS.

(2) technical issues related to comparisons of global-mean temperature trends derived from different instrumental measurements with different physical, spatial, and temporal sampling characteristics; and (3) data uncertainties, as neither of these measurement systems was specifically designed for long-term climate monitoring (NRC, 1999b).

A chronic difficulty in creating a continuous, consistent climate record from satellite observations alone is that satellites and instruments have a finite lifetime of a few years and have to be replaced, and their orbits are not stable. Nine satellites, and the follow-on AMSU, compose the current operational record, and the methods of merging the data from these different satellites and channels are complex. Moreover, the satellite data record is continually evolving as newly discovered problems are accounted for and corrected.

NRC (2000d) provides a summary of the principal sources of uncertainty in trend estimates of MSU temperatures. They include systematic measurement errors, and, in particular, poorly understood problems with radiometer gain (the ratio of the perceived to the actual signal), which seriously impact the ability to intercalibrate the series of MSUs. During periods of overlap, which are sometimes far too short, temperatures measured by two different MSUs are compared, and offsets of 0.4ºC in magnitude are typical. Different adjustment methods to these offsets produce a spread in trends of about 0.1ºC decade$^{-1}$. There is also uncertainty in determining each satellite's bias relative to some reference value. Even in the absence of measurement errors, orbital decay (decrease in satellite altitude), orbital (diurnal) drift (change in the local time), and other effects have the potential to introduce spurious signals into the MSU temperature record. The effect of such uncertainties on trends is apparent in two independent MSU tropospheric temperature reconstructions (Figure 2-2), which differ by about 0.10ºC decade$^{-1}$ over 1979-2002.

It is critically important to the climate community that assessments of the retrieval methodology and assumptions that are used to compute the MSU temperature record continue to be performed and documented. Additionally, different versions of the temperature time series need to be available in an easily accessible form.

### Lessons Learned

- The satellite data record is continually evolving as newly discovered problems are accounted for and adjusted, requiring reprocessing to attempt to create consistent and stable data records.

- Systematic measurement errors are not constant in time—biases between instruments on different satellites change—and are thus difficult to remove, especially when overlap periods between satellites are short.
- Adjustment methods different from these offsets can be rationalized, yet each method leads to different results/conclusions.
- It is critically important that independent processing efforts of the radiance data have been performed. Through multiple independent efforts, important insights into the precision of the MSU temperature records are being obtained.

## SATELLITE PRECIPITATION

Temperature and precipitation are the most fundamental elements in defining climate for a region, for the world, or for a specific time period. Temperature data are often more representative in time and space for a given observational network than precipitation data, and this has made it much more difficult to produce precipitation data representative of climate. The significant difficulty in measuring solid precipitation adds to the problem. However, long-term and broad-area averaging does allow for reasonable classification of climatic types and trends from gauge analyses for those areas with installed gauges.

Recent (over the last 10 years) development of satellite capabilities to retrieve data related to precipitation amounts over broad space and relatively long time periods have provided the opportunity to observe ungauged areas (especially key ocean areas for climate, such as the tropical western Pacific) and also provide global coverage (Figure 2-3). Understanding Earth's global energy and water cycles and their variability and changes depends upon an accurate representation of precipitation and related latent heating profile.

The origin of satellite retrievals to estimate precipitation began in the late 1960s, and retrievals using visible, IR, microwave, and combination techniques have been developed with varying degrees of success. One important early algorithm that is still frequently applied uses the correlation between IR-based cloud-top temperature and precipitation for time and space scales of a day and 250 km or more. More recently (1987) Special Sensor Microwave/Imager (SSM/I) data permitted the estimation of precipitation using multiple microwave channels, providing a more accurate, but sparser global dataset. However, the microwave data are obtained from polar orbiters and thus have temporal sampling problems. Estimates from TOVS and Outgoing Long-wave Radiation (OLR) data provide estimates in snowy-surface regions, where both IR and microwave schemes fail.

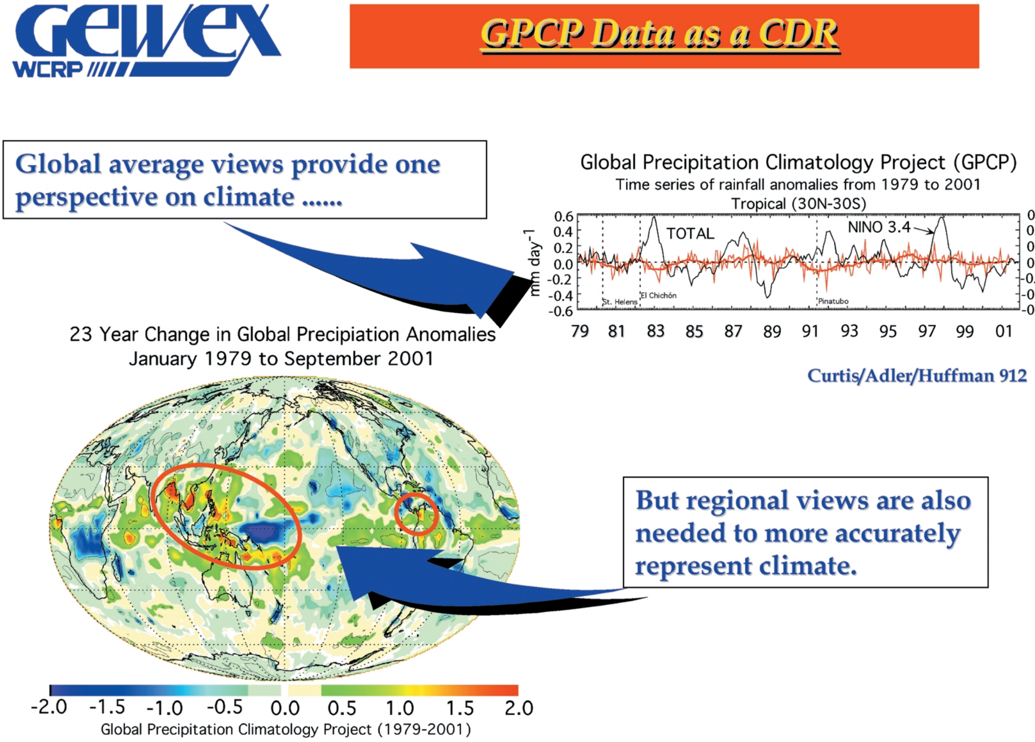


FIGURE 2-3 The Global Precipitation Climatology Project (GPCP) data highlight the importance of generating regional and globally averaged CDRs.

The other piece of the puzzle was the development of analyses to combine the various satellite estimates, and then merge in the rain gauge data into a homogeneous global dataset. This approach formed the basis for the international World Climate Research Programme (WCRP), the Global Energy and Water Cycle Experiment (GEWEX), Figure 2-3), and the Global Precipitation Climatology Project (GPCP) monthly 2.5 × 2.5° product. The GPCP has also encouraged other analysis procedures: 1 × 1° daily and pentad, and even working towards tri-hourly representations of precipitation. In each case spatial coverage requires the merging of multiple satellite datasets and uses in situ gauge data.

Most recently intercomparison and intercalibration with the Tropical Rainfall Measuring Mission (TRMM) satellite data has provided improvements over the GPCP retrieval schemes, but these improvements have not yet been applied to harmonizing the various generations of satellite data over the 23-year record to provide the needed accuracy for water balance climate change trends.

### Lessons Learned

- Although the GPCP methodologies can be used to produce monthly 2.5 × 2.5°, daily 1 × 1°, and even tri-hourly 0.25 × 0.25° precipitation data, the random errors grow significantly larger as the scales become smaller.
- Definitive characterization of bias error in regions lacking rain gauges, including oceanic and underdeveloped areas, is an unsolved problem and a matter of current research.
- Random errors, and sometimes bias errors, are strongly dependent on the frequency of sampling by high-quality (e.g., microwave) sensors, particularly at fine scales.
- Solid precipitation, high-latitude precipitation and precipitation over complex terrain remain extremely difficult to retrieve from satellite data. Gauge analyses face challenges in the same regions, when they are available at all, compounding the problem.
- It is critical to have accurate in situ validation or reference data in a variety of climate regimes to facilitate the development and long-term quality assurance of reliable satellite-based precipitation estimates.
- It is important to have as many accurate in situ observations as possible to provide the tie points that are critical to maximizing the accuracy of the final "best" combination products.

## EARTH RADIATION BUDGET AND CLOUDS

Two closely related and overlapping sets of climate variables are required for characterizing clouds and quantifying Earth's radiation budget. The radiation budget involves monitoring and understanding the fate of both solar radiation incident on the planet and terrestrial radiation emitted by it. Clouds play a major role in the radiation budget, contributing significantly both to the planetary albedo and to the greenhouse effect. In fact, uncertainties involving the role of clouds and cloud radiation feedbacks are centrally important to climate prediction on decade to century time scales. It has long been true that climate models differ by about a factor of three in their sensitivity to greenhouse gas concentrations, as measured by the "global warming" or equilibrium surface atmospheric temperature change in response to doubling carbon dioxide. For these reasons CDRs concerned with clouds and Earth's radiation budget are exceptionally important to climate research.

ERBE (Earth Radiation Budget Experiment) was an outstandingly successful program. It provided data on the spatial and temporal variability of quantities such as Earth's albedo and emitted energy (Figure 2-4). It also documented the radiative effects of clouds and provided a definitive answer to a critical question: clouds both cool Earth by contributing to the planetary albedo and warm Earth by contributing to the greenhouse effect, but which effect dominates? ERBE proved quantitatively that the global-mean net radiative effect of Earth's cloud cover is a significant cooling.

In the opinion of key participants an important factor in the success of ERBE was its small and dedicated science team. The foundation of that team was a group within NASA that claimed ownership of the project, augmented by some non-NASA scientists. Virtually all the team members were committed to the project and were focused on measuring two quantities,

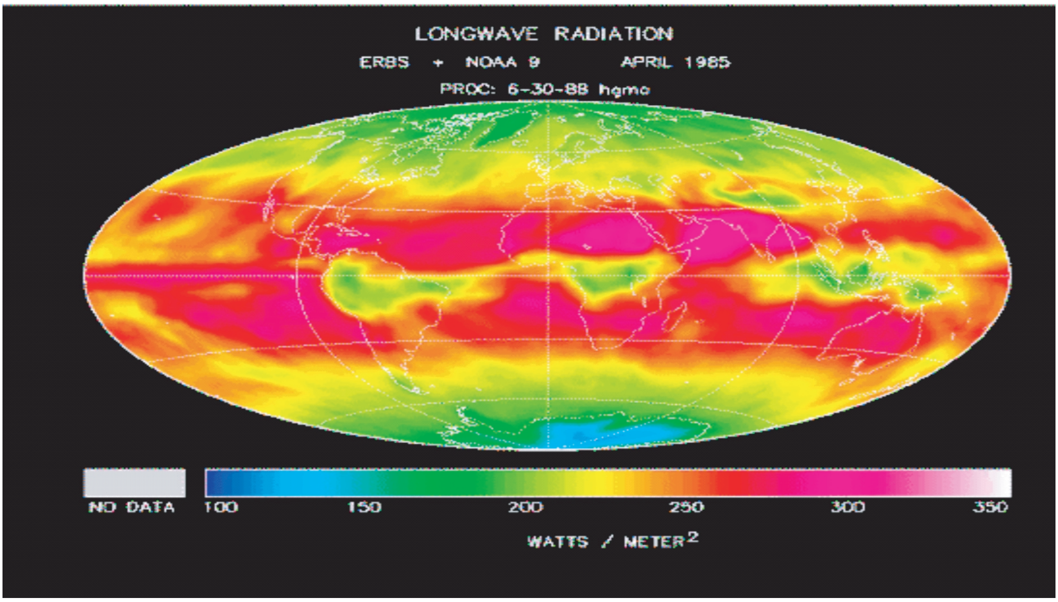


FIGURE 2-4 Outgoing longwave radiation from ERBE.

reflected shortwave and emitted longwave energy, with an accuracy that was unmatched at that time.

There was also a sustained effort undertaken by the science team to understand the behavior of the ERBE instrument. In ERBE, NASA supported the non-NASA science team members financially for up to 12 years; so long-term research was facilitated. ERBE also escaped the fate of having its research budget redirected to cover hardware cost overruns. NASA kept the ERBE science team budget separate from the hardware budget and protected it.

Two different but complementary approaches to observing clouds have been taken by the International Satellite Cloud Climatology Project (ISCCP) and by the Atmospheric Measurement Program (ARM). ISCCP is primarily a NASA-supported program, established in 1982, to collect and analyze satellite radiance measurements to infer the global distribution of clouds, their properties, and their diurnal, seasonal, and interannual variations. Data collection began in 1983 and continues to the present. The resulting datasets and analysis products have been used to improve understanding and modeling of the role of clouds in climate, with the primary focus being the effects of clouds on the radiation balance. ISCCP data are also used to support many other cloud-related studies, including several large projects aimed at an improved understanding of the hydrological cycle. ISCCP has devoted significant resources to data management. A Correlative Data Center coordinates the delivery of other satellite and conventional weather data. A Satellite Calibration Center normalizes the calibration of the geostationary satellites with respect to a polar orbiter satellite standard. All ISCCP data products are archived at an ISCCP Central Archive.

The ARM Program was begun in 1989 by the U.S. Department of Energy (DOE). ARM is part of DOE's effort to resolve scientific uncertainties about global climate change with a specific goal to improve the performance of climate models used for climate research and prediction. The development of better parameterizations, derived from an observation-based improved understanding of the cloud radiation problem, is the focus of ARM. In pursuit of its goal ARM has established field research sites in several climatically significant locations. At these sites data on the effects and interactions of clouds and radiation have been obtained over extended periods of time from a wide variety of instruments. The most developed ARM site is the Southern Great Plains site, which straddles the Kansas-Oklahoma border. As in the case of ISCCP and ERBE the ARM data have been made widely available. Substantial resources have been devoted to data processing, analysis, quality control, and other necessary aspects of making large heterogeneous datasets useful to a broad range of scientists.

### Lessons Learned

- ERBE’s success is related to consistently sustained resources over a substantial period of time.
- Participants outside the funding agency are extremely helpful.
- Collaborative programs receptive to advice from all sources yield success.
- A strong data management effort is important.
- The science team can play a key role in directing the program.
- Accurate calibration is essential.

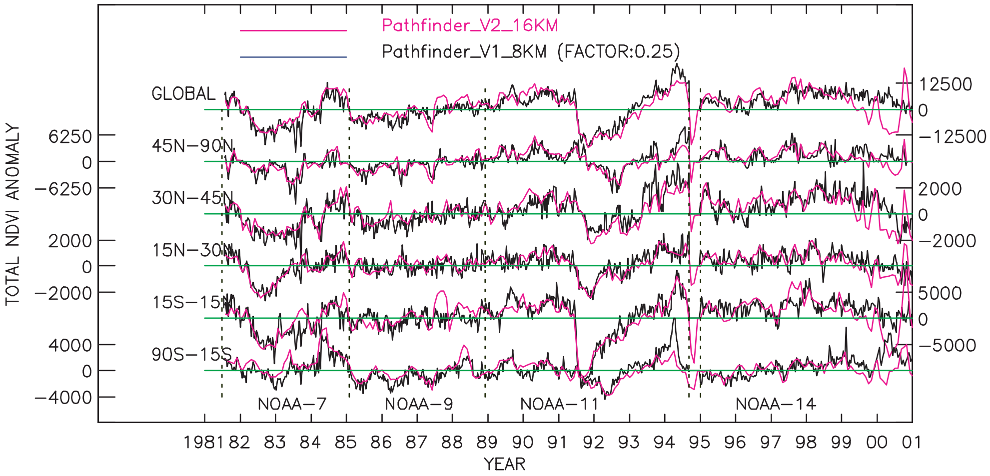
## VEGETATION DYNAMICS AND LAND COVER

NOAA currently produces an experimental product CDR termed “vegetation condition” from Polar Operational Environmental Satellite/Advanced Very High Resolution Radiometer (POES-AVHRR) sensor data. This CDR contains two geophysical variables: the Normalized Difference Vegetation Index (NDVI) and the Drought Index. This product serves a host of useful purposes in various sectors of our economy and supports valuable science on monitoring vegetation state and activity.

Given that the AVHRR sensors were never meant for vegetation monitoring, the many uses of these sensors’ data in documenting the human impact on global vegetation is a tribute to the collective creative efforts of a large scientific community; it laid the foundation for continued study of vegetation from space with next-generation sensors, such as Earth Observation System/Moderate Resolution Imaging Spectroradiometer (EOS-MODIS), NPP, and NPOESS-VIIRS. The AVHRR data will remain the start and a first important segment of the data record for vegetation monitoring. At the present time this data record is unique and comprises a nearly 20-year global record at a spatial resolution of 8 km (Figure 2-5).

### Lessons Learned

- A clear outline of the various steps involved in the development of a CDR of vegetation activity is needed.
- Algorithm developers and processing centers selected through peer review provide a mechanism to generate good products.
- Research on unresolved problems should be targeted through announcements and committed funding with a view toward integrating such research into CDR production.



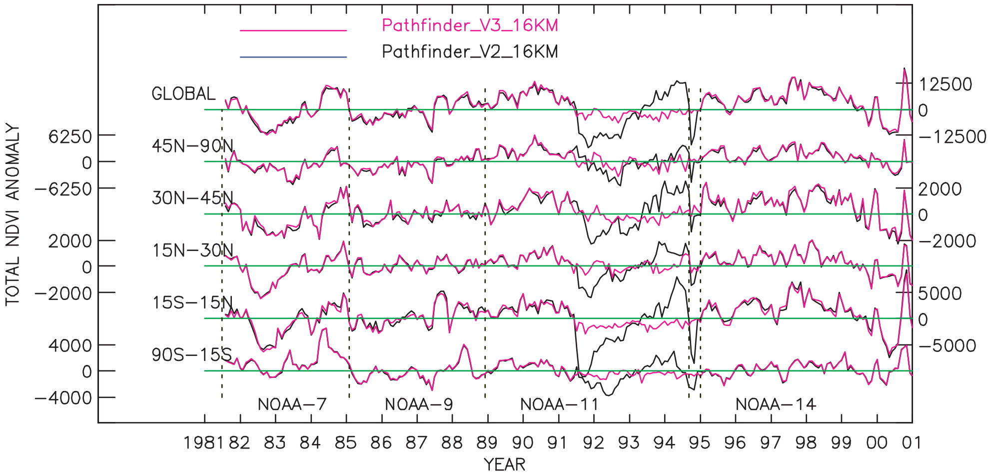


FIGURE 2-5 Vegetation index monthly anomaly time series for the period July 1981 to December 2000 for broad latitudinal bands and the globe (spatial resolution is 8 km$^2$). The time series from the Pathfinder AVHRR Land is shown here as Version 1 (V1). The data from different AVHRR sensors on NOAA-series satellites are indicated. The original time series shows high frequency variations due to residual cloud cover. The impact of satellite drift is clearly noticeable, especially in the case of NOAA 11 and 14. Likewise, the impact of the Mt. Pinatubo eruption in June 1991 and El Chichon in March 1982 is also discernable. The break in data in late 1994 is filled in with data from NOAA 9. Successive corrections to the data by spatial and temporal compositing (V2) and through correlations with climate data (V3) alleviate some of these problems.

- Archival of raw data at the highest possible resolution and accompanying detailed documentation of CDR algorithms are important for multiple reprocessing of data.
- Uncertainty estimates for various steps in CDR production and at pixel levels are needed for overall uncertainty estimates.
- Validation of algorithm physics and geophysical products through comparisons with in situ data must be performed.
- Community acceptance of CDRs occurs by involving a broad segment of the scientific and applications communities.

## NOAA SNOW MAP PRODUCT

The past three decades have seen the emergence of more accurate and complete information on the spatial extent and physical state of snow. This is leading to a better understanding of the variability of snow cover on annual to decadal scales, of cryosphere-climate interactions, and of the role snow may play in regional and global climate change. Snow data for these investigations come from a variety of satellite and in situ sources, including visible and microwave satellite products and point and snow course ground observations. Ultimately a blended snow product will best map the distribution, depth, and water equivalent, taking advantage of the strengths of each of the three sources of information.

Throughout the satellite era the premier dataset used to study snow extent on regional to hemispheric scales has long been the weekly visible wavelength satellite maps of Northern Hemisphere snow cover produced by NOAA (Figure 2-6). These maps constitute the longest satellite-derived environmental dataset available (Figure 2-7). This snow product is also unusual among satellite climate data records, as it involves a considerable ongoing manual effort to produce, and since its inception, has been produced in an operational mode for weather forecasting purposes.

The NOAA weekly snow map series began in late 1966. Since then, trained meteorologists have created maps in an operational weekly, and since 1999, daily mode. Mappers primarily relied on visual analyses of polar-orbiting satellite imagery to identify the location of snow cover across Northern Hemisphere lands. Secondary data sources included Geostationary Operational Environmental Satellites (GOES), Geostationary Meteorological Satellites (GMS), and Meteosat imagery. Map quality is predicated on the availability of clear sky visible imagery and the meteorologist's experience. In June 1999, production of the weekly maps ceased, and were replaced with a daily Interactive Multisensor Snowmap (IMS) product. IMS maps

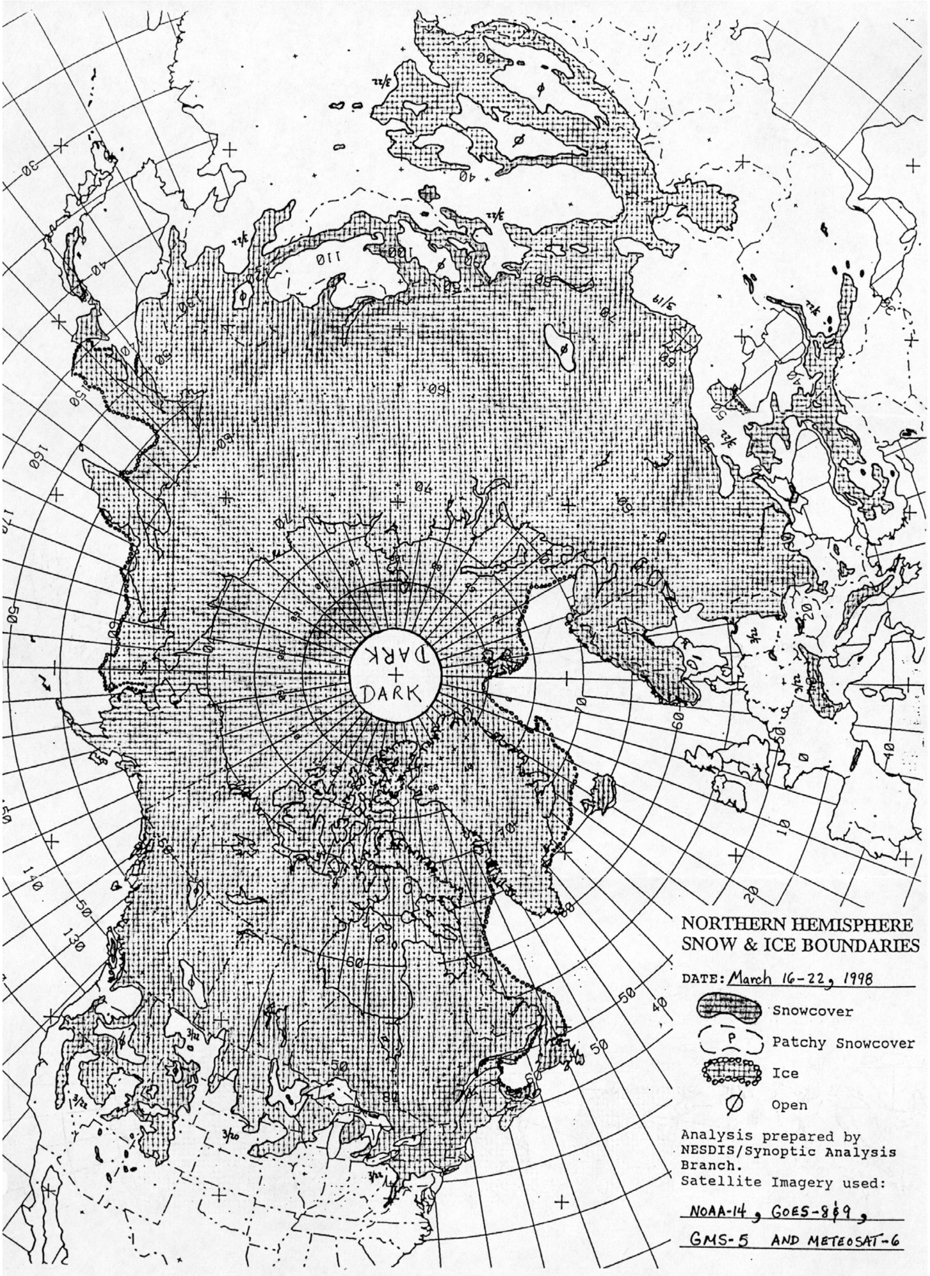


FIGURE 2-6 NOAA weekly snow map for March 16-22, 1998. Sea ice coverage is also shown.

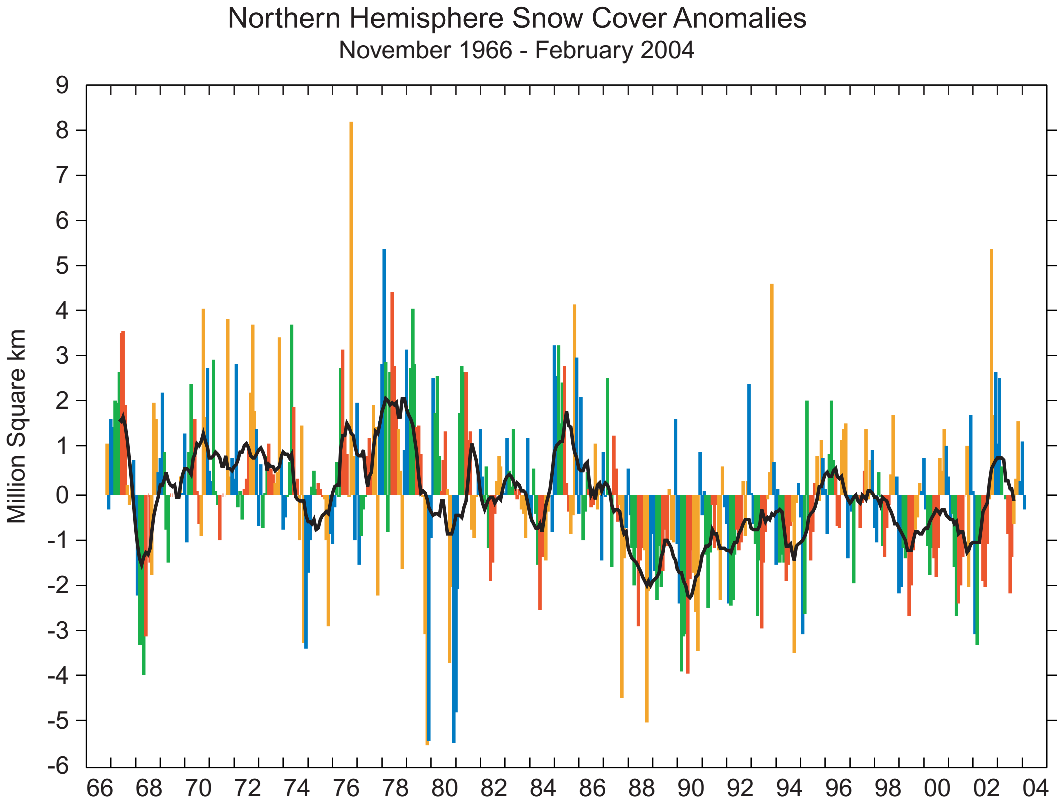


FIGURE 2-7 Anomalies of monthly snow cover extent over Northern Hemisphere lands (including Greenland) between November 1966 and February 2004. Also shown are 12-month running anomalies of hemispheric snow extent, plotted on the seventh month of a given interval. Anomalies are calculated from NOAA snow maps. Monthly anomalies are color coded by season: fall: orange; winter: blue; spring: green; summer: red.

still rely primarily on manual analyses of visible imagery, and are digitized to a 1024 × 1024 hemispheric grid, much finer than the 128 × 128 grid used previously.

An advisory board of climatologists from government, academic, and private sectors has yet to be appointed by NOAA to cooperate with the operational sector producing the maps. Despite this there have been some beneficial ad hoc efforts that, for instance, fostered cooperation between the forecasting and climate communities during the transition from the weekly to daily product.

### Lessons Learned

- Gratitude is owed to early satellite scientists for developing the weekly snow map product, to NOAA meteorologists for maintaining product production since then, and to NOAA for supporting the two-year overlap of the former NOAA weekly and new daily IMS products.
- Relying on the meteorology community as the primary driving force and production source of a climate data record has led to problems with data quality and continuity. The absence of sufficient metadata has led to delays in recognizing and correcting inconsistencies. Confidence limits and error margins have yet to be ascribed to the maps for any period of their existence.

## SEA SURFACE TEMPERATURE

As one of the mostly widely observed oceanographic parameters, sea surface temperature (SST) had its humble beginnings as measurements of the temperature of a bucket of seawater sampled from a sailing vessel with mercury-in-glass. Over time, canvas buckets replaced the wooden buckets. This switch was reflected in the time series of SSTs, which shows a drop in temperature between the late 1890s and 1940 (Figure 2-8). Studies by the Hadley Centre in England revealed that this negative bias in the bucket SST measurements was caused by cooling characteristics of the canvas buckets that were different from the wooden buckets, which insulated the surface sample much better than the canvas buckets.

The transition from sailing ships to powered vessels saw an increase in ship speed that made it impractical to collect bucket samples, and it was decided instead to use the seawater collected to cool the engines. This had two important consequences; first, the intake of this cooling water was significantly below the sea surface being between 3 and 5 m deep, and second, the engine room where the temperature of this cooling water temperature was read was a very warm location on the ship. A consequence of this latter effect was a sharp increase in the SST measured with this method. This can be seen in the large jump in SSTs in Figure 2-8 after 1940, which covers the period of this change from bucket samples to ship cooling water temperatures called “ship injection” SSTs.

FIGURE 2-8 Annual anomalies from 1951-80 average of uncorrected SST (green lines) and corrected (red lines) for (a) Northern and (b) Southern Hemisphere, 1856-1992.

Ship SSTs remained the primary source of SST measurements until the early 1970s. At that time a number of infrared sensors, which measure skin temperatures rather than bulk temperatures, were operating on NOAA weather satellites. The first operational SST product from a satellite was based on measurements by the Scanning Radiometer (SR)—a cross track instrument with an 8 km resolution in a single 11 µm thermal infrared channel. Radiances of the SR were evaluated against ship SSTs to derive a linear transformation algorithm to compute SST from the satellite data.

This was followed by the AVHRR, which became a standard of SST measurement from satellites. Most of these SST applications required in situ SST measurements originally from ships but later from drifting and moored buoys emphasizing the need for both satellite and in situ measurements. Recently the availability of passive microwave sensors, which measure temperature in approximately the top centimeter of the ocean, for the computation of SST has demonstrated a clear capability to sense SST through clouds and atmospheric water vapour. Since temperature measurements differ based on the recording instrument, new algorithms are being developed to blend thermal infrared and passive microwave satellite measurements of SST and relate them to in situ SST measurements.

### Lessons Learned

- Sea surface temperature measurements have evolved dramatically in the past 200 years. It is possible to extend the time series back in time by taking into account the observational methods of the past and correcting the data for these methods.
- Satellite SST is fundamentally different from in situ SST, and this difference has largely been ignored in the computation of routine SST products during the past two decades.
- New methods are being developed to mix thermal infrared and passive microwave satellite measurements of SST and to relate them to in situ measurements of SST.
- The definition of an SST product as a CDR is based on community use and application.
- New SST products will be defined and calculated in the near future and many of these may find wider acceptance as CDRs. This process depends on the people generating the products and quantifying the biases and errors, and their ability to communicate the value of their product and the interest of the user community.

## OCEAN COLOR

In many respects ocean color is the marine analog to vegetation dynamics and land cover, and it is motivated by similar concerns. Marine phytoplankton accounts for approximately half of the global annual primary production, and thus plays a significant role in the global carbon cycle. There are many reasons to expect that climate change could affect phytoplankton productivity: through changes in ocean circulation, temperature, atmospheric deposition of iron-rich dust, and surface irradiance, to name a few. Because of their fast turnover rates, phytoplankton have the potential to affect carbon uptake and release much faster than terrestrial vegetation. Thus, it is as important to monitor phytoplankton as it is to monitor vegetation dynamics on land.

Satellite measurements of phytoplankton biomass are based on the effect that phytoplankton pigments have on the color of the water. The most widely used product derived from ocean color measurements is chlorophyll concentration (mg $m^{-3}$), a measure of phytoplankton biomass. Ocean color remote sensing has been a successful area of technological development since the Coastal Zone Color Scanner (CZCS), a proof-of-concept sensor launched in 1978 onboard Nimbus 7. The CZCS operated for eight years but had a limited duty cycle (operating only 10 percent of the time), thus its coverage was spotty. Its calibration was suspected of drifting during the mission, but there were no means to calibrate the sensor in orbit. It is therefore not possible to determine trends from the eight-year CZCS record. There was a gap of 10 years before the next ocean color sensor, the Ocean Color and Temperature Sensor (OCTS), launched onboard the Japanese Advanced Earth Observing Satellite (ADEOS) satellite. The U.S. launched the Sea-viewing Wide Field-of-View Sensor (SeaWiFS) on the Orbview 2 satellite in August 1997.

SeaWiFS has collected ocean color data since September 1997, and continues to produce global products that are the best examples of climate-quality ocean color data currently in existence. MODIS, onboard Terra and Aqua, also provides ocean color data products. Both MODIS and SeaWiFS require a regular program of vicarious calibration whereby atmospherically corrected satellite radiances are compared with in situ water-leaving radiances. Adjustments are then made to the satellite top-of-atmosphere radiance calibrations to force agreement. SeaWiFS has remained remarkably stable, requiring only one postlaunch adjustment, whereas the MODIS on Terra has experienced a number of abrupt (and in some cases unexplained) changes in its radiometric calibration.

Algorithms for atmospheric correction and for deriving such water properties as chlorophyll concentration have evolved significantly since the CZCS era. Each ocean color mission now uses a different algorithm for chlorophyll, and in fact, MODIS has produced three products that are called "chlorophyll." Figure 2-9 illustrates differences among the chlorophyll products produced by the various algorithms.

In 1997 NASA initiated the Satellite Intercomparison and Merger for Biological and Interdisciplinary Ocean Science (SIMBIOS) program to address the problem of merging ocean color data into a seamless time series. This project supported an extensive validation effort (primarily in situ measurements made throughout the world's oceans). There were also investiga-

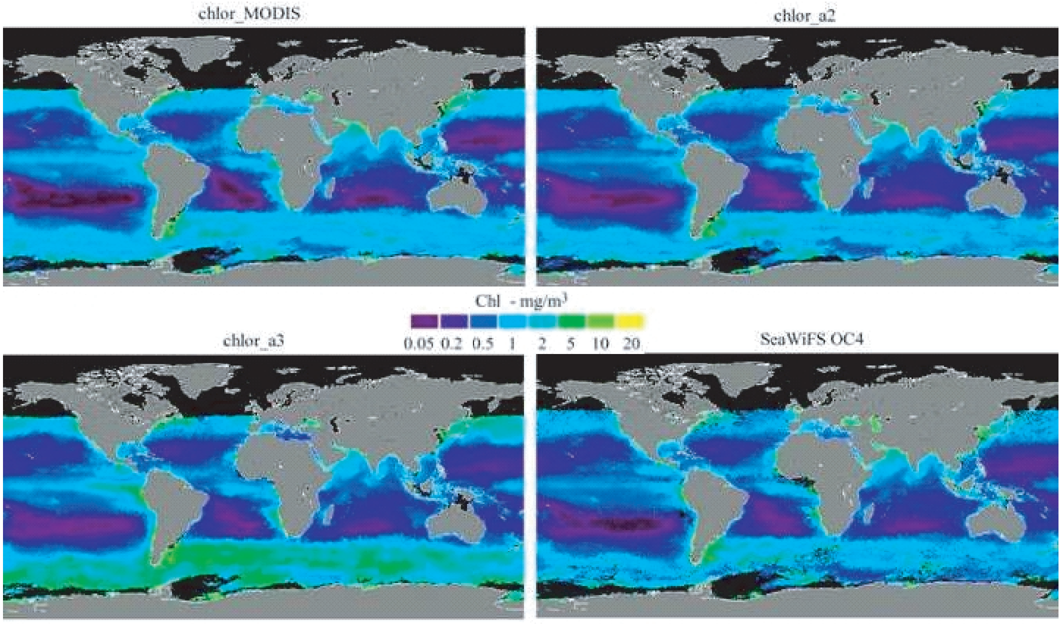


FIGURE 2-9 Comparison of MODIS and SeaWiFS chlorophyll products for December 2000. Images depict the monthly average chlorophyll concentration based on four different algorithms; chlor_MODIS, chlor_a2, and chlor_a3 are MODIS algorithms. The chlor_a2 algorithm is designed to be similar to the SeaWiFS OC4 algorithm. Chlorophyll concentration (mg $m^{-3}$) is a measure of phytoplankton biomass produced by ocean color sensors.

tors supported to define strategies for merging data from difference sensors. The SIMBIOS project represents a model for generating one class of FCDRs and an associated family of derived TCDRs.

Ocean color FCDRs are the calibrated radiances derived by ocean color sensors together with the supporting in situ data used to calibrate the radiances. The Marine Optical Buoy (MOBY) located off Lanai, Hawaii, is currently used by several ocean color missions (U.S., Europe, and Japan) for vicarious calibration.

### Lessons Learned

- Algorithms will change over time but will always rely on carefully calibrated, top-of-atmosphere radiances. It is most important to assure that the top-of-atmosphere radiances are well calibrated, geolocated, and preserved for future generations.
- Ocean color radiance data will require an ongoing program of vicarious calibration whereby radiances are compared with in situ measurements.
- It is better to produce a modest number of ocean color variables (e.g., chlorophyll, colored dissolved organic material, suspended sediment) that enjoy wide acceptance than to create a large number of variables that have no community support.
- Input from the science community regarding data quality and ways to improve the quality can be received at yearly workshops. Special workshops can be held to select consensus algorithms for producing the CDRs.

## SEA ICE

Sea ice time series derived from multichannel passive microwave data are among the most consistent and longest continuous satellite-derived geophysical records, extending almost three decades. Data from passive microwave radiometers have been available since December 1972 with the launch of the Nimbus 5 Electrically Scanning Microwave Radiometer (ESMR). In 1978 the Scanning Multichannel Microwave Radiometer (SMMR) operated onboard Nimbus 7 and provided data until 1987, followed by the successive Defense Meteorological Satellite Program Special Sensor Microwave/Imager (DMSP SSM/I) sensors. The Advanced Microwave Scanning Radiometer (AMSR) will continue this relatively long history of polar remote sensing. The greater spatial resolution, additional spectral channels, and enhanced system performance of AMSR will further contribute to polar studies through the generation of improved and additional polar ocean products.

Among the many sea ice variables derived from satellite passive microwave imagery, ice extent (the area within the ice-ocean margin) is the one parameter whose variability and trends are most firmly established (Figure 2-10). Two of the most widely used passive microwave sea ice

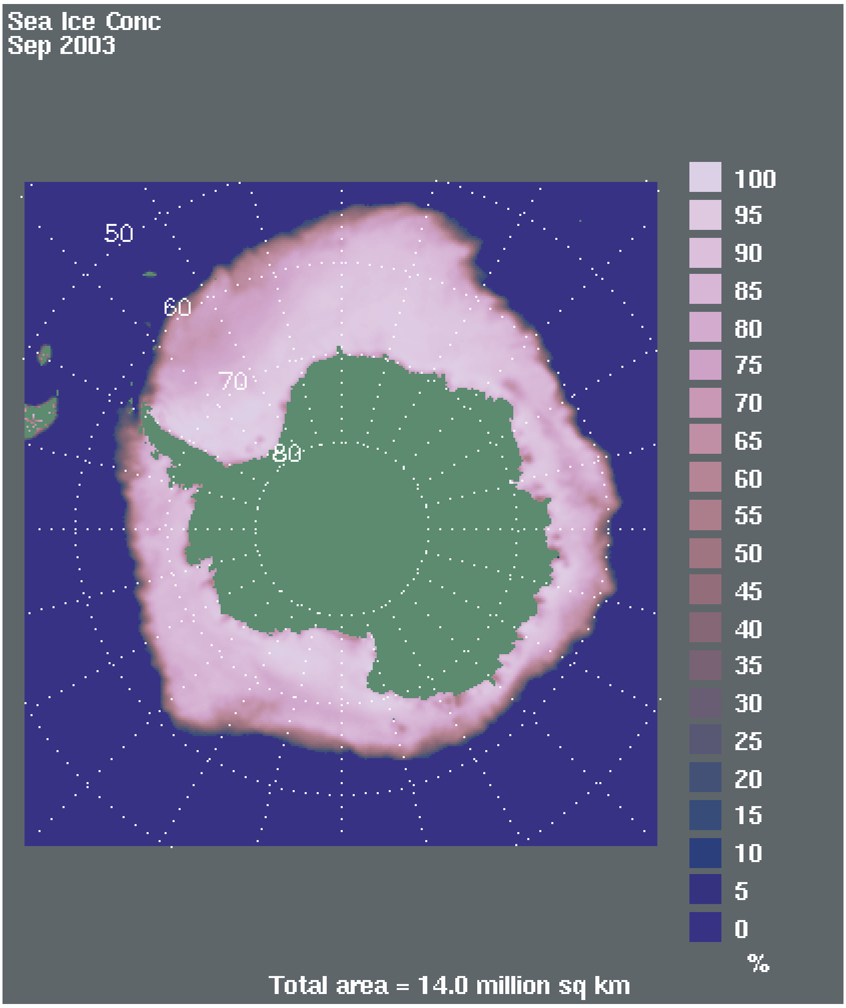


FIGURE 2-10 Mean September 2003 sea ice concentration in the Southern Ocean around Antarctica, based on satellite passive microwave data. SOURCE: Courtesy Ken Knowles, University of Colorado.

algorithms are the NASA Team (NT) and the Bootstrap (BS) algorithms. An enhancement of the NT sea ice algorithm, called NT2, was recently developed and helps overcome the problem of low ice concentration biases associated with surface snow effects that are particularly apparent in the Southern Ocean.

All the sea ice algorithms however suffer when melt ponds form on the sea ice, as the flooded floes appear as open water to the passive microwave instrument. Thus, false underestimation of ice concentration can be expected during summer melt, although in terms of albedo and heat fluxes, the passive microwave estimate may be more meaningful. In addition, special "weather" filters and land and ocean boundary filters are typically applied to the passive microwave data to avoid false ice as a result of weather and the spill over effects of the land in coastal regions of a geographic mask.

Analysis based on SMMR data found a slight negative trend in Arctic sea ice extent from 1978 to 1987 (2.4 percent per decade). Data from the subsequent SSM/I has provided the basis to follow up the SMMR trends and has revealed a greater reduction in Arctic sea ice area and extent during the SSM/I period: decreases from 1987 to 1994 were ~4 percent per decade compared with ~2.5 percent per decade from 1978 to 1987. Large interannual variability, coupled with the brevity of the individual SMMR and SSM/I records, compelled researchers to produce longer time series for more robust trend estimation.

The merging of SMMR and SSM/I data involves satellite intercomparison and radiometric adjustments based on the six-week overlap period in 1987 when both sensors operated. Two independent analyses of merged SMMR–SSM/I data established the trend in arctic ice area and extent (1978-1995) to be about $-3.0 \times 10^5$ km$^2$ per decade, corresponding to ~3 percent per decade. The consensus is that the annual sea ice extent in the Northern Hemisphere has shown a steady decline of ~3 percent per decade from the late 1970s to the late 1990s. Overall, the Southern Hemisphere shows a slightly positive trend of 0.17 percent per decade (1978-2000), although the trend is insignificant. Because the overlap period between the various sensors (e.g., SMMR, SSM/I) is at times only one month, the error due to changes in sensor characteristics, calibration, and thresholds for the 15 percent ice edge may be significant.

## Lessons Learned

- The algorithms developed for SMMR at NASA Goddard Space Flight Center and SMMR products were only transferred to the National Snow and

Ice Data Center more than halfway through the Nimbus 7 mission, limiting their climate value.

- Early planning for the transition to sea ice products from SMMR to SSM/I is crucial for satellite intercomparisons and for creating more reliable records.
- SMMR ice extent and concentration products and SMMR-SSM/I gridded brightness temperatures were designated as Polar Pathfinders. This funding was crucial to the development, release, and distribution of a time series of sea ice data.
- The Polar Distributed Active Archive Center (DAAC) User Working Group (PODAG) provided valuable oversight and guidance to the NSIDC in the planning and implementation of sea ice products from its inception in 1991 to the present with 20 meetings held to date.
- There are several standard algorithms for deriving sea ice concentration, and no single algorithm is "best" for all user communities.

## SUMMARY OF LESSONS LEARNED

The preceding examples were examined to identify redundancies and develop a set of lessons learned that represent the combined wisdom gained in all of these earlier projects. In summary, the committee finds 12 attributes that contributed to the success of these examples:

1. **CDRs must be accepted by the community**. This is the most fundamental characteristic of a CDR. This acceptance is evaluated after a period of community use to correct unforeseen problems, identify new applications, and apply algorithm improvements.
2. **Long-term funding for CDR production is required**. This is a need common to all CDR activities, and it is clear that without stable funding it is impossible for a CDR to be created or sustained long enough to be of value. Historically this funding has come from a variety of sources; but insofar as federal agencies are concerned, NASA and NOAA have funded these types of data analyses.
3. **Science advisory review provides critical oversight**. Without science input it is difficult to create or identify a CDR and to maintain its production and utilization.
4. **User workshops allow broad community input**. One method for generating user input is by initiating workshops.
5. **Long-term trends are often unreliable because of input changes, calibration errors, algorithm evolution, and instrument changes**. Again,

science guidance is needed to assess errors in long-term time series and correct or compensate for physical changes in long time series.

6. **CDR stewardship is needed throughout the entire process**. A critical requirement is that the science overview and involvement is end-to-end and that oversight will be needed continuously for the duration of the period to ensure that the CDR is generated and maintained.
7. **Reprocessing and reanalysis is an ongoing activity**. The development of a CDR is often an evolutionary process and repeated reprocessing of the entire input dataset is often necessary. The CDR algorithm will change over time as measurement hardware and software and understanding of the measurement process evolve.
8. **Raw data must be archived**. The reprocessing requirement for satellite data means that the raw satellite data must be archived for future reprocessing.
9. **Multiple, independent processing can provide important insights into the CDR**. Multiple efforts provide information on the precision of a CDR, and biases based on one processing method may not be obvious without comparison with other independent methods.
10. **Data management and oversight are critical components of a successful program**. CDR datasets should be managed in a way that allows easy, meaningful access by users of varying technical sophistication. Web access could include browse images, read software, subsetting tools, and online plotting
11. **Validation and overlap of successive satellite missions is critical for developing consistent CDRs over time**. Satellite measurements are by their very nature “remote” and thus in situ observations are needed to validate remotely sensed data and monitor sensor degradation, while overlap is needed to reduce satellite biases
12. **Orbital and instrument decay need correction for consistent long-term CDRs**. Another source of change that requires in situ validation measurements is caused by orbital and instrument degradation decay. This can change the atmospheric path length for each instrument, alter the Sun-Earth sampling geometry, or introduce aliasing of the diurnal cycle.

# 3

# Elements of a Successful Satellite CDR Generation Program

Developing a successful Climate Data Record (CDR) generation program poses many challenges because of the varied uses of climate data, the complexities of data generation, and the difficulties in sustaining the program over extended periods of time. Many of these challenges are not unique to a climate data record program, but are common elements faced in creating most long-term science programs (e.g., NRC, 2002). The previous chapter described the experiences from programs where satellite data are the primary global data records. These represent only a fraction of such experiences and yet the investment to date in just these programs alone has been enormous. This chapter outlines 14 key elements that the committee believes are important for the successful generation of CDRs. The first section discusses four elements related to organization, emphasizing the importance of having a high-level coordinating body within NOAA, broad involvement of stakeholders, and mechanisms for review (and redirection) by the science community. The second section presents elements related to the generation and stewardship of the fundamental CDRs (FCDRs) and the thematic CDRs (TCDRs). The final section discusses elements related to sustaining a CDR generation program.

## ORGANIZATION

In devising a program for generating CDRs, NOAA would benefit greatly from developing an organizational framework that includes mechanisms for providing scientific oversight and advice, encouraging feedback from user communities, and allowing opportunities to redirect the program based on advice and feedback. The task of generating CDRs is ambitious, but NOAA does not need to accomplish everything on its own because CDR expertise lies within other agencies, academia, and the private sector as well. Therefore, in developing a management and administration component, NOAA can help to ensure success if it involves scientists with a vested interest in CDRs, finds committed people to generate the CDRs, develops technical and science support for broad involvement, and creates teams that are reviewed and renewed regularly, consisting of NOAA personnel, outside scientists, industry, and others. The following elements lay out a framework of responsibilities that should be accomplished and the kinds of groups needed for each role. The committee believes that the following 14 elements will help NOAA to create a successful CDR program.

**1: A high-level leadership council within NOAA is needed to oversee the process of creating climate data records from satellite data.**

A leadership council of NOAA management personnel would receive input from the climate research community and other stakeholders through an advisory council of internationally recognized climate experts and would have the authority to approve plans and commit resources to generate the CDRs. The leadership council would adopt responsibility for overall stewardship of the CDR program, determining whether the FDCRs and TCDRs are effective and if not, working with partners to correct problems.

**2: An advisory council is needed to provide input to the process on behalf of the climate research community and other stakeholders.**

An advisory council would advise the leadership council concerning the generation of CDRs. The function of the advisory council would be to

- recommend and prioritize the variables that are developed into TCDRs;
- oversee the calibration of FCDRs and validation of TCDRs;
- evaluate proposed new TCDRs and refinements of existing TCDRs as measurement capabilities improve or scientific insights change over time;
- review the utility and acceptance of TCDRs and recommend the elimination of those that are not successful; and
- review and oversee NOAA's stewardship of the CDR program.

Members of the advisory council should include participants from within NOAA, other federal agencies, academia, and industry. Given the importance of the advisory council, compensation for their services would be appropriate. Respondents to the community survey and attendees at the workshop were all very clear in stating that without some form of compensation, NOAA will not be able to attract the best scientists for the advisory council, which in turn could limit the success of the program.

### 3: Each FCDR should be created by an appointed team of CDR experts.

The expertise needed to create and validate FCDRs is different from that needed to produce TCDRs. The FCDR teams should involve engineers and spacecraft specialists, as well as representatives from the thematic science teams, because feedback from the generation of TCDRs is essential. NOAA already has in-house staff (either within NOAA centers or in NOAA-funded cooperative institutes) to create FCDRs from satellite data, but history indicates that staff not familiar with the product may not be aware of problems with the data. Therefore, NOAA should involve scientists from government, academia, and the private sector who have expertise in how the FCDRs are to be used, and are familiar enough with the variables to know what values are reasonable. The main functions of the FCDR teams include instrument monitoring and production of FCDRs, and they should document their work extensively so that future generations can easily assess and understand what they have done. Proper documentation is also important for users, because all FCDRs will have limitations and errors, and these must be well documented.

The size of an FCDR team will be based on financial resources, but guidance from studies examining why projects succeed (Standish Group, 1999) indicates that smaller teams are more likely to achieve their goals.

### 4: Science teams should be formed within broad disciplinary theme areas to prescribe algorithms for the TCDRs and oversee their generation.

While the FCDR generation teams are focused on creating accurate and precise radiance measurements, the program will not be successful unless scientists are actively utilizing the FCDRs to create TCDRs. The committee recommends the formation of thematic science teams within broad disciplinary theme areas (Table 3-1) to oversee the generation of TCDRs and evaluate TCDRs created by outside groups. These teams historically have formed around specific technologies, but this may not be the best approach in the long term; for example, users of active microwave remotely sensed data have not typically interacted with users of passive microwave, visible, and infrared remotely sensed data. An alternative approach would be to form teams around science themes as previously recommended by the NRC

TABLE 3-1 NASA and NOAA Thematic Groupings of Space-based Observations

| | |
|---|---|
| 24 NASA/Earth Observing Systems (EOS) measurements are divided into 5 groups.<br>1. Atmosphere<br>2. Land<br>3. Ocean<br>4. Cryosphere<br>5. Solar radiation | National Polar-orbiting Operational Satellite System (NPOESS) divides its 61 Environmental Data Records (EDRs) into 7 "parameter" groups.<br>1. Key parameters<br>2. Atmospheric parameters<br>3. Cloud parameters<br>4. Earth radiation budget parameters<br>5. Land parameters<br>6. Ocean and water parameters<br>7. Space environmental parameters |

(2000b). The committee believes that establishing CDR science theme teams, not instrument teams, will be essential to the goal of generating successful CDRs.

These teams should be led by recognized scientists who are actively engaged in research generating, utilizing, or validating the TCDRs. The team leaders should be ultimately responsible for the quality of the TCDRs, and they should provide the advisory council with periodic updates. The thematic teams should include research scientists funded by or employed by NOAA and other agencies, organizations, or private sector companies who use the data, and they should have some representation from the FCDR teams. The TCDR science teams should be competitively selected, with limited (but renewable) terms. Team members should also be compensated for their work, similar to the advisory council. Funding for the thematic teams should be broadly based and could be orchestrated by the Climiate Change Science Program (CCSP) or a partnership of federal agencies (see Chapter 5).

## CREATING FUNDAMENTAL CLIMATE DATA RECORDS

The distinction between FCDRs and TCDRs[1] is important and to a large extent unique to the generation of CDRs from satellite data. The heart of NOAA's efforts to create a successful CDR program lies in the creation of the FCDRs. It is vitally important that NOAA appreciate the steps required to create the FCDRs, as the success of the CDR program hinges on creating reliable, consistent, and stable FCDRs.

---

[1] In some cases, such as microwave brightness temperatures, the FCDR is the TCDR.

### 5: FCDRs must be generated with the highest possible accuracy and stability.

As explained in Chapter 1, the FCDRs are the time series of calibrated signals (e.g., top-of-atmosphere radiances, brightness temperatures, radar backscatter) for a family of sensors (e.g., Advanced Very High Resolution Radiometer [AVHRRs]), together with the ancillary data used to calibrate them. In some cases, extensive in situ datasets might need to be included as part of the FCDRs if these ancillary data are needed to regenerate the time series at a future date. These in situ datasets also are occasionally improved, so efforts are needed to ensure that the most up-to-date in situ dataset is utilized for the FCDR analysis. The FCDRs will be used to create a variety of TCDRs for various disciplines. Where possible, NOAA should adhere to the Global Climate Observing System (GCOS) climate-monitoring principles (Box 3-1). In some cases it will not be possible to meet all the requirements. For instance, Requirement 1, "Complete sampling within the diurnal cycle (minimizing the effects of orbital decay and orbit drift) should be maintained," will not always be met when using historic polar orbiter data to create CDRs. Clearly TCDRs should be developed, to the extent possible, to account for diurnal effects, but in some instances it will not be possible. A TCDR record tied to a narrow segment of the diurnal range might still prove valuable decades hence, when the physical understanding of some processes become better understood.

### 6: Sensors must be thoroughly characterized before and after launch, and their performance should be continuously monitored throughout their lifetime.

Verification of instrument performance requires a comprehensive understanding of the physics behind the measurement. Satellite sensors must have a thorough prelaunch characterization and the ability to measure important instrument properties on orbit, including the ability to calibrate the sensor after launch. Procedures should be in place to monitor sensor performance in near real time.

The satellite and sensor engineers, working with the FCDR team, should do the best job possible to ensure the accuracy of the calibrated data used to create the FCDRs. An integral part of this process is a full characterization of instrument performance and stability, and continuous monitoring of the observing system for changes in the sensors. Since most of NOAA's operational satellites were created as weather rather than climate platforms, this

## BOX 3-1 CLIMATE MONITORING PRINCIPLES

The international GCOS Panel (GCOS, 2003) developed the following principles for monitoring climate variables from satellites, which the committee endorses and suggests that NOAA implement:

1. Constant sampling within the diurnal cycle (minimizing the effects of orbital decay and orbit drift) should be maintained.
2. A suitable period of overlap for new and old satellite systems should be ensured for a period adequate to determine intersatellite biases and maintain the homogeneity and consistency of time-series observations.
3. Continuity of satellite measurements (i.e., elimination of gaps in the long-term record) through appropriate launch and orbital strategies should be ensured.
4. Rigorous prelaunch instrument characterization and calibration, including radiance confirmation against an international radiance scale provided by a national metrology institute, should be ensured.
5. On-board calibration adequate for climate system observations should be ensured and associated instrument characteristics monitored.
6. Operational production of priority climate products should be sustained and peer-reviewed new products should be introduced as appropriate.
7. Data systems needed to facilitate user access to climate products, metadata, and raw data, including key data for delayed-mode analysis, should be established and maintained.
8. Use of functioning baseline instruments that meet the calibration and stability requirements stated above should be maintained for as long as possible, even when these exist on decommissioned satellites.
9. Complementary in situ baseline observations for satellite measurements should be maintained through appropriate activities and cooperation.
10. Random errors and time-dependent biases in satellite observations and derived products should be identified.

step is particularly relevant for NOAA to address. Changes in satellite characteristics, such as orbital drift, system calibration, sensor degradation, and instrument failure compromise the ability to create high-quality, consistent CDRs over time. NOAA should assure that procedures are in place to monitor the observing system for irregularities that could corrupt the long-term value of the FCDRs. Such a diagnostic scheme will allow the FCDR team to distinguish between artificial changes related to the observing system and real changes due to climate.

It is essential that a period of overlap between successive sensors be incorporated into launch schedules to assess and correct for differences between sensors, thus assuring continuity of the long-term climate records. Over time as new systems are developed the FCDR team also should be charged with determining what impact the new sensor has on observations. NOAA must ensure that introduction of new observing techniques does not result in artificial changes in the FCDR. The committee recommends that NOAA continue the Polar Operational Environmental Satellite (POES) performance-monitoring activities with some modifications (in italics), namely,

- inventorying data, filling in missing periods with other data in NOAA archives *provided that the other data is fully compatible with the FCDR (that is, other data cannot cause spurious trends or variability);*
- converting internal satellite quality data into useful information for end users; and
- providing easy-to-use Web-based tools that link end user quality control information to more detailed instrument information.

**7: Sensors should be thoroughly calibrated, including nominal calibration of sensors in-orbit, vicarious calibration with in situ data, and satellite-to-satellite cross-calibration.**

For the FCDRs to be useful for future applications and to maintain a consistent record based on multiple satellites over several decades, the sensors must be well calibrated. Sporadic efforts have been mounted by various groups to correct existing radiance biases, but this function should be centralized, standardized, and routinely performed by NOAA. The GCOS panel recommended that steps be taken to make radiance calibration, calibration monitoring, and satellite-to-satellite cross-calibration of the full operational constellation a part of the operational satellite system. The International Satellite Cloud Climatology Project recognized that calibration is an iterative process, and they defined three types of calibration: (1) nominal calibration; (2) vicarious calibration; and (3) satellite intercalibration.

Nominal calibration involves determining the calibration of a single sensor on a single platform. Prelaunch calibration is a standard procedure, but the calibration of the instrument must be monitored in orbit, and if necessary adjusted. Depending on the instrument, this may involve onboard calibration lamps or other means of onboard calibration, but often there is uncertainty associated with the “known source” onboard the spacecraft.

Alternative methods involve maneuvers to view deep space, or the Sun or Moon as a constant radiation source by which to monitor the stability of an instrument during its orbital lifetime. This does not necessarily replace the need to calibrate the instrument, but it does provide a method of monitoring instrument stability over time (Figure 3-1).

Vicarious calibration is accomplished by measuring a known target

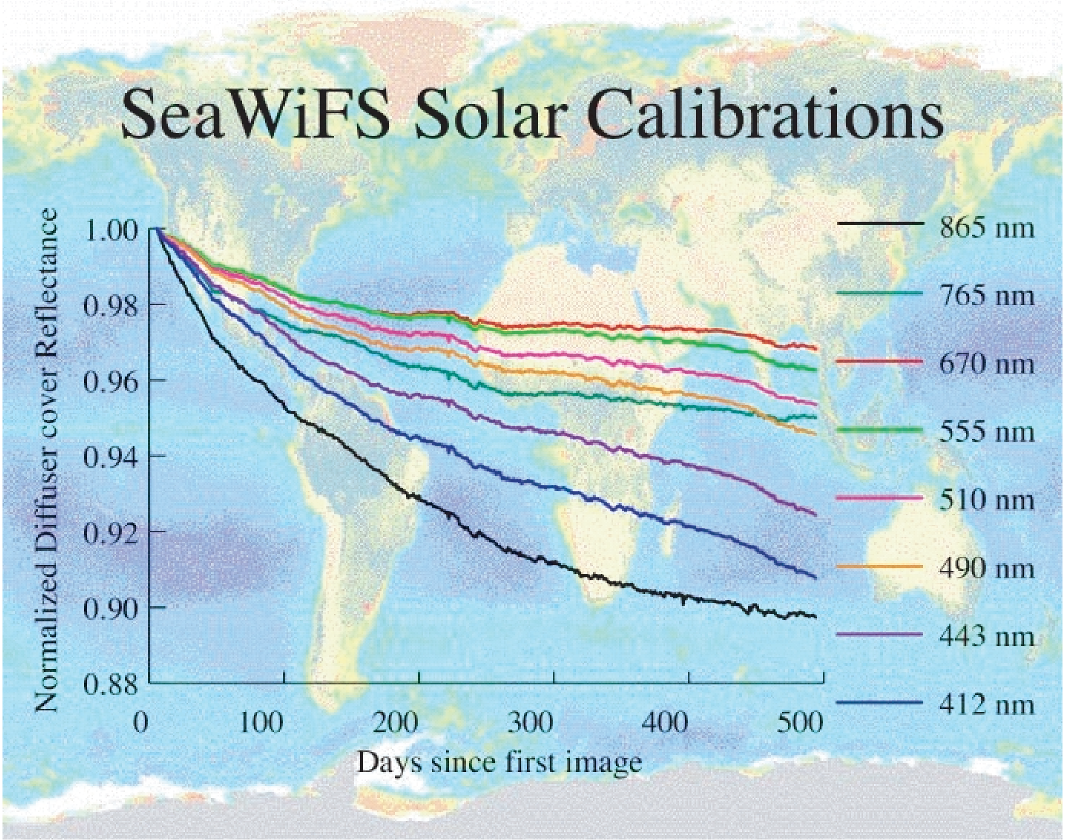


FIGURE 3-1 The Sea-viewing Wide Field-of-View Sensor (SeaWiFS) detectors are monitored by periodically viewing the Sun through a diffuser plate. In this figure measured detector responses in the eight channels have been normalized to their initial values to indicate how the responses have changed over time. This information is used to correct the radiance measurements for changes in the sensitivity of the detectors. The absolute calibration of SeaWiFS is done vicariously. SOURCE: NASA Goddard Space Flight Center.

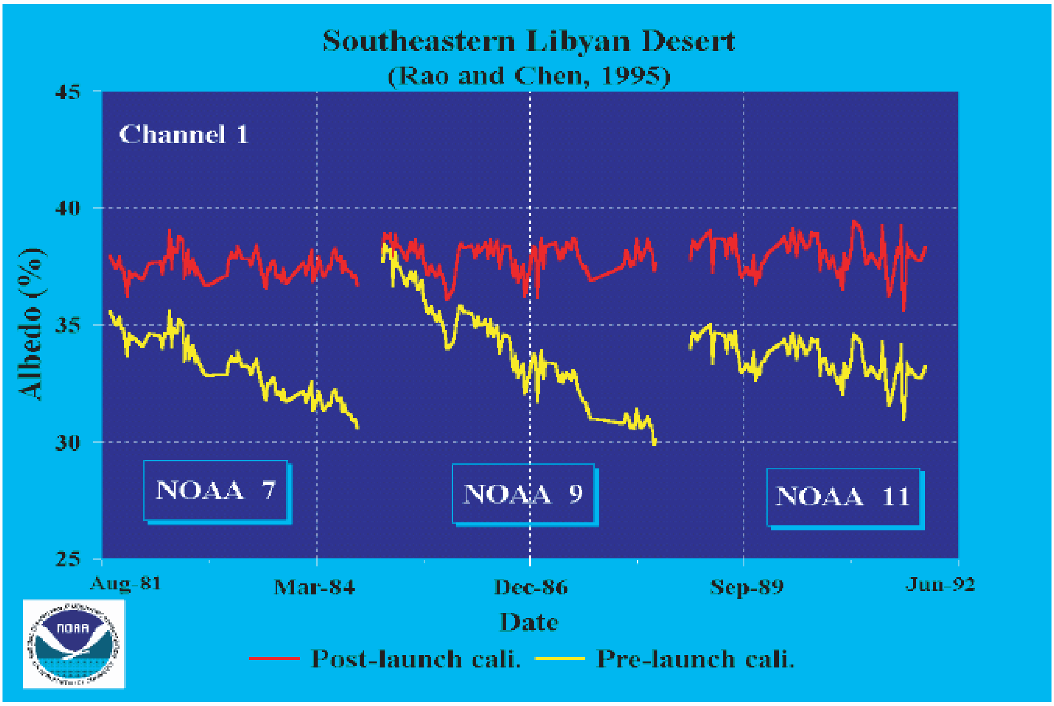


FIGURE 3-2 Comparison of albedo measurements with and without vicarious calibration. Without proper postlaunch calibration, spurious trends in the data can occur. SOURCE: Rao and Chen, 1995.

(e.g., desert) or comparing the satellite signal with simultaneous in situ balloon, radiosonde, or aircraft measurements. The satellite calibration is then adjusted, after correcting for atmospheric effects, to agree with the target. All CDRs should be vicariously calibrated at regular intervals, regardless of onboard calibration.[2] Without vicarious calibration, sensor data can drift over time (Figure 3-2). Vicarious calibration also can serve as an additional means for monitoring instrument stability. In addition, the future experience gained in the vicarious calibration of well-calibrated satellite

---

[2] The GCOS Climate Monitoring Principles (see Box 3-1) do not recommend vicarious calibration, but this committee believes it is an important calibration-monitoring technique.

instruments will undoubtedly prove to be useful in extending CDRs backward in time.

The third form of calibration is satellite intercalibration. This involves adjusting several same-generation instruments to a common baseline, such as calibrating all SSM/I sensors to one baseline. This is particularly important for long-term studies, as each sensor will have slightly different baselines even if built to the same specifications. Processes that lead to slight differences between sensors include uncertainties in prelaunch instrument response characterization, deposition of contaminants on sensors, drift in the satellite orbit and altitude, electronic noise, and cross-talk. Each bias source must be addressed, preferably with reference to known physical mechanisms. The best assurance of sensor intercalibration is to guarantee a period of overlap between successive sensors. Even though we are in the fifth decade of satellite studies, we have more to learn about radiometric calibration and data analysis. Overlap is a crutch to help us to continue to produce useful results while we are still learning how to use the data. Without intercalibration, long-term trends in CDRs likely will lead to spurious trends (Figure 3-3).

## CREATING THEMATIC CLIMATE DATA RECORDS

Although the FCDRs represent NOAA's long-term legacy, the majority of users will use the thematic CDRs (Box 3-2). Since NOAA cannot create all possible TCDRs, mechanisms must be in place to select an appropriate number of TCDRs to generate. Once created these TCDRs must have rigorous validation and estimated uncertainty levels.

> **8: TCDRs should be selected based on well-defined criteria established by the advisory council.**

The advisory council should begin by identifying thematic areas and establish science teams to recommend generation or acceptance of existing TCDRs in each discipline area. A number of possible "themes" can be defined (e.g., Table 3-1). One simple method that NOAA could employ to select the TCDRs is to select team members based on their ability to generate data records; the selection of the science teams and TCDRs would automatically determine which variables are produced. An alternative approach would be for NOAA to issue a Request for Information (RFI) to formally solicit ideas rather than proposals as a first step. NASA did this a few years ago to determine which future satellite missions should be in-

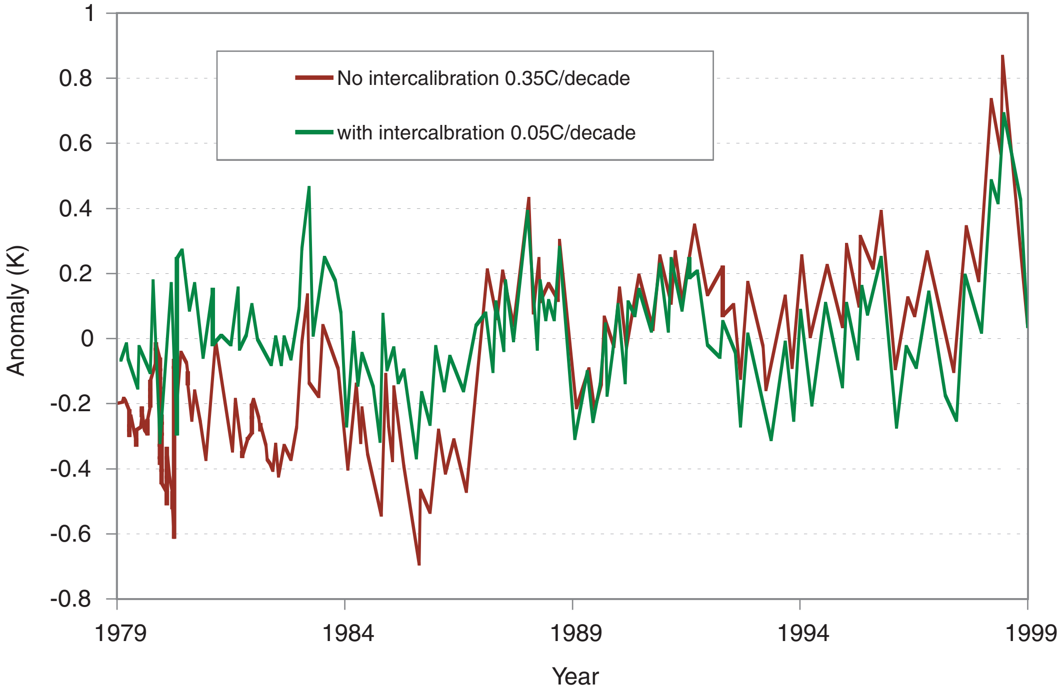


FIGURE 3-3 Example of satellite intercalibration from monthly global Microwave Sounding Unit (MSU) Channel 2 anomalies. If sensors on different platforms are not calibrated with one another, spurious trends can appear. SOURCE: NASA Global Hydrology and Climate Center.

## BOX 3-2 NEEDS AND USES OF CDRS

With an increased appreciation for the impact of weather and climate on daily activities (e.g., NRC, 2003a), society's need for climate data grows rapidly. Many of the major users and uses of climate information have been illustrated in the recent NRC reports *Making Climate Forecasts Matter* (NRC, 1999c) and *Fair Weather: Effective Partnerships in Weather and Climate Services* (NRC, 2003b). In brief, some of the sectors using CDRs include education, research, water resources, energy, agriculture, forestry, transportation, defense, health, insurance, recreation and tourism, manufacturing, and retail. It is likely that CDRs will become more valuable and more used over time, requiring NOAA to do the best possible job to create reliable and consistent CDRs.

cluded in its Earth science strategic plan. The RFI was administered very much like a request for proposals. Mission concepts were submitted and panels were assembled to review the concepts.  An RFI might be issued to solicit recommendations for TCDRs. The council should subsequently work with thematic team leaders to arrive at a prioritized list of TCDRs. Variables chosen for TCDR development should address key questions about the climate system, leading to clear improvements in (1) scientific understanding of the climate system, enabling climate variability and change to be better documented; (2) projections for future climate states; (3) regional, national, and international climate assessments; and/or (4) the nation's ability to respond to climate variations. The benefits of the TCDRs should address applications of climate information as well as the scientific understanding of the climate system. The criteria for selection of TCDRs should also include the technological readiness of the record. In most instances NOAA likely will select one TCDR for a given parameter, but in certain situations NOAA may fund a baseline TCDR and encourage the creation of other TCDRs for comparison. In exceptional situations NOAA may even fund several TCDR efforts for the same parameter (recall the lessons learned from the MSU temperatures).

**9: A mechanism should be established whereby scientists, decision makers, and other stakeholders can propose TCDRs and provide feedback that is considered in the selection of TCDRs.**

Since the TCDRs will be used by scientists, decision makers, and other stakeholders, the process of selecting TCDRs will provide the greatest long-term utility if it is fully documented and open to input from user communities. The science community ultimately will submit proposals to NOAA and other funding agencies for generating TCDRs, and to be truly successful NOAA needs a buy-in from the science community from the start. As noted by Weisberg et al. (2000), "The most successful programs have been those with clearly defined users for the data they produce, which requires early interaction between the scientists responsible for designing the program and targeted data users." When these communities are brought together early to identify their needs, rather than just being asked to approve or comment on what the science community has planned, there is a much greater chance of incorporating the needs and concerns of user communities into planning the TCDRs, and thus the opportunity to design a more satisfactory program that enjoys long-term user support.

Open science meetings have been used with great success to assist in determining programmatic priorities, and NOAA could convene such conferences on the creation of TCDRs. These meetings could be held in conjunction with conferences held by other organizations, such as the American Meteorological Society or the American Geophysical Union, which would open the conference to broader attendance and reduce costs. Community surveys are also useful in generating a list of needs from various decision makers and user groups.

### 10: Validated TCDRs must have well-defined levels of uncertainty.

The process of validating a TCDR derived from satellite measurements is not simply a matter of "ground truthing" a satellite-derived product. It is the process of establishing uncertainty levels for the TCDR based on principles of error propagation and comparisons with independent correlative measurements. The identification, quantification, and minimization of biases and errors helps users understand how much confidence they should have in the TCDRs and whether the data are appropriate for their applications. More specifically, the uncertainty associated with a TCDR determines how much of a trend can be detected with the record.

An understanding of the measurement error structure is critical. Error propagation from raw data to the final derived product must be understood. Instruments and data processing will introduce systematic artifacts in the data. Error sources are not necessarily Gaussian, nor are errors uncorrelated spatially and temporally. The steps involved in deriving a geophysical variable (e.g., geolocation, calibration, and correction for atmospheric effects) must be identified. Each of these requires algorithms specific to the task and will introduce uncertainties in the geophysical product. Knowing the magnitude of the processing task will help to choose the right algorithms and ultimately quantify the uncertainty in the final geophysical product.

In discussing CDR uncertainty it is important to distinguish between geophysical quantities and indices. Geophysical quantities, the products of the TCDRs, have in principle an uncertainty that can be determined for a given time and space scale. Estimates of the quantity will lie within the uncertainty if measured by any observing system capable of measurements for the particular time and space scale. Indices, on the other hand, are defined by the instrumentation and the algorithm used to construct the index. Normalized Difference Vegetation Index (NDVI) is a typical example. There may not be correlative measurements for comparison, and thus uncertainty estimates are strictly based on instrument and algorithm properties. Nevertheless, indices provide valuable insight into the workings of the climate system that as yet defy predictive capabilities through modeling. Ultimately, understanding will form links between indices and physical quantities. The significance of the latter is that with such understanding, historic data can be reanalyzed, given the new insight, to extend CDRs backward in time.

**11: An ongoing program of correlative in situ measurements is required to validate TCDRs.**

The process of validating a TCDR is an ongoing activity. There should be a program of in situ measurements established and maintained for this purpose. In some cases assessments to determine the amount and location of in situ data needed to estimate uncertainty may have to be performed. NOAA should also consider incorporating other satellite data as another source of correlative measurements. Data from geostationary satellites could provide information not available from the polar-orbiting satellites, in turn reducing uncertainty in the data. In particular, data from geostationary satellites can help with sampling the diurnal cycle. NOAA should also consider examining several algorithms and techniques to determine whether there is agreement among multiple routes to obtaining the same results. Blended products may be used to develop TCDRs, although this may make it more difficult to understand how to reconstruct datasets and how to account for errors.

## SUSTAINING A CDR PROGRAM

Lessons from the past suggest that sustaining a long-term research program is difficult and is rarely achieved. In the short term NOAA will create successful FCDRs and TCDRs by following the steps discussed in this chapter. Since the CDR program is conceived with a long-term vision, it is important to recognize several elements related to sustaining the CDR program. Producing ongoing climate products (such as CDRs) is contingent on having stable funding both for obtaining data and for conducting the sustained scientific research that will necessarily underlie the production of CDRs. NOAA has traditionally been a mission agency with responsibility for weather monitoring and prediction and support of this mission through satellite operations, and data archiving and management. In comparison, a CDR generation program requires efforts above and beyond NOAA's traditional role in weather forecasting.

NOAA has long supported climate research, but much of the research it supports is done in-house or through research programs that are generally located in research units (e.g., Office of Atmospheric Research Laboratories) in the agency rather than operational units like the National Environmental Satellite, data, and Information Service (NESDIS). Research on CDRs must be directly linked with the production of CDRs and must be an integral part of the CDR program. Climate, unlike weather prediction, is a subject with widely distributed expertise across disciplines within academic, government, and private industry sectors. CDRs will inevitably serve the needs of multiple highly disparate communities ranging from scientists to individuals with responsibility for public policy and management and will draw on a changing array of satellites and sensors, requiring overlap and intercomparison of measurements. To succeed in the CDR mission NOAA will need to obtain a new higher level and probably new sources of sustained financial support for this valuable mission that will extend over many decades. If long-term funding is not sustained, history suggests that this the program will fail.

**12: Resources should be made available for reprocessing the CDRs as new information and improved algorithms become available, while also maintaining the forward processing of data in near real time.**

Over time, errors will be uncovered in the FCDRs and TCDRs, new algorithms will be developed for the TCDRs, and new technologies will be available. To ensure the success of any long-term program, mechanisms are needed to address deficiencies, correct problems, and ensure continuity, and these actions require adequate resources. As noted by NOAA's Climate and Global Change Working Group report, "Given the continuing improvement in our understanding of climate observations and the need for long time series, reprocessing is a hallmark of every climate observing system." As our understanding of the climate and satellite instruments improves, reprocessing could also be useful for extending CDRs back in time; for instance, CDRs created from NPOESS could be extended back to the Television Infrared Operational Satellite-Next Generation (TIROS-N) series, with knowledge that uncertainties in the earlier records may be larger (emphasizing the need for stating uncertainly levels).

In determining the resources needed to generate CDRs it is therefore essential to include the capability (e.g., computer processors, storage devices, personnel) to reprocess the data at periodic intervals. The FCDRs will need to be reprocessed as new information is acquired or better calibrations are made, but these records will eventually become stable. On the other hand, the TCDRs will continue to change indefinitely as new or improved algorithms are developed and improved applications are made of the FCDRs. NOAA should consider development and maintenance of a two-track approach: (1) a commitment from an organization to implement, document, and disseminate (free and open) a Version X of a CDR under the guidance of advisory and science teams; and (2) a funded extramural research program to validate, assess, and provide improvements (upgraded algorithms or blending procedures) on which to base future versions and reanalyses. Mechanisms also should be in place for active data users, on Track 1 to be fully informed about basic findings, progress, and tentative plans for Track 2.

**13: Provisions to receive feedback from the scientific community should be included.**

Community feedback is important in developing a successful program. There should be a continual dialog about utility, quality, and problems between those who make observations and those who use them. This dialog will improve the quality of the data and foster their continuation. Users are the ones exercising the data, and as people use the data, problems are uncovered. Systematic metrics should be collected for all TCDR data streams to determine the utility of the TCDR. If the metric suggests that a TCDR is not being utilized at a high enough level to warrant continuation, it should be scuttled and funds directed to another TCDR. Since the FCDRs will still be generated, the scuttled TCDR could be recomputed at a future date. NOAA should publish acceptable levels of use so that users know the cutoff point (decommissioning) of TCDR production.

Regular user workshops are a meaningful way to convert user comments into new policies and procedures. NOAA could convene a workshop of the investigators who are using TCDRs in a particular theme area. This would be an “open science meeting” in which scientists would share their findings (e.g., give science talks) and also discuss limitations and recommendations for improving the TCDRs. Regular opportunities for dialog between the scientists and decision makers or other users of the data are also valuable.

**14: A long-term commitment of resources should be made to the generation and archival of CDRs and associated documentation and metadata.**

NOAA cannot afford to create all the desired CDRs; and thus, careful consideration should be given to the process of prioritizing the list of TCDRs. Once priorities are established, the resources needed to process them should be determined as realistically as possible, and then a long-term commitment must be made to provide the resources to sustain them and archive them, which is covered in more detail in the next chapter.

Stable support is an essential characteristic of a successful CDR generation program; thus inflationary increases should be programmed into budget planning. Operating cost increases or other factors often require flexibility and adjustments by the system operators to maintain data flow while long-term solutions are sought.

There should be a commitment to support research that utilizes the TCDRs. We do not believe that NESDIS should be responsible for supporting the research of all or even most science team members. There must be a commitment made by other agencies (e.g., NASA, NSF) and other NOAA line offices (e.g., National Marine Fisheries Service) to support the research community. The Climate Change Science Program could serve as the organizing entity for ensuring this commitment. Chapter 5 outlines the further details of partnerships that NOAA should explore.

## CONCLUDING REMARKS

How does a CDR become a community standard? How can NOAA ensure that the CDRs are responsive to user needs? The elements of a successful CDR generation program described above would assure this over time, because the steps outlined involve a community of people who use the records and allow them real input into their creation. The involvement of the climate research community will not happen simply by producing CDRs and making them available. It must be fostered by support for research specifically involving the CDRs and organized meetings of funded investigators. It also requires a long-term commitment on the part of NOAA to the generation of the CDRs and to their integrity and validity.

# 4

# Data Management Requirements

Early attention to data management and archiving is a critical step in ensuring the success of a long-term Climate Data Record (CDR) program. Datasets, and ancillary information such as metadata, must be preserved for decades and stored in ways that promote (1) access as data needs change; (2) reprocessing as errors are discovered or calibration is improved; (3) integration as new data products, algorithms, and data technologies are developed; and (4) user-friendly access tools. Climate research problems will inevitably require that scientists use combinations of datasets from many sources: satellite, aircraft, in situ, and even socio-economic data.

It will be critical to facilitate the integration of these multiple types of data. To extract the full scientific and societal value, the data must be available in appropriate formats for scientists, public and private sector decision makers, and managers. Each of these user groups requires different types of information from the original data, which adds complexity and cost to the data management system. Satellite-derived CDRs present special problems stemming from the great volumes of data collected, the multiple sensors and channels involved, and the need in many cases to incorporate surface validation information or to integrate in situ and satellite data sources for the CDR.

Because the ultimate legacy of long-term CDR programs is the data left to the next generation, the cost of data management and archiving must be considered as an integral part of every CDR program. For reference, other large science programs with multidecadal data access and preservation requirements can spend as much as 20 percent of their budget on data management (NRC, 2002). Over time, CDR programs will require significant resources for both continued collection and ongoing management of the data. This chapter discusses data management principles and outlines some key elements to help NOAA to maintain a high-quality CDR program.

## REQUIREMENTS FOR DATA MANAGEMENT OF CLIMATE DATA RECORDS

The preceding chapter highlighted 14 elements that contribute to a successful CDR generation program. An underlying requirement of many of these elements is a sound data plan for stewardship, management, access, and dissemination of CDRs; for instance, fundamental CDR (FCDR) and thematic CDR (TCDR) data obtained from satellites will involve huge volumes of data, but those who need the CDRs will generally not utilize large datasets. A balanced suite of TCDR products will meet the needs of most users, although there will be occasions when portions of FCDR time series, or even raw data, will be needed for independent research; for instance, they also will require other information, such as guides to the data, explanatory metadata based on community standards, fact sheets, frequently asked questions (FAQ) lists, browse images, and searchable archives (by location, time, and phenomenon). To preserve the integrity of the data series and the flexibility needed to constitute new CDRs from the same underlying data, the original data must be stored and available for scientific reanalysis over time. This requires full documentation, including instrument documentation (e.g., CDR information, hardware documentation, firmware documentation, engineering models, and computer models), platform documentation (e.g., overview), and algorithm documentation (e.g., Algorithm Theoretical Basis Developments, “gray” books). Long-term success for the CDRs also will depend critically on sufficient metadata, in standard formats, including metadata fully describing the product line and metadata to discuss CDR limitations and to aid in data management (dataset lineage, version control, and unique identification parameters). The committee cautions that the cost of metadata generation and maintenance can be a significant part of the overall data management costs.

## SYSTEM DESIGN

A carefully designed, efficient data system is fundamental for ensuring success of the CDR program. Since CDRs will be stored, analyzed, and reprocessed in an environment of changing technology and user requirements, the system design should focus on simplicity and endurance. The more complex the system, the more difficult, time-consuming, and costly system upgrades will be. The lessons from Earth Observing System Data and Information System (EOSDIS) and reports from the Standish Group (1999) also suggest that large, complex systems are more prone to failure than smaller, more specialized data systems (Box 4-1). The need to maintain the CDR data systems over long periods will require either large complex systems, medium size and intermediate-complexity systems, or numerous small technically simple but organizationally complex systems. The institutional structure that is used to manage the data will play a critical role in determining who will use CDRs and how they will be used.

The value of standard data management practices cannot be overemphasized because data quality, ease of access, accuracy of documentation, easy problem reporting procedures, and other elements of data management will either promote or hinder current and future utility of CDRs. For instance, the success of the National Center for Environmental Prediction and the National Center for Atmospheric Research (NCEP/NCAR) reanalysis effort compared with the European Centre for Medium-Range Weather Forcasts (ECMWF) reanalysis effort highlights the need for accessible data; the NCEP/NCAR report effort resulted in the second most cited paper in the Earth sciences (E. Kalnay, personal communication). The NCEP/NCAR success is not a result of having superior data but of having data that are more available (Box 4-2).

Advances in computing and networking capabilities are creating new methods of resource sharing, data access, and scientific collaboration. Systems should be designed to permit analysis of multiple or merged CDRs and data-mining strategies.

The following sections elaborate on the general aspects of data management related to data quality, formats, access, policy, and security, and note some features specific to satellite–derived CDRs.

## Data Quality

To ensure that the FCDRs and TCDRs are of the best scientific quality, research scientists who understand the data and the meaning of changes in the data and use the data for their own research should play a major role in data management through active involvement in the development of data products and associated documentation. This practical engagement of scientists also allows for the infusion of their scientific perspectives as actual and prospective data users into the data production process.

The TCDR science teams introduced in the previous chapter must work under a well-defined protocol to implement operating procedures that cover all phases of product development from start to finish, providing standards that all CDRs must meet. The science team approach ensures the communication of all essential production information among the data producers,

## BOX 4-1 THE EOSDIS EFFORT

The NASA Earth Observing System (EOS), and in particular the Aqua satellite and its system of instruments, affords a highly similar collection of instruments being planned for the NPOES Preparatory Project (NPP). Moreover, the scientific products produced by the EOSDIS are forerunners of the environmental data records expected from NPP. Thus, the lessons learned from the development and conduct of the EOSDIS offers the National Environmental Satellite, Data, and Information Service (NESDIS) an unparalleled opportunity to benefit from that experience in planning the production of the National Polar-orbiting Operational Satellite System (NPOESS)/NPP CDRs. The following lessons learned from the EOSDIS illustrate some overarching management considerations for meeting evolving customer needs over decades as the evolution of technology, scientific requirements, and budgetary constraints change:

**Science Investigator Processing Teams.** A programmatic change in early 1998 transferred the responsibility for most EOS data processing from the Distributed Active Archive Center (DAAC) (and the EOSDIS Core System) to EOS science instrument teams and their facilities. These teams included both scientists who generated data and scientists who used TCDRs. This transfer, accomplished through a call for proposals, was a major reason for the success and timely delivery of the EOS standard data products and accounted for the high degree of scientific community acceptance of these products.

**Planned Evolutionary Upgrades.** The EOSDIS has changed significantly in terms of architecture design and implementation since its original planned configuration. Planning for the infusion of evolving information technologies over the course of the development of EOSDIS has made it possible to support the scope of data products and services without compromising the functionality under the ensuing budgetary constraints over the years. If anything, the functionality has expanded to support a much larger community than originally envisaged. It is the expanded functional evolution that has led to the recognition that EOSDIS is a more open and distributed architecture both in terms of science processing and user applications. By adopting the EOSDIS Clearing House (ECHO), costly revised system versions or scrapping of systems and restarting from a clean slate will be avoided for many products. ECHO allowed for a limited open source architecture concept to address the current needs and capabilities as a natural evolution. ECHO supports various searches of the metadata so that individual communities can tailor the user interface to their own needs and access methods.

**Program and Project Management.** Creating widely acceptable CDRs from NOAA operational satellites will be as difficult a science challenge as managing a data information system as complex as the EOSDIS. Garnering the full support of a diverse and broad representation of the science community from the initial proposed concept, plan, scope, and implementation is critical to the success of this NOAA undertaking. Unfortunately, in developing the EOSDIS the science community was not completely supportive from the start, and was unsatisfied with the centralized design approach of an EOSDIS core system with the DAAC selection process and with its role in the scientific processing of higher-level data products and the one-size-fits-all approach. Allowing users to gain ownership of requirements through sponsored workshops to reach community consensus and initiating processes to enable users to prioritize requirements allowed stakeholders to take an active role in the design and thus improve their level of comfort with the EOSDIS core system.

### Lessons Learned

1. The data management system must be designed initially with scientific user input actively sought and given a paramount role in establishing the design criteria. The best technological input from information technology (IT) experts is important, but scientific guidance is also essential.
2. The system cannot be designed once and frozen; provision must be made for it to evolve. At all stages in the process, feedback from the user community must be continually solicited and heeded, and the community must be encouraged to buy in and consider itself a full partner in the effort.
3. Encourage “pull” and discourage “push.” A sure sign of a poor system design is the tendency for the designers and builders to tell the customers that they know what’s best and can predict how customers will use the data.
4. Bigger need not mean better; the idea that several compatible, smaller, and more agile systems may be preferable to one big one.
5. Not all users are alike, and optimal use of the data means ensuring that the system is easy to use for the novice single user and for large groups and experts.

## BOX 4-2 NCEP/NCAR REANALYSIS

The Reanalysis Project is a cooperative effort of the National Center for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR) to produce a 50-year (1948-1997) record of global analyses of atmospheric fields. This effort, which started in 1989, grew out of the need in the research and climate monitoring community for a climate data assimilation system (CDAS) that would be unaffected by changes in numerical weather prediction operational systems. The CDAS Advisory Panel recommended that a long-term reanalysis be carried out in conjunction with the development of the CDAS (Kistler et al., 2001).

The design, development, and implementation of the reanalysis project occurred during 1990-1994. It involved the recovery of land surface, ship, rawinsonde, pibal, aircraft, satellite, and other data, and quality controlling and assimilating these data. Data collection, a major task that was performed mainly at NCAR, required the cooperation of international agencies including the U.K. Meteorological Office, Japanese Meteorological Agency, and the European Centre for Medium-Range Weather Forecasts (ECMWF). The data assimilation system is kept unchanged over the reanalysis period, although it is still affected by changes in the observing systems.

The main outputs from the NCEP/NCAR reanalyses are four-dimensional gridded fields and observations. Gridded output variables are classified into four classes, based on the degree to which they are influenced by the observations or the model. An archive of five decades of observations has been encoded into a common format (denoted Binary Universal Form for the Representation of Meteorological Data [BUFR]), including metadata.

NCEP conducted two major reanalyses, one from 1948 to the present (Kalnay et al., 1996; Kistler et al., 2001) and a second from 1979 to the present (Kanamitsu et al., 2000). The long, consistent datasets from reanalysis have been extremely valuable to an impressive range of scientific studies and applications, including climate monitoring, climate prediction, applied climatology such as prediction and monitoring of climate related health problems, stratospheric transports and chemistry, and boundary conditions for regional models. It is estimated that 5,000-15,000 papers and studies have been carried out just in the last few years using the reanalyses, and their use is growing exponentially (E. Kalnay, personal communication).

Two other major global reanalyses have been undertaken: the ECMWF 15-year reanalysis covering 1979-1993 (Gibson et al., 1997) and the NASA/Data Assimilation Office 17-year reanalysis

data operations staff, data analysts, documentation writers, user services, and archive and distribution team members.

NOAA should also determine the relevance of the Data Quality Act (67 CFR 8452, http://www.noaanews.noaa.gov/stories/iq.htm) to satellite data and resolve issues related to watermark, provenance, reproducibility of data, peer review, integrity of data, and supporting information.

covering 1980-1996 (Schubert et al., 1993). The University of Maryland has performed a preliminary 50-year reanalysis of the oceans.

### Lessons Learned

1. Early and continuous input from major types of users is vital. An advisory panel that guided the project through the first four years (1989-1993) provided stability and advised on such matters as data dissemination (e.g., recommendation on issuing CD-ROMs compatible with multiple platforms including PCs). After 1993 the advisory panel was replaced by a user advisory committee. Annual workshops held since 1991 serve as another conduit for feedback on the use of the data.
2. Funding support from other NOAA offices (the National Weather Service and the Office of Global Programs) and the National Science Foundation was indispensable.
3. Global use of the reanalysis data has been facilitated by free distribution of the data products on CD-ROM and online, and the provision of the data in formats that are commonly used (e.g., netcdf). The CD-ROMs, with subsets of the data that satisfied the needs of a large percentage of users and included user-friendly software, were distributed with 13,000 copies of two major articles in the Bulletin of the American Meteorological Society. The NOAA Climate Diagnostic Center (CDC) has a user-friendly interface for accessing the data, creating subsets, and displaying the data online. The same data are also available directly from NCAR, NCEP, and other sites, such as the International Research Institute for Climate Prediction at the Lamont-Doherty campus in Palisades, New York.
4. The system design must include a capability for reprocessing and error correction; for example, because of changes in the observing systems (mainly satellites), model parameterizations had to be accommodated and errors in the longitude assigned to synthetic observations of sea-level pressure over the Southern Hemisphere had to be corrected. The availability of other analyses for intercomparison has been valuable for assessing the reliability of the NCEP/NCAR reanalyses.

### Data Format

Different categories of users will require different data formats, and these will change over the decades. If the CDRs are available in multiple, flexible, and well-documented formats or in a form that permits the use of alternate formats, NOAA will be able to meet the needs of future users, particularly several decades hence when the data are still valuable. The committee does not believe that this report should recommend which data format to use for CDRs, because technology development will change the available options. It is, however, important that CDRs be made available in interoperable formats with certain common standards, such as being self-describing, and that there be periodic reconsideration of CDR formats as the underlying data technologies change. Many current format tools have been written by data users. This is an inefficient use of research funds, as the same problems must be solved over and over again by individual researchers. It would be better for the data provider, in consultation with data users, to promote standard formats for CDRs or to provide the tools necessary to use the data. In the case of the latter, simplistic visualization is not enough, and extraction and format conversion are essential.

If NOAA is to fulfill its goal of increasing the understanding of climate and climate impacts, integration of satellite data with other types of data, such as in situ, geospatial, and socio-economic data, must be simple and easy to perform. A currently effective means of accomplishing this is through Geographic Information Systems (GIS). Common geospatial standards for data that permit interoperable use of various types of software and hardware are critical. Specifications from the Open GIS Consortium (http://www.opengis.org/) should be used in CDRs to promote interoperability across geospatial data types. While GIS systems might be considered as a means for distribution, calibrated, geolocated, and time-stamped observations likely will remain the primary data source.

### Data Accessibility

The key to data access is the ability to provide data to the scientists and other users that is as practical and cost-effective as possible. With the increase in satellite data resolution, and the corresponding increase in data volume, providing users with no more than what they want has become increasingly important. Two primary ways of reducing the amount of unwanted data delivered to the users are (1) to increase the accuracy of the search and (2) to provide subsetting services. Mechanisms for providing each will help NOAA to ensure that the CDR generation program is successful.

- *Subsetting.* NOAA should ensure rapid access to meet the user subsetting needs. Specifically, they should provide the capability to subset the CDR in multiple formats including row/column bounding box, similar to the

EOSDIS data pool concept (http://nsidc.org/data/data_pool/index.html). In addition to subsetting by time and space, subsetting by parameter (e.g., cloud fraction or channel radiance) should also be available.

- *Temporal Search*. Users should have the capacity to search by day or year, time of orbit, and temporal subsampling (data from every nth day), separately or in combination.

- *Spatial Search*. Users should have the capacity to search by a variety of spatial specifications (e.g., latitude and longitude sectors and spherical polygons).

Given recent and expected future advances in networking, storage capabilities, and technologies for data access (e.g., mobile devices and wireless networks), NOAA should endeavor to make CDRs available online through user-friendly, automated procedures. NOAA should take advantage of developments in other agencies with regard to data distribution initiatives, grid computing, and online collaborative tools (e.g., NSF cyber-infrastructure and middleware initiatives).

Web service infrastructures are rapidly evolving to permit data access through such digital libraries as the Digital Library for Earth Science Education (DLESE) and the National Science Digital Library (NSDL). Such software infrastructure as Unidata's Thematic Real-time Environmental Distributed Data Services (THREDDS) is a tool for accessing archived environmental datasets from distributed server sites. NOAA can benefit from these technologies by designing a system to accommodate data mining and data discovery.

To ensure that the CDRs are used in multiple fields of science, CDR products should be promoted and distributed by multiple channels and mechanisms: electronically through online media, person to person at scientific meetings, and through scientific publications and presentations. To meet the needs of differently equipped users located around the world, data products should be distributed by multiple electronic paths, as well as made available on a variety of media (e.g., CD-ROM, DVD, DLT, flash drives). As data distribution technologies evolve the means of disseminating CDRs must also evolve.

The ability to promote and distribute the data depends greatly on metadata standards. Although data may be held in various formats broad data discovery and access will depend on metadata standards, or metadata that can be automatically mapped to a standard. NOAA should monitor and take an active role in groups working on this problem so that their needs for the CDRs are represented in the standards.

### Data Policy

Usually implemented through the data management process, the CDR data policy will have to build upon longstanding policies and develop some new aspects as the development process develops, probably determining the applicable policy for each CDR (e.g., some CDRs will involve multiagency or international sources [see chapter 5] with different policies applicable). NOAA already has significant experience with these types of policy issues. The primary principle should be open and unrestricted access to all data and in compliance with the recommendations of World Meteorological Organization (WMO) Resolution 40 and all applicable federal regulations. The irreplaceable primary or FCDR data must be preserved in perpetuity, and data policies for superceded TCDR versions should be established so that they may be decommissioned and deleted and the storage resources recovered.

### Data Security

Data management systems must ensure the security of stored data. The primary means of data security currently involves having authenticated system backups. Redundancy is essential, and backup copies must be regularly placed in widely-separated geographical locations.

## DATA STEWARDSHIP AND LONG-TERM ARCHIVE

Various scientific and policy-making groups have reviewed and defined the requirements for essential data systems and services needed to ensure a long-term satellite data record in support of climate research (NRC, 2000b; GCOS, 2003). Recently the Earth Observing System Science Working Group on Data offered recommendations relevant to the Earth Science Data Lifecycle that have been modified appropriately here.

- NOAA, in conjunction with NASA and the DOD, should determine the nature of “stewardship.” How does it work (at the various stages in the life of the CDR)? Who is responsible for it? Who funds it?

- NOAA has initiated a Comprehensive Large Array-data Stewardship System (CLASS), which is an electronic library of environmental satellite data. The web site provides capabilities for finding and obtaining those data. CLASS is an operational component of NOAA's Office of Satellite Data Processing and Distribution (OSDPD) and NOAA's National Climatic Data Center (NCDC). Its success is dependent on provision of adequate resources. It may provide useful lessons and capabilities for CDR data management.

Planning for CDR data management must take place within a framework that considers the full range of data management issues over multiple decades. The data lifecycle approach provides a broad view of data stewardship that represents a fundamentally new concept. The transfer of FCDRs and TCDRs to a long-term archive (LTA) should begin when there is community agreement on the validity of a CDR, although with reprocessing as a hallmark of the TCDRs they may not be suitable for an LTA. The current concept for EOS that the transition occurs following the end of a mission is not valid for the planned multidecadal life span of NOAA's CDR program. The primary distinction between an active archive and an LTA designation is the level of user support provided to a dataset and ease of access (e.g., rapid ftp versus copied from physical media). In particular, as long as reprocessing is likely, data should remain in the active archives. Coordinated schedules and goals should be set up for working with the other agencies to effect initial CDR agreements, planning, and eventual transfer. The proposed advisory council and the science thematic teams should participate in advisory panels and committees within NOAA to specify and administer the CDRs.

Each TCDR team should develop guidelines to manage the data stream throughout the data lifecycle.[1] These guidelines will provide the NPP and NPOESS mission science teams with a roadmap for the orderly transition of the data from production to an active archive and ultimately on to an LTA facility. New operational satellite missions must plan for an orderly process that addresses data archiving, metadata collection, data access, and data delivery as the data progresses through its full lifecycle.

---

[1] Pre-FCDR and FCDR data should be preserved forever. TCDRs will go through numerous improvements and upgrades, so there is a life cycle for specific versions of TCDRs. This is not a small issue for data managers and storage specialists, and it plays heavily into the need for metadata for version and lineage control as well as data policy for discontinuing versions.

The permanent preservation of the CDR must be assured for at least a century; therefore, policies and strategies should be in place to facilitate the long-term viability of the CDRs. This requires methodologies that address

- migration (copying and reprogramming applications to new hardware system);
- encapsulation (an e-document explaining how to recreate software and hardware systems to decode the bits);
- emulation (software running on new platforms that mimic the hardware processing, prior software and applications systems, and virtual computers);
- standards (an ISO standard reference model that provides a conceptual framework and defines a consistent set of standards for all major archive functions and services); and
- peer-to-peer file sharing (distributed, ubiquitous computer servers networked in a dependable infrastructure that can support nomadic data access and retrieval).

## INFRASTRUCTURE

### Institutional Structure

Institutional options for structuring data archive and dissemination functions include both central and distributed archives and can be located either totally within NOAA or completely within a nongovernmental center. There are advantages and disadvantages to each approach and risks to each that must be considered when planning the institutional structure for managing CDRs. There must be carefully defined and documented agreements to ensure continuity of data preservation and provision for transfer of data to other archiving centers if this continuity cannot be assured. Each of the approaches discussed here has different institutional and financial requirements. CDRs need a system that is cost-effective, provides the flexibility required by the disparate CDR user groups, and has the stability required for permanent data services and preservation. There are at least four ways this could be done:

1. A single, archive within a NOAA/NESDIS center would provide NOAA with complete responsibility and control over the data storage system, including maintenance and upgrades. The disadvantages of a single system are the volume of data that would need to be managed and the

potentially serious impact of a single point of failure. Large systems also tend to be more inflexible over time and more difficult to adjust, which also threatens success (Standish Group 1999). The diversity of the CDRs that will be managed in a single archive may affect the quality of the data, and it is possible that, like EOSDIS, a one-stop shop will not satisfy user needs.

2. A second approach, storing data in distributed archives across NESDIS disciplinary centers, namely the National Climatic Data Center (NCDC), National Geophysical Data Center (NGDC) and its linked National Snow and Ice Data Center (NSIDC), and National Ocean Data Center (NODC), also retains responsibility for data management entirely within NOAA. The advantage of this structure is that each topical data center would focus on a more limited number of TCDRs. Another strength of this approach is that the in situ data is usually close at hand, along with experts from the associated fields. Somewhat smaller data storage units may also be more adaptable to technology development. One disadvantage of this scheme is that it requires strong coordination across NOAA data centers and requires additional oversight to ensure that goals are met on schedule. CDRs in several disciplinary areas (biosphere, cryosphere, hydrosphere) are not formally represented in the three NESDIS centers. Because some CDRs may fit equally well in two or more data centers, a Web portal across the data centers would permit users to identify what they want and could supply the software to locate data from the appropriate center.

3. A distributed archive that spanned both NESDIS and external centers, such as NASA DAACs, the Federation of Earth Science Information Partners, International Council of Scientific Unions (ICSU) World Data Centers (WDCS), or university data centers is another option. Such an archive could encompass multiple agencies and government and nongovernmental data centers. Many of the NASA DAACs, for example, not only archive data generated by external groups (including products derived from NOAA polar orbiters), but they also create their own CDRs and have well-organized user communities. NOAA could utilize these functioning structures for CDR creation and data management, and develop specific partnerships, rather than create a new organization. The advantage of this approach is a further reduction in the number of CDRs a location manages, which may result in more informed data stewardship. As with option two, the potential problems incurred are the more complicated organizational structure and the need for a transparent data portal. In addition, continuity over time in terms of both data management and funding for data maintenance may be more difficult to accomplish with this more diverse institutional structure. Effective short-term (possibly for mission duration only) archive centers could be

located at facilities with appropriate experts. The additional condition would be that these centers have a long-term archive agreement to transfer all data, metadata, and associated documentation to a permanent archive center.

4. A fourth option is a central archive that is subcontracted outside the government. The advantage of this scheme is that proposals can be solicited for the project, which may lead to innovative new data archival and dissemination schemes. The winning proposal would also have full stewardship of the system, which could help ensure success. However, NOAA would have less control over the maintenance and upkeep of the system and costs could rise significantly over time. User services might also be a bigger problem with a fully external central archive.

There are compelling reasons to avoid Options 1 and 4. In the case of option 2, not all NOAA centers have experience with satellite data streams. Option 3 has the advantage of entraining a wider range of expertise. Regardless of which of the four options is selected, there will be a need for strong oversight, periodic reconsiderations of scientific advances and user needs, and frequent assessments of the adequacy of data management procedures and responsiveness to technological changes.

### Levels of Service

An important step in data dissemination is the decision about the levels of service for each CDR. These levels should be assigned for different functional activities: ingest, processing, documentation, archiving, access and distribution, search and order, and user services (see Appendix C for more details based on EOSDIS). For data ingest there are two primary alternative modes: operational (time-critical) ingest with immediate verification of data integrity and quality, and routine ingest and verification of data quality and integrity without tight time constraints. Data processing options include such alternatives as operational products generated within two, seven, or thirty days of ingest or availability of required inputs. Since users have markedly different acceptable processing times, NOAA should survey user communities to determine appropriate time delays.

As noted in Chapter 3 and earlier in this chapter, documentation of the CDR generation process (metadata) is critical for future reprocessing efforts and for using the data appropriately. Data and product holdings (including multiple versions of products and corresponding documentation as needed) should be documented to the adopted standard for long-term archiving, including details of processing algorithms and processing history; documentation should be sufficient for current use (e.g., product type descriptions, product instance [e.g., granule] descriptions including version information, FAQs, "readme" directions, Web pages with links to metadata, user guides, and references to journal articles describing the production or use of the data or product).

### The User Community

Society's need for climate data has grown rapidly, along with better computing capabilities in user communities, better access to data, and an increased appreciation for the impact of weather and climate on daily activities (e.g., NRC, 2003a). At NASA DAACs total requests for products increased from under 41,000 in fiscal year 1996 prior to the launch of Terra to over 208,000 user requests in FY 2002 (F. Fetterer, personal communication). The NOAA data centers have witnessed a similar increase in data requests, volume delivered, and products stored. Since 1996, NOAA data centers have received nearly an order-of-magnitude increase in data requests (Figure 4-1), with marked percentage increases in NODC requests, although NCDC continues to receive the most requests. The volume of data delivered to users has increased at an even higher rate (Figure 4-2). The increase in user requests and data volume accompanies an increase in the number of products stored at the NOAA data centers, from roughly 800 in 1998 to nearly 1600 in 2003 (Figure 4-3). Most of the products requested are from the private sector (Figure 4-4).

## CONCLUDING REMARKS

Long-term archives of FCDRs, derived products, and complete documentation must be preserved. This will facilitate reprocessing and user access to create new TCDRs over the entire record, including the archiving of the required ancillary data, instrument, project and dataset documentation, and the science production software.

The institutional structure chosen for data management must meet the criteria set out for CDR generation, archiving, access, and distribution. The overall system design should be flexible and enduring. An archive should be identified for each data product, including both dissemination responsibilities and long-term archiving.

A clear policy is needed from the beginning to ensure continuity in the

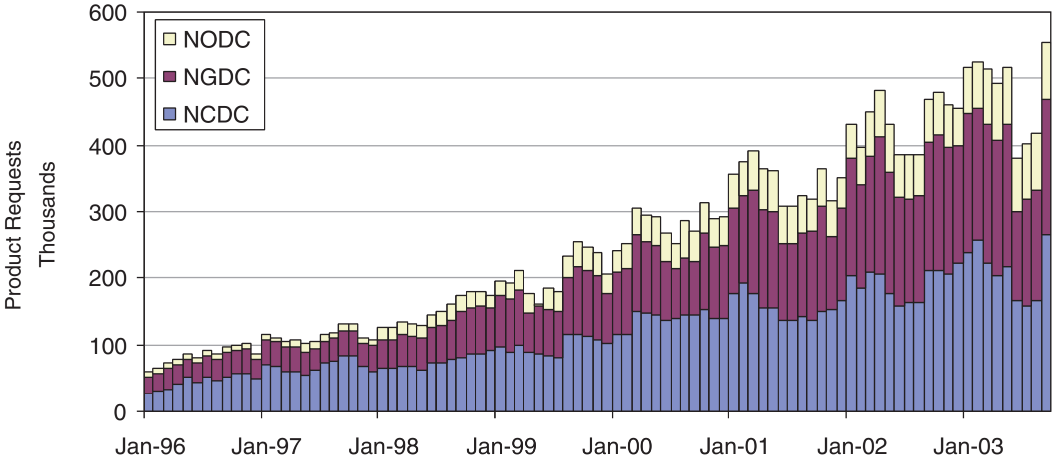


FIGURE 4-1 Product requests at the NOAA data centers. This figure and others below represent the true user community. Data-sharing practices afforded by unrestricted data, especially in the academic and government research communities, ensures there are more users than documented in these figures. This is a hidden benefit to the community, but makes the metrics a little more uncertain. SOURCE: Compiled from NOAA data by S. Drobot, National Research Council.

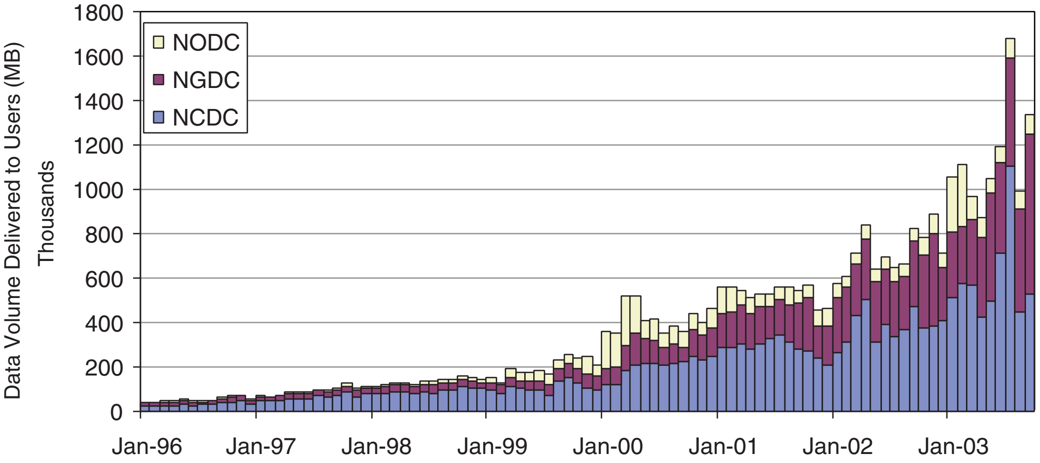


FIGURE 4-2 Data volume delivered to users. SOURCE: Compiled from NOAA data by S. Drobot, National Research Council.

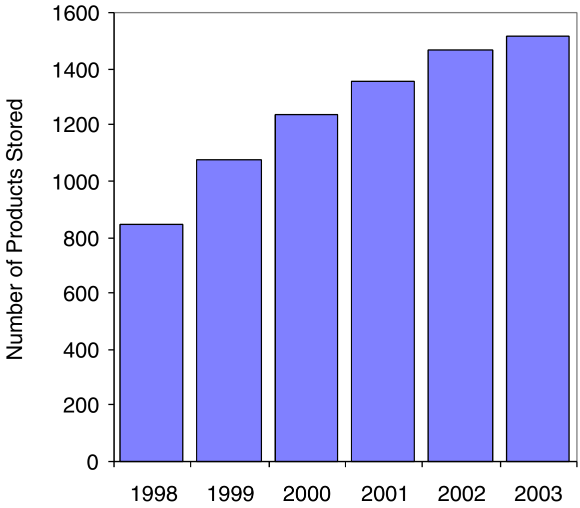


FIGURE 4-3 Products stored at the NOAA data centers. SOURCE: Compiled from NOAA data by S. Drobot, National Research Council.

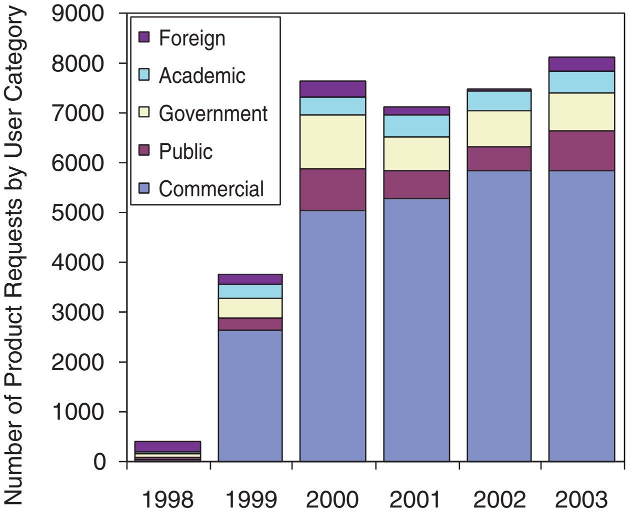


FIGURE 4-4 Product requests by user category. SOURCE: Compiled from NOAA data by S. Drobot, National Research Council.

data record as well as full and open data exchange and access. Distribution must encompass multiple electronic paths and a variety of media. Data must be available in formats appropriate for a variety of uses, including geospatial and socio-economic applications.

Life cycle data management from initial planning, through development and implementation is needed. This must involve cooperation among researchers, data and archive managers, data collectors, and primary users. To assist in making decisions on data stewardship in a resource constrained environment, a process should be established for the science assessment of the long-term potential of data and data products.

Given the large satellite data volumes, it is critical that the NOAA infrastructure provide tools to enable the user to do spatial and temporal searches and arbitrary subsetting. Levels of service must be determined and implemented in the design of the system infrastructure. Preserving complete documentation along with the data is of absolute importance for successful reprocessing of archived data to produce improved or new geophysical products. The use of CDRs by policy makers, resource managers, educators, and planners will require the NOAA CDR system to provide them with the capability for deriving high-level information products from the CDRs.

# 5

# Partnerships Essential for Implementation

The new emphasis on climate within the mission of NOAA and the accompanying responsibility for stewardship of climate data will require an increased focus on partnerships. Because Climate Data Records (CDRs) require consistency and continuity to provide insights into climate variability and change, they require a much broader input from the research community than has been necessary for the operational data required to support weather forecasting. As NOAA looks to the future the recognition of expanding capabilities in handling and processing high-volume data rates, necessary for addressing satellite and Numerical Weather Prediction (NWP) assimilation processing (such as the expanding use of data grids, shared processing, virtual laboratories, high-bandwidth data transmission), combined with the capabilities for distributed support from user communities can increase the ability for partnerships to more directly address the development, analysis, reanalysis, and research of CDRs.

The nature of climate and the family of data records necessary to describe and potentially predict its variability and change requires a global view, which places a focus on CDRs from satellites. As noted in previous chapters, creating a program to develop, produce, archive, and disseminate CDRs will involve a large investment in monetary and human resources. NOAA alone cannot create high-quality CDRs that satisfy the broad user communities of today and provide climate data stewardship for future generations. Fortunately NOAA's plan to create CDRs is of interest to a variety of national and international programs that share similar goals. To maximize the effect of NOAA's finite resources it should develop partnerships with other groups whose goals relate, at least in part, to those of NOAA. In developing a CDR plan and in taking on its stewardship role it is crucial for NOAA to take a proactive leadership role in international and interagency partnerships and leverage the limited funding available to support this type of effort. This will require an open and collaborative environment with full participation from the national interagency and international climate science community.

This chapter addresses the role of interagency and academic teams and partnerships, the role of international partnerships and programs, and the potential need for a change in the present NOAA structure to engage the broader research community and increase the extramural research necessary to achieve success in the long-term CDR process and acceptance of CDRs by the community.

## NATIONAL PARTNERSHIPS

### Interagency Partnerships

Because a number of government agencies share climate-related missions, the CDR process requires strong interagency partnerships. CDRs involving multi-agency participation are strengthened by a diverse funding basis and oversight, and leveraged by human resources provided through those partnerships. Several existing mechanisms can be used to strengthen the multi-agency interactions required in the development of the CDR process; organizations such as the following have some existing leadership roles and responsibilities that should prove useful: U.S. Climate Change Science Program (CCSP[1] ), Office of the Federal Coordinator for Meteorological Services and Supporting Research (OFCM), and the National Polar-orbiting Operational Satellite System (NPOESS) Integrated Program Office (IPO).

For the United States, the CCSP has a mission with goals, objectives, and management infrastructure that addresses CDRs with involvement of both Climate Data Science Teams (CDSTs) and Climate Data Science Councils (CDSCs). The CDST teams "are composed of a group of scientists and engineers whose purpose is to convert raw instrument data into CDRs, including calibration, algorithm development, validation, error analysis, quality control, and data product design" (CCSP, 2003), which corresponds closely to the role and responsibilities of the Fundamental Climate Data Record (FCDR) teams. The CDSCs are responsible for climate observations in support of CCSP research themes, similar in scope and responsibility to the Thematic Climate Data Record (TCDR) teams.

---

[1] The Climate Change Research Initiative (CCRI) and the U.S. Global Change Research Program (USGCRP) were combined into the Climate Change Science Program. The USGCRP supports the long-term objective to build a climate-observing system.

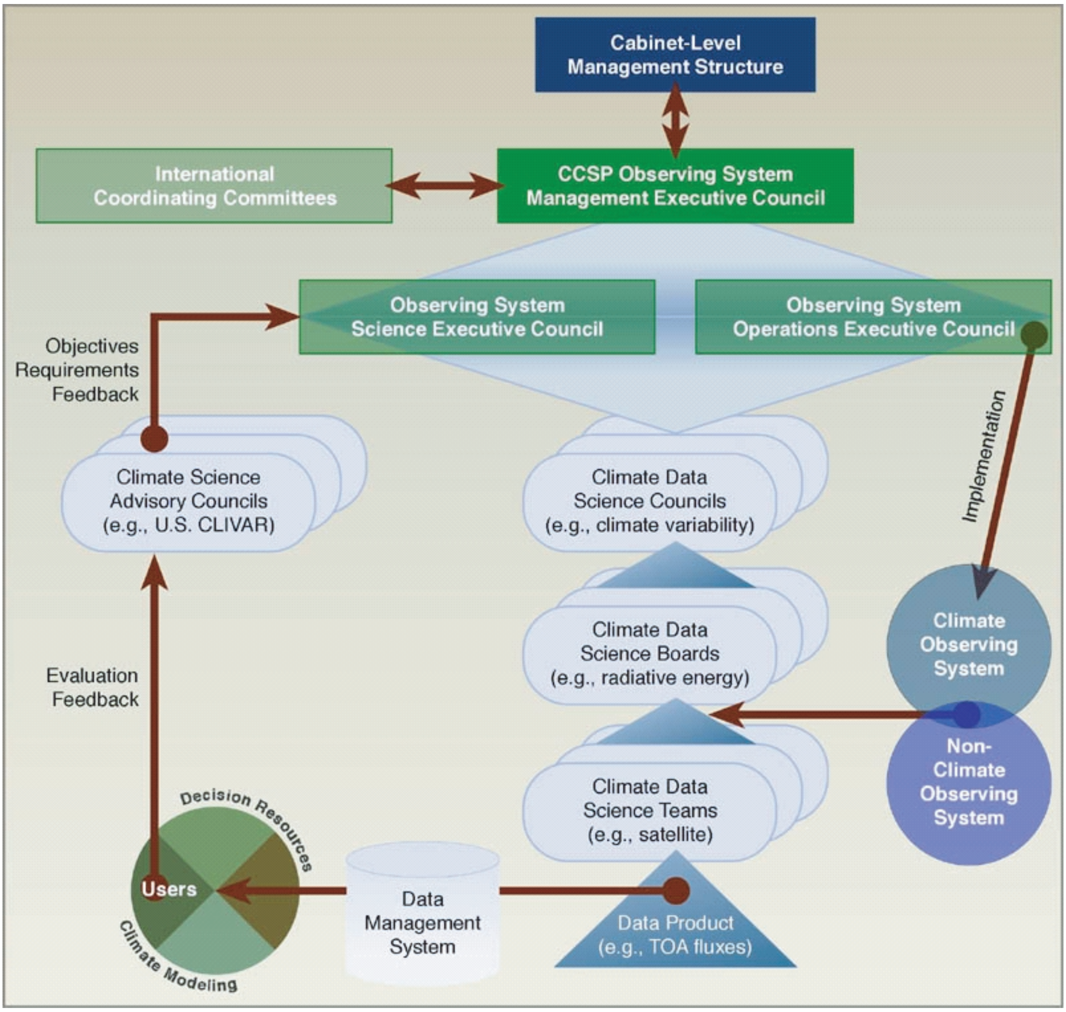


FIGURE 5-1 The organizational structure of the Climate Change Science Program is designed to facilitate wide agency participation and to make research directly useful to decision makers.

The organizational structure of CCSP (Figure 5-1) involves wide agency participation, and its access to the highest levels of government provides a framework that addresses most of the concerns (lessons learned) and issues highlighted in previous chapters for NOAA. As a key partner and participant in the CCSP process already, NOAA has both the capability and responsibility to seriously consider approaching the CDR process within this CCSP framework.

The Office of the Federal Coordinator for Meteorological Services and Supporting Research (OFCM) has the mission to coordinate meteorological activities for 15 federal agencies, and it could be considered for supporting or assisting in managing the CDR process. OFCM has a stronger tie to the operational services than CCSP, and it is increasing its climate services. It has a strong success record for coordinating past multi-agency activities (e.g., Next Generation Radar [NEXRAD] and Automated Surface Observing System [ASOS] procurement and deployment) and there are some advantages to developing the CDR process through the OFCM. OFCM has broad operational weather responsibilities, however, and shifting focus to climate services and CDRs may not be as easily achieved through OFCM as through CCSP.

### Bilateral and Multilateral National Partnerships Involving NOAA, NASA, and DOD

NOAA and NASA have a history of cooperative activities related to CDR generation, including the NOAA/NASA Pathfinder program, the joint NOAA and NASA support of development of datasets for NOAA's Climate Change Data and Detection project, NOAA scientist participation on NASA science teams that produce CDRs, and NOAA/NASA cooperation on the generation of long-term ozone data. Unfortunately the agencies lack a formal, systematic procedure for ongoing cooperation.

Previous NRC reports have outlined several key guidelines for partnerships and programs related to climate data (NRC, 2000a,c; 2003a). Given that the operational meteorological community now formally recognizes climate as a mission, along with the related aspects required to develop climate data records, several of the previous recommendations can be restated from the CDR perspective. NOAA should approach NASA to improve and formalize the process of developing and communicating CDR requirements and priorities (see Appendix E for a listing of previous NRC recommendations). The research and operational communities also should be more alert for new and unexpected applications of NASA's exploratory research and establish a process of assistance for discovering these applications. With the proposed advisory council, science theme teams, and open science workshops, NOAA will have regular contact with user communities, and they should pass the user recommendations on to NASA so that NASA is aware of user concerns as well. This creates a more formal NOAA process to identify user requirements for NASA research and it would benefit both agencies.

A formal process for evaluating all NASA Earth science missions for potential climate applications would provide a solid foundation for developing effective plans for transitioning activities. The advisory council envisioned here could interact with NASA to gather information and advise NOAA on upcoming NASA missions. Regardless of how NOAA forms a plan for interagency communication, the committee stresses that this is a key step in creating a successful long-term CDR program.

The NPOESS IPO[2] (NOAA, DOD, and NASA) is a good first step in formalizing some aspects of the cooperative process. Since the polar-orbiting satellites provide the largest portion of the U.S. data for CDRs, the existing NOAA partnerships with NASA and the DOD (including the IPO) should be strengthened and expanded to ensure the specific aspects of CDRs are systematically addressed.

To effect a smooth transition between the research-oriented Earth Observing System (EOS) program and its continuation as an operational program, NASA has successfully initiated a bridging mission called the NPOESS Preparatory Program (NPP). The primary objectives are to provide NASA with continuation of global change observations after EOS Terra and Aqua and to provide NPOESS with risk reduction demonstration and validation of critical NPOESS sensors, their algorithms, and their processing strategies. NASA's experience with climate-quality observations is that detailed characterization of satellite sensors must be made during development and testing, and frequent calibration during each mission is required to match a satellite's observations to a preceding satellite to avoid measurement degradations that are indistinguishable from climatic trends. This mission will accordingly address a limited set of FCDRs that would carryover into the NPOESS program.

Because of the focused Polar Operational Environmental Satellite (POES) acquisition and operational mission responsibilities, the breadth of the IPO mission would have to be expanded to appropriately address the FCDR and TCDR generation from in situ, nonpolar, and non-US satellites. This may not be sufficient to achieve the needed focus on the necessary CDR development process.

---

[2] The NPOESS IPO was established on October 3, 1994. The IPO organizationally resides within the Department of Commerce, NOAA National Environmental Satellite, Data and Information Service (NESDIS), and is staffed with personnel from DOD, Department of Commerce, and NASA.

### Other National Agencies

NOAA should expand communications with other agencies whose responsibilities include sustained climate observations or climate impacts, namely, the U.S. Geological Survey, the U.S. Department of Energy, and the U.S. Department of Agriculture. Although these agencies are not considered major contributors to satellite climate data record efforts, they do represent large and important user communities of CDRs, and they could provide useful insight needed for creating CDRs. For some CDR needs they may also be willing to share costs for the generation of the CDR, and they may be able to provide in situ data that is essential for verifying and improving CDRs.

The National Ocean Partnership Program (NOPP) is an interagency program worth examining as a potential model for how federal agencies might organize themselves to create CDRs. NOPP was formed in 1997 as an organization of federal agencies that fund oceanographic research and operations. The purpose of NOPP is to coordinate this funding by establishing priorities for research initiatives, avoiding duplication, and leveraging resources from the various agencies to address priorities. NOPP has a small administrative staff, and its major functions are carried out by three groups: (1) the National Ocean Research Leadership Council (NORLC) made up of high-level representatives from the federal agencies. There are now 16 agencies represented on the NORLC, but the chairmanship rotates among the four major funding agencies (NSF, NOAA, Navy, and NASA); (2) an Interagency Working Group that meets once a month and is made up of working-level program managers from each agency; and (3) an Ocean Research Advisory Panel (ORAP) consisting of recognized experts from outside the government who advise NOPP concerning science priorities. At least once a year NOPP issues a Broad Agency Announcement (BAA) soliciting proposals in a particular research area. The funding for each BAA usually is drawn from several agencies. NOPP staff helps administer the proposal review process; awards for projects are then made by the agencies providing the funds.

Considering NOPP as a model for how the U.S. government agencies might organize themselves for creating CDRs, the leadership panel mentioned in earlier chapters might actually be subsumed by such an organization that is much larger than NOAA. NOAA is leading the effort for the NORLC to develop the Integrated Ocean Observing System (IOOS), which is a coordinated national and international network of observations, data management, and analyses that systematically acquires and disseminates data and information on past, present, and future states of the oceans and the nation's Exclusive Economic Zone. Under this NOPP model NOAA may be able to utilize some IOOS assets and management structure for the CDR program.

As an internal NOAA activity, the management of the CDR process implementation could be addressed within the existing NOAA management structure; as an interagency or international cooperative process, the leadership and advisory councils would necessarily have to reflect the participation and cooperation of the appropriate interagency and international partners. This would probably require an interagency/international working group similar to that mentioned here within the NOPP. There would then be three groups analogous to those forming the NOPP. A leadership council would involve high-level representatives from the federal agencies who have the authority to make commitments for their agencies. At this level the agencies would decide on the role each is willing to take and the resources that are brought to the table for a collective effort. An interagency working group would oversee the funding of CDR generation, and an external scientific advisory panel would advise as to the selection of datasets that meet the criteria for becoming CDRs. Using an organizational model such as NOPP would establish the CDR development process as a new, independent cooperative structure among the interested and contributing agencies.

## NOAA and Academic Partners

Although NOAA's relationships with academic partners have been primarily focused on applied research for operations, academic partners have conducted climate research under NOAA funding through several key Cooperative Institutes and grants funded by such NOAA programs as the Office of Global Programs, Office of Research and Technology Applications, National Marine Fisheries Service, Coastal Ocean Program, National Undersea Research Program, and the Saltonstall-Kennedy Grant Program.

The development of satellite CDRs involves a significant and sophisticated understanding of the end-to-end CDR process: instrument capabilities, space platform characteristics, retrieval methods, calibration and validation issues, and processing methods; therefore, a number of different skills and in-depth knowledge of each are required. NOAA's Cooperative Institutes collectively have the breadth of special expertise required and can support CDR development by contributing (1) scientific expertise to stewardship teams; (2) in situ data useful for CDR development and verification; (3) their data processing and computing infrastructure as needed (e.g., 20+ years of International Satellite Cloud Climatology Project (ISCCP) processing and 15 years of Scanning Multichannel Microwave Radiometer and Special Sensor Micrwave/Imager (SMMR/SSM/I) processing at the NSIDC); and (4) graduate students and postdocs who have gained CDR experience and have built NOAA's capacity in this area.

These Cooperative Institutes have been successful in advancing research, operations, sustained observations, applied climatology, and in training students, postdocs, and junior scientists for NOAA missions. The existing institutes show how strong and continuing connections with the university community (not just through the institutes) will be necessary and beneficial to the entire CDR process. Two other examples of NOAA's academic related programs are The National Weather Service (NWS) Collaborative Science, Technology, and Applied Research (CSTAR) Program, which focuses on collaborative university and NWS research applied to forecast operations, and the Cooperative Program for Operational Meteorology, Education and Training (COMET) Outreach Program, which funds research applied to forecast operations.

Participation of academic partners in instrument teams, science teams, user groups, and advisory panels and committees is also essential for ensuring the success of CDRs. Feedback on the utility of CDRs in climate applications from the academic and private sector communities is critical to their success. A related and critical aspect of CDR development is data management; as an example of the previous recognition of this importance, NOAA/NESDIS since 1976 has been providing support for the data management in the NSIDC/WDC for Glaciology in Boulder, Colorado.

### Other Partnerships

With a wide variety of proven and potential societal benefits related to applications based on CDRs, there are also private sector interests in CDR applications (Figure 4-4), and relationships with these interests should be developed and maintained. There is a long history of private sector environmental data relationships with NOAA and NASA that may be considered for use with CDRs.

Among the NOAA programs through which private sector partners can be funded include the National Sea Grant College Program (a partnership and bridge between government, academia, industry, scientists, and private citizens) and the Saltonstall-Kennedy Grant Program (financial assistance for research and development projects to strengthen and develop the U.S. fishing industry). The NWS had private sector partners for delivery of NEXRAD radar data, a program that ended in January 2001. More recently the NWS embarked on another restructuring of NEXRAD data distribution, using the Unidata Local Data Manager. Private sector representatives have been involved formally in the strategic planning and have provided recommendations for data types and distribution mechanisms.

With regard to other types of climate data, NOAA has private citizen partners in the NOAA Cooperative Observers Network (http://www.nws.noaa.gov/om/coop/). This network produces observations that can assist in the ground validation of satellite CDRs. NOAA is encouraged to continue its modernization (as in the case of the Climate Reference Network) and maintenance of this service through training, improved instrumentation (adding such nontraditional but critical sensors as soil moisture sensors), and expanding the spatial coverage of the network in order to provide the best in situ data needed to answer critical climate questions.

In the area of regional climate NOAA is working with the Western Governors Association to plan a drought monitoring network. The association is sponsoring a drought bill to Congress that calls for establishment of a national drought council as well as the National Integrated Drought Information System (NIDIS), a nationwide drought monitoring network (a partnership of federal and state agencies and external partners) to measure parameters from sky to soil (soil moisture at several depths). The drought monitoring includes the satellite vegetation index and surface wetness measurements.

The developing NOAA Climate Transition Program (NCTP) also aims to expand regional climate services by developing information tools for decision makers and providing education and outreach capacity for new products.

The NASA Sea-viewing Wide Field-of-View Sensor (SeaWFS) ocean color project also has a private sector partnership. According to the SeaWFS Project Web site (http://seawifs.gsfc.nasa.gov/SEAWFS.html):

> The SeaWiFS mission has been made affordable and timely because of its unique private vendor, data purchase structure. As part of the contract between NASA and Orbital Sciences Corporation [OSC], NASA retains all rights to data for research purposes, and ORBIMAGE retains all rights for commercial and operational purposes. There has been an embargo period of 2 weeks from collection for general distribution of data to research users to protect ORBIMAGE's commercial interest. Three exceptions to the 2-week embargo are a) field experiments requiring data for ship positioning, b) operational demonstrations to prove feasibility and usefulness, and c) assessment of calibration/validation and instrument performance by NASA. Access to the NASA data archive has been permitted for research purposes by authorized users only. After five years, the data may be used without restriction.

There are also significant potential private partner connections related to CDRs in the energy, insurance, agriculture, financial (e.g., weather derivatives market) and private weather service industries. Involvement of the American Meteorological Society's Private Sector Board and other private sector organizations (e.g., Commercial Weather Services Association) should be encouraged to ensure participation this stakeholder community. For most satellite CDRs private sector partners would be valuable participants in advisory committees and user groups, and as collaborators with academic partners.

## INTERNATIONAL PARTNERSHIPS

Climate has been a global concern since the first World Climate Conference in 1979 and following the organization of the World Climate Research Programme (WCRP) jointly by the World Meteorological Organization (WMO), International Oceanographic Commission (IOC), and International Council for Scientific Unions (ICSU). The organization of meteorological observing networks and coordination and sharing of satellite observations internationally is an indispensable component of global climate research.

The Committee on Earth Observation Satellites (CEOS), currently chaired by the NOAA/NESDIS director, Greg Withee, and the International Global Observing Strategy partnership (IGOS) form the basis for strong and continuing international cooperation on the acquisition and development of CDRs from the international satellite systems. As a result of the recent Earth Observation Summit, an ad hoc Group on Earth Observations (GEO) was established to prepare a 10-year implementation plan for a coordinated, comprehensive, and sustained Earth observation system (or systems). This provides additional focus and support for further development of the CDR process internationally. It is important for NOAA/NESDIS to take advantage of these broad, high-level relationships, since CDR development may require investments by meteorology and space agencies from other countries to ensure reliable and consistent CDR global datasets.

For satellite data there are also international bilateral partnerships between NOAA/NESDIS and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). NASA also has these type partnerships with the European Space Agency (ESA), Radarsat International, and the Japan Aerospace Exploration Agency (JAXA), formerly the National Space Development Agency of Japan, all of which can facilitate development of many of the FCDRs required for supporting the family of CDRs. In addition, the National Climatic Data Center (NCDC) is the World Data Center for

Meteorology. However, the development of the thematic CDRs that focus on the climate parameters blend FCDRs from a variety of international sources and data types (satellite and in situ) and, therefore, require a significant involvement in the international projects currently developing these types of global climate products. NOAA should proactively focus on broad data sharing and exchanges, as necessary, with international partners. International data access will allow the greatest climate science knowledge advancement.

The WCRP projects Cryosphere and Climate (CliC), Climate Variability and Predictability (CLIVAR), Stratospheric Processes and Their Role in Climate (SPARC), and especially Global Energy and Water Cycle Experiment (GEWEX) are integrating FCDRs from countries and organizations worldwide, leveraging the funds required and demonstrating the added value of international collaboration in developing climate products. Several of the WCRP projects have well-established and published procedures supported by the international community as well as data management working groups coordinating their production. Several have data commitments made under the WCRP international agreement process. Examples include ISCCP (clouds), Global Precipitation Climatology Project (GPCP) (precipitation), Global Water Vapor Project (GVaP), Global Aerosol Climatology Project (GACP), International Satellite Land-Surface Climatology Project (ISLSCP) all under GEWEX, as well as the Arctic Precipitation Data Archive at the GPCC, and the International Arctic Buoy Program under the Arctic Climate System (ACSYS, now CliC). NOAA currently has both a data supply, participatory involvement, and data archival role in the ISCCP and GPCP projects, but not the broad proactive, leadership role across the climate community that is required to take on the stewardship necessary for achieving global acceptance of the family of CDRs.

A strong and sustaining mechanism is required within NOAA to take advantage of international activities in satellite programs and disciplinary climate system programs and projects. While NOAA has well-established operational links with other national space agencies and climate services, it is less well established in leadership roles for development of the broad range of global CDR projects of the WCRP.

## CONCLUDING THOUGHTS

Effective partnerships will be essential in the CDR process. The broad need and uses of CDRs throughout the interdisciplinary, interagency, and international community require the involvement of these groups in the development of CDRs through both science advisory teams and in management, coordination, and funding. The difficulty in recommending an interagency or international organizational structure for the CDR process is that the greater the participation and involvement of organizations outside of NOAA, the greater the influence in decisions and outcome that come from beyond NOAA. While we believe this broader involvement is certainly necessary, this could threaten the long-term stewardship and leadership role we also believe NOAA must play to ensure a successful CDR process. NOAA must take ownership of the overall process to be a true long-term steward; however, it clearly cannot do this alone. We have previously recommended the basic organizational structure of a leadership council, supported by an advisory council and expert and science teams for the FCDRs and TCDRs. While this basic functional organization is necessary, broadening into the multi-agency and international arenas will require an additional interagency and international working group for cross-organization implementation management. We also believe that building on existing organizations is preferred to establishing new and independent organizational structures, and have discussed such existing organizations as CCSP, OFCM, IPO, IGOS, WCRP, and NOPP. While none of these organizations were designed to serve just NOAA, all were created and serve a function based on multiple agency or international involvement and funding support. The key element is for NOAA to retain leadership, stewardship or, in essence, act as the executive agent for the climate community and request support from one of these existing organizations to accept the role of implementing organization for the CDR process on a full participatory basis. If the implementing organization were to fail to function properly or were to go out of existence, NOAA would retain the responsibility for maintaining the CDR process within its own or another organizational structure, while containing the basic recommended organization elements.

While existing international organizations could be used, we believe the existing U.S. multi-agency organizations should be considered first, rather than a new or probably more complicated international structure. Keeping in mind that as the current CEOS and IGOS partner activities develop and themes become closer to implementation, NOAA must become a part of this process to ensure appropriate international understanding, cooperation, and support for the CDR process. At this time the most appropriate structure to initiate the CDR process under NOAA stewardship appears to be the current CCSP structure with its CDSTs and CDSCs. It may be that NOAA should take on the observations and data management portion of the CCSP (at a minimum for CDRs that are primarily satellite-based) as the lead or executive agent for the implementation of these parts of the

CCSP. Successful development, maintenance, research support, and long-term commitment to CDRs will require strong, sustained funding support for this process and associated agreement from all CCSP agencies. Even with interagency and international support, this is a new commitment to support the broad climate science community, and will need greater funding than NOAA has previously committed.

# 6

# Conclusions and Recommendations

NOAA's new vision is to "move into the 21st century scientifically and operationally, in the same interrelated manner as the environment that we observe and forecast, while recognizing the link between our global economy and our planet's environment." This vision includes a new mandate to understand climate variability and change to enhance society's ability to plan and respond. Given the inherent complexity of climate, reliable and stable long-term observations are needed to describe and potentially predict climate. This naturally encompasses a global view and highlights the importance of using satellite data. However, great care must be taken to ensure that the climate data records (CDRs) based on satellite observations have the necessary reliability and consistency to distinguish between artificial changes related to the observing system and real changes in climate. Developing a successful satellite CDR generation program poses many challenges owing to the varied uses of climate data, the complexities of data generation and storage, and the difficulties in sustaining the program indefinitely. This chapter highlights the key findings of the committee and addresses the important recommendations that will help to ensure that NOAA designs a plan to guide satellite CDR generation from existing and new satellites for understanding, monitoring, and predicting climate variations and changes. We present one overarching recommendation and six supporting recommendations.

## OVERARCHING RECOMMENDATION

**NOAA should embrace its new mandate to understand climate variability and change by asserting national leadership for satellite-based climate data record generation, applying new approaches to generate and manage satellite climate data records, developing new community**

**relationships, and ensuring long-term consistency and continuity for a satellite climate data record generation program.**

NOAA already is well established as a national leader in weather services, and NOAA also provides leadership for weather-based satellite data. NOAA's climate mandate is a new function, and NOAA will need to embrace and be proactive in providing leadership for climate data in order to fulfill its mandate. A successful CDR generation program requires a long-term commitment and efforts above and beyond NOAA's traditional role in weather forecasting. The task and the structures being proposed for NOAA in this report are considerably more complex, costly, and demanding than those currently in place. Unless there is the highest level of commitment within the agency to institute and fund these changes, there is considerable doubt in the science community that the CDR agenda as described in the report will succeed.

The committee's review of some previous efforts reveals a number of key lessons learned relating to the involvement of user communities in all program aspects and adhering to several guidelines for creating, storing, and reprocessing fundamental climate data records (FCDRs) and thematic climate data records (TCDRs).[1] Particular care is needed to store all data with thorough metadata and in easily accessible formats. NOAA should not feel obligated to be solely responsible for generating all the nation's CDRs; many other agencies and communities have similar interests and expertise, and by enhancing and expanding community involvement in the program NOAA can help to ensure community acceptance and creation of the best possible CDRs.

## APPLYING NEW APPROACHES TO GENERATE AND MANAGE SATELLITE CDRS

**Supporting Recommendation 1: NOAA should utilize an organizational structure where a high-level leadership council within NOAA receives advice from an advisory council that provides input to the process on behalf of the climate research community and other stakeholders. The advisory council should be supported by instrument and science teams responsible for overseeing the generation of climate data records.**

---

[1] See Figure 1-2 for the distinction between FCDRs and TCDRs.

NOAA would benefit greatly from utilizing an organizational framework responsive to advice and feedback from user communities, and where there are mechanisms for redirecting the program based on advice and feedback. In particular, NOAA can help to ensure success if it involves scientists with a vested interest in CDRs, finds committed people to generate the CDRs, develops technical and science support for broad involvement, and creates teams that are reviewed and renewed regularly.

*An advisory council should establish criteria for selecting climate variables to become satellite-derived TCDRs and recommend which variables should be developed into TCDRs based on proposals from thematic science teams.* Since NOAA cannot create all possible TCDRs, mechanisms must be in place to select an appropriate number to generate. Based on input from user communities, an advisory council of internationally recognized experts can recommend to NOAA which TCDRs should be created and subsequently whether these TCDRs are accepted and utilized by the community.

*The generation of FCDRs should be carried out by a team of engineers and scientists, with representatives from the thematic science teams to ensure feedback from the generation of TCDRs.* The ultimate legacy of the CDR program is the data passed on to the next generation. To ensure that the FCDRs are generated with the highest possible accuracy and stability, the FCDR teams should monitor satellite characteristics and they should document their work extensively so that future generations can easily assess and understand what they have done.

*TCDR Science Teams formed within broad interdisciplinary areas should prescribe algorithms for TCDR development and oversee TCDR generation.* Most users will utilize TCDRs, not FCDRs, and the success of NOAA's program is dependent on creating reliable and stable TCDRs. The TCDR teams should be led by recognized scientists who are actively engaged in using the data. These teams should include research scientists funded by or employed by NOAA and scientists from other agencies, academia, and private industry who use the data, and they should be competitively selected, with limited (but renewable) terms.

**Supporting Recommendation 2: NOAA should base its satellite-based climate data record generation program on lessons learned from previous attempts, which point out several unique characteristics of satellite climate data records, including the need for continuing calibra-**

**tion, validation, and algorithm refinements, all leading to periodic reprocessing and reanalysis to improve error quantification and reduce uncertainties.**

*NOAA should make radiance calibration, calibration monitoring, and satellite-to-satellite cross-calibration a part of the operational satellite system*. Changes in satellite characteristics, such as orbital drift and sensor degradation, compromise the ability to create high-quality, consistent CDRs over time. Procedures must be in place to monitor the observing system for irregularities that could corrupt the long-term value of the FCDRs. A suitable period of overlap between new and old satellite systems is also vital to determine inter-satellite biases and maintain the consistency of time-series observations. Since most of NOAA's operational satellites were created as weather rather than climate platforms, this is notably relevant for NOAA to address.

*An ongoing program of validation should be carried out to determine the uncertainty associated with TCDRs*. The process of validating a TCDR derived from satellite measurements is not simply a one-time activity carried out in a limited number of locations. It is the process of establishing rigorously derived uncertainties for the TCDR using independent correlative measurements conducted throughout the time period of record and over global scales, which in turn determines whether a true climate trend can be detected.

*NOAA should establish a two-track CDR generation program, including an upgradeable baseline CDR track and a second (mostly extramural) funded research program to validate, analyze, assess, and reduce uncertainties in future base versions*. The two-track approach helps foster a culture where scientists and users know that future improvements will be available. The FCDRs will need to be reprocessed as new information is acquired or better calibrations are made, but these records will eventually become stable. The TCDRs will continue to change indefinitely as new or improved algorithms are developed and improved applications are made of the FCDRs.

**Supporting Recommendation 3: NOAA should define satellite climate data record stewardship policies and procedures to ensure that data records and documentation are inexpensive and easily accessible for the current generation and permanently preserved for future generations.**

*The data management system should provide free and open access to data, facilitate the reprocessing of CDRs, and allow for new satellite CDRs to be created.* There is a need for a clear policy from the beginning to ensure continuity in the data record as well as full and open access to data and metadata, including the ancillary data required to reprocess the CDRs, project and dataset documentation, the science production software, and easy error-reporting procedures. Preserving the documentation with the data is important for future reprocessing of archived data. A variety of users will access data, and NOAA can ensure a more robust program if the data are available in formats appropriate for a variety of uses, including geospatial and socio-economic applications.

*NOAA should ensure a data management infrastructure that can accommodate specific user requests.* In view of the large satellite data volumes that a CDR program will create, the NOAA infrastructure needs to provide tools enabling the user to do spatial and temporal searches and arbitrary subsetting. These levels of service should be determined and implemented in the design of the system infrastructure.

*NOAA should establish a process for scientifically assessing the long-term potential of data and data products.* Lifecycle data management from initial planning through development and implementation is needed for a successful program, and this should involve cooperation among researchers, data and archive managers, data collectors, and primary users. Given the limited resources that programs face, scientific assessments of the data can help NOAA to organize its archive so that data dissemination is efficient in terms of personnel and financial resources.

## DEVELOPING NEW COMMUNITY RELATIONSHIPS

**Supporting Recommendation 4: NOAA should develop new community relationships by engaging a broader academic community, other government agencies, and the private sector in the development and continuing stewardship of satellite climate data records.**

*NOAA should annually convene an "open science meeting" where users share their findings (i.e., give science talks) and discuss limitations and recommendations for improving the TCDRs in a particular theme area.* Regular opportunities for open dialog between those creating CDRs and those using them will improve the quality of the data and foster support for continuing the CDR generation program. These meetings could be held in conjunction with conferences held by other organizations, such as the American Meteorological Society or the American Geophysical Union, with benefits being cost savings and broader attendance.

*NOAA should include other agencies and user communities in development, analysis, and reprocessing of CDRs.* A high level of commitment within NOAA and a number of changes at multiple levels within the agency will be needed to institute and fund the various components of CDR stewardship, but it will still be essential to aggressively seek out partnerships. The expertise for creating satellite CDRs lies within the broad academic, government, and private sectors, and through partnerships with these entities NOAA can ensure a more successful CDR generation program. By including the other sectors in the CDR generation, analysis, and reprocessing program, NOAA can also engender community acceptance of CDRs. The committee notes that NOAA may also need to develop new ways of working with partner organizations. Well-defined procedures for interagency communication and collaboration are essential for long-term stewardship and the success of the CDR program.

*NOAA should solicit a commitment from other agencies (e.g., NASA, NSF) and utilize many line offices to support research that utilizes the TCDRs.* A commitment to support research that utilizes the TCDRs will help to ensure the program's success, but this need not be the sole responsibility of NOAA.

> **Supporting Recommendation 5: NOAA should consider existing U.S. multi-agency organizations for implementation of the climate data record program, rather than devising a new structure. The most appropriate organization is the Climate Change Science Program (CCSP).**

NOAA need not implement an entirely new management structure for generating CDRs. The goals and management structure of the developing Climate Change Science Program are similar to NOAA's climate goals, and NOAA could assert leadership by volunteering to be the lead or executive agent for the observations and data management portion pertaining to satellite CDRs. The CCSP structure already has built-in interagency interactions that NOAA could leverage, and by taking the lead for satellite CDRs, NOAA could advance its new climate mandate.

## ENSURING LONG-TERM STABILITY FOR A SATELLITE CDR GENERATION PROGRAM

**Supporting Recommendation 6: NOAA should pursue appropriate financial and human resources to sustain a multidecadal program focused on satellite climate data records.**

With a coordinated CDR effort under the CCSP, NOAA could provide the nation with the needed leadership to develop CDRs. Even if NOAA leverages funds and personnel from other agencies, academia, and private industry, the committee believes that NOAA will have to be aggressive in developing avenues for additional funds to provide the needed capital to successfully generate, analyze, reprocess, store, and disseminate CDRs for decades, taking inflationary increases into account. Developing a satellite CDR program is fundamentally important to the nation, and it is imperative that the effort not be inhibited by a lack of human or financial resources.

# References

CCSP (Climate Change Science Program). 2003. Strategic Plan for the U.S. Climate Change Science Program. Washington, D.C.: Climate Change Science Program.

Christy, J.R., R.W. Spencer, W.B. Norris, W.D. Braswell, and D.E. Parker. 2003. Error estimates of Version 5.0 of MSU/AMSU bulk atmospheric temperatures. Journal of Atmospheric and Oceanic Technology 20:613-629.

GCOS (Global Climate Observing System). 2003. The Second Report on the Adequacy of the Global Observing Systems for Climate in Support of the UNFCCC. Available at http://www.wmo.ch/web/gcos/Second_Adequacy_Report.pdf. Accessed May 4, 2004.

Gibson, J.K., P. Kallberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano. 1997. ERA description. ECMWF Reanalysis Project Rep. 1, 72 pp. Available at http://www.ecmwf.int/assets/icons/webapp/tgz.gif. Accessed May 4, 2004.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, B. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph. 1996. The NCEP/NCAR 40-Year Reanalysis Project. Bull. Amer. Meteorol. Soc. 77:437-472.

Kanamitsu, M., W. Ebisuzaki, J. Woolen, J. Potter, and M. Fiorino. 2000. An overview of NCEP/DOE Reanalysis-2. Proc. 2nd Intl. Conf. on Reanalyses, Reading, England, 23-27 Aug. 1999. WCRP-109 (WMO/TD-985) WMO, Geneva, 1-4.

Keeling, C.D. 1998. Rewards and penalties of monitoring the Earth. Annu. Rev. Energy Environ. 23:25-82.

Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne, and M. Fiorino. 2001. The NCEP/NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation. Bull. Amer. Meteorol. Soc. 82:247-268.

NOAA (National Oceanic and Atmospheric Administration). 1995. NOAA Strategic Plan: A Vision for 2005. Washington, D.C.: NOAA.

NOAA. 2001. Final Report of the NOAA Science Advisory Board's Panel on Strategies for Climate Monitoring. Washington, D.C.: NOAA.

NRC (National Research Council). 1999a. Global Environmental Change: Research Pathways for the Next Decade. Washington, D.C.: National Academy Press.

NRC. 1999b. Adequacy of Climate Observing Systems. Washington, D.C.: National Academy Press.

NRC. 1999c. Making Climate Forecasts Matter. Washington, D.C.: National Academy Press.

NRC. 2000a. Issues in the Integration of Research and Operational Satellite Systems for Climate Research: Part I. Science and Design. Washington, D.C.: National Academy Press.

NRC. 2000b. Ensuring the Climate Record from the NPP and NPOESS Meteorological Satellites. Washington, D.C.: National Academy Press.

NRC. 2000c. Issues in the Integration of Research and Operational Satellite Systems for Climate Research: Part II. Implementation. Washington, D.C.: National Academy Press.

NRC. 2000d. Reconciling Observations of Global Temperature Change. Washington, D.C.: National Academy Press.

NRC. 2000e. Improving Atmospheric Temperature Monitoring Capabilities. Washington, D.C.: National Academy Press.

NRC. 2000f. The Oil Spill Recovery Institute: Past, Present, and Future Directions. Washington, D.C.: National Academy Press.

NRC. 2001. Climate Change Science: An Analysis of Some Key Questions. Washington, D.C.: National Academy Press.

NRC. 2002. A Century of Ecosystem Science: Planning Long-Term Research in the Gulf of Alaska. Washington, D.C.: National Academy Press.

NRC. 2003a. Satellite Observations of the Earth's Environment. Washington, D.C.: The National Academies Press.

NRC. 2003b. Fair Weather: Effective Partnerships in Weather and Climate Services. Washington, D.C.: The National Academies Press.

NRC. 2003c. Planning Climate and Global Change Research: A Review of the Draft U.S. Climate Change Science Program Strategic Plan. Washington, D.C.: The National Academies Press.

Rao, C.R., and J. Chen. 1995. Inter-satellite calibration linkages for the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on the NOAA-7,9, and -11 spacecraft. International Journal of Remote Sensing 16:1931-1942.

Schubert, S.D., R.B. Rood, and J. Pfaendtner. 1993. An Assimilated Data Set for Earth Science Applications. Bull. Amer. Meteorol. Soc. 74:2331-2342.

Standish Group. 1999. CHAOS: A Recipe for Success. Available online at http://www.standishgroup.com/sample_research/PDFpages/chaos1999.pdf. Accessed March 5, 2004.

Vinnikov, K.Y., and N.C. Grody. 2003. Global warming trend of mean tropospheric temperature observed by satellites. Science 302:269-272.

Weisberg, S.B., T.L. Hayward, and M. Cole. 2000. Towards a U.S. GOOS: A synthesis of lessons learned from previous coastal monitoring effort. Oceanography 13:54-61.

# APPENDIXES

# A

# Workshop Agenda and Participants

## WORKSHOP AGENDA

### Thursday, August 21

| | |
|---|---|
| 8:00 – 8:30 | Continental Breakfast |
| 8:30 – 8:45 | Welcome [Dave Robinson] |
| 8:45 – 9:30 | Plenary Talk #1: NOAA Satellite CDR Plan [Mitch Goldberg, John Bates] |
| 9:30 – 9:45 | Plenary Talk: Lessons from the NCEP/NCAR Reanalyses [Eugenia Kalnay] |
| **10:00 – 12:30** | **Session #1: Meeting User Needs** |
| 10:00 – 10:35 | Plenary Talk: Issues in Climate Data Records from satellite observations [Kevin Trenberth] |
| 10:35 – 10:45 | Comments from Greg Withee |
| 11:05 – 12:30 | Breakout Sessions<br>1 – Climate monitoring<br>2 – Model validation and development<br>3 – User applications |
| 12:30 – 1:30 | Lunch |
| **2:00 – 4:30** | **Session #2: Attributes of Successful CDRs** |
| 2:00 – 2:30 | Plenary Talk: What are the key attributes of successful CDR generation programs? [Graeme Stephens] |
| 2:45 – 4:15 | Breakout Sessions<br>1 – CDR principles<br>2 – Data management<br>3 – Assimilation/integration (w/other satellites, in situ measurements, multivariate data) |
| 4:30 – 4:45 | Day 1 Closing Remarks [Dave Robinson] |

### Friday, August 22

| | |
|---|---|
| 8:00 - 8:30 | Continental Breakfast |
| 8:30 – 9:00 | Breakout Session #1 Group Discussion |
| 9:00 – 9:30 | Breakout Session #2 Group Discussion |
| **9:30 – 12:30** | **Session #3: CDR Production Strategies** |
| 9:30 – 10:00 | Plenary Talk: What are the advantages and disadvantages of different models or strategies for producing CDRs, such as using partnerships among government, academia, and the private sector, different blends of space-based and in situ data (e.g., all space-based versus some balance) or other approaches? [Bill Rossow] |
| 10:15 – 11:45 | Breakout Sessions |
| | 1 – Biosphere |
| | 2 – Hydrosphere |
| | 3 – Energy |
| 12:00 – 12:30 | Breakout Session #3 Group Discussion |
| 12:30 | Closing Remarks [Dave Robinson] |

### WORKSHOP PARTICIPANT LIST

Alvin Miller, NOAA
Andrew Heidinger, NOAA
Antonio Busalacchi, University of Maryland
Arnold Gruber, NOAA
Chet Koblinsky, NASA
Chris Elvidge, NOAA
Dan Tarpley, NOAA
Dave Thompson, Colorado State University
Dorothy Hall, NASA
Dudley Chelton, Oregon State University
Ed Harrison, PMEL
Eugenia Kalnay, University of Maryland
Forrest Hall, University of Maryland
Frank Muller-Karger, University of South Florida
George Ohring, NOAA
Gerald Dittberner, NOAA
Graeme Stephens, Colorado State University
Greg Withee, NOAA
Herb Jacobowitz, Short and Associates

Herb Kroehl, NOAA
Jim McGuire, NPOESS IPO
Jim Miller, NOAA
Jim Silva, NOAA
John Janowiak, NOAA
John Townshend, University of Maryland
Juri Knyazikhin, Boston University
Ken Casey, NOAA
Ken Knapp, NOAA
Kenneth Casey, NOAA
Kent Hughes, NOAA
Kevin Trenberth, NCAR
Kory Priestley, NASA
Lisa Botluk, NOAA
Marie Colton, NOAA
Martha Bodden, Mitretek
Mary Ann Esfandieri, NASA
Mike Haas, NPOESS/IPO
Mitch Goldberg, NOAA
Norm Loeb, NASA
R.E. Murphy, NASA
Rachel Pinker, University of Maryland
Russ Rew, UCAR
Shobha Kondragunta, NOAA
Sky Yang, NOAA
Steve Mango, NPOESS/IPO
Sue Russell, Mitretek
Todd Mitchell, University of Washington
Watson Gregg, NASA
William Rossow, NASA

# B

# Biographical Sketches of Committee Members

**David Robinson** is the chair of the Geography Department at Rutgers University. He received his Ph.D. from Columbia University. Dr. Robinson has expertise in the collection and archiving of accurate climatic data, and he is interested in climate change (particularly state and regional climate issues), hemispheric and regional snow cover dynamics, interactions of snow cover with other climate elements, and the dynamics of solar and terrestrial radiative fluxes at and close to the surface of Earth. He is the author or co-author of approximately 130 articles, over half in peer-reviewed journals and book chapters. Dr. Robinson also is the State Climatologist for New Jersey.

**Roger Barry** is a professor of geography and the director of the National Snow and Ice Data Center World Data Center for Glaciology, Boulder, and he is rostered in the Cooperative Institute for Research in Environmental Sciences at the University of Colorado. He received his Ph.D. from the University of Southampton (U.K.). His major interests are in Arctic climate, cryosphere-climate interactions, mountain climate, and climatic change. Dr. Barry is a fellow of the American Geophysical Union and a foreign member of the Russian Academy of Natural Sciences. He serves as co-vice-chair of the Scientific Steering Group for the World Climate Research Programme's project on Climate and Cryosphere and is a member of the Terrestrial Observations Panel for Climate. He also serves on the editorial boards of *Physical Geography* and *Polar Geography*. Dr. Barry is fluent in French, German, and Russian, and he has been a Fulbright teaching scholar at Moscow State University in Russia. He has also held visiting appointments at ETH (Zurich), the Alfred Wegener Institute for Marine and Polar Research (Bremerhaven), the Climatic Research Unit at University of East Anglia (U.K.), the Institute of Astronomy and Geophysics at the University of Louvain-la-Neuve (Belgium), the Department of Geography at the University of Canterbury (New Zealand), and the Department of Biogeography and Geomorphology at the Australian National University (Canberra).

**Janet Campbell** and her research team are developing techniques for studying biological and biogeochemical processes in the ocean using satellite remote sensors. Their primary sources of data are ocean color satellite sensors such as the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and the Moderate Resolution Imaging Spectroradiometer (MODIS). They are modeling the effects of phytoplankton, dissolved organic materials, suspended sediments, and other particles on the spectral radiance measured by these satellites, and are exploring inversion techniques for using the satellite ocean color data to map these substances. Techniques are being developed for estimating primary productivity in coastal waters, and for blending regional models for coastal applications. Dr. Campbell is a member of NASA's SeaWiFS and MODIS science teams. As a member of the MODIS team she is responsible for developing algorithms and strategies for monitoring chlorophyll and primary productivity in coastal ocean, estuarine, and inland waters. Dr. Campbell has been an associate research professor at the University of New Hampshire (UNH) since 1993, and is a member of the Graduate Faculty. Between 1997 and 1999, she served as the NASA program manager for ocean biology and biogeochemistry. Before coming to UNH she was a research scientist at the Bigelow Laboratory for Ocean Sciences in Boothbay Harbor, Maine (1982-1993), where she established and directed the remote-sensing computer facility. She previously worked as an aerospace technologist and engineer at the NASA Langley Research Center in Hampton, Virginia. She holds a Ph.D. in statistics from Virginia Polytechnic Institute.

**Ruth DeFries** is an associate professor at the University of Maryland, College Park, with joint appointments in the Department of Geography and the Earth System Science Interdisciplinary Center. She investigates the relationships between human activities, the land surface, and the biophysical and biogeochemical processes that regulate Earth's habitability. She is interested in observing land cover and land use change at regional and global scales with remotely sensed data and exploring the implications for such ecological services as climate regulation, the carbon cycle, and biodiversity. Dr. DeFries obtained a Ph.D. in 1980 from the Department of Geography and Environmental Engineering at Johns Hopkins University and a bachelor's degree in 1976 from Washington University with a major in Earth science. Dr. DeFries has worked at the National Research Council with the Committee on Global

Change and has taught at the Indian Institute of Technology in Bombay. She is a fellow of the Aldo Leopold Leadership Program.

**William J. Emery** is a professor at the Colorado Center for Astrodynamics Research in the Department of Aerospace Engineering at the University of Colorado. He received his Ph.D. from the University of Hawaii. His research interests are in satellite remote sensing of the ocean and land surface vegetation. Ocean applications include skin sea surface temperature, the computation of surface currents from satellite images, mapping of geostrophic currents from satellite altimetry, and general air-sea interaction studies. The goal of Dr. Emery's research is to make satellite data a source of quantitative information that can be incorporated into numerical models of the phenomena controlling these systems. Dr. Emery has served on many panels looking into creation of long-term climate records from satellite data.

**Milton Halem** holds an emeritus position as Distinguished Information Scientist with the Earth Science Directorate at the NASA Goddard Space Flight Center (GSFC). Dr. Halem formerly served as assistant director for information sciences and as chief information officer for the GSFC. Dr. Halem has also served as chief of the Earth and Space Data Computing Division, where he was responsible for the development and management of the NASA Center for Computational Sciences, one of the world's most powerful complexes for scientific data intensive supercomputing and massive data storage. He acquired his Ph.D. in mathematics from the Courant Institute of Mathematical Sciences at New York University in 1968. He joined NASA in 1971 as the Global Atmospheric Research Program (GARP) project scientist and later headed up the Goddard Global Modeling and Simulation Branch. His personal achievements include more than 100 scientific publications in the areas of atmospheric and oceanographic sciences and computational and information sciences. He is most noted for his groundbreaking research in simulation studies of space-observing systems and for the development of four-dimensional data assimilation for weather and climate prediction. He has earned numerous awards, including the NASA Medal for Exceptional Scientific Achievement (twice), the NASA Medal for Outstanding Leadership (NASA's highest award), the NASA Distinguished Service Medal, and an honorary doctorate from Dalhousie University (Canada).

**James Hurrell** is a scientist at the National Center for Atmospheric Research's Climate and Global Dynamics Division, Climate Analysis Section. His research interests include climate variability and anthropogenic climate change. He has contributed to the Intergovernmental Panel on Climate Change assessments, and is actively involved in the international research program on Climate Variability and Predictability. Dr. Hurrell has a Ph.D. in atmospheric science from Purdue University. He has received the Clarence Leroy Meisinger Award from the American Meteorological Society, the Distinguished Alumni Award from the University of Indianapolis, and he is a fellow of the Royal Meteorological Society. He participated in the NRC's Panel on the Global Energy and Water Cycle Experiment.

**Arlene Laing** is an assistant professor in the Department of Geography at the University of South Florida. She received a Ph.D. in meteorology from Pennsylvania State University. Her research interests are in mesoscale convective systems, wildfire forecasting, satellite estimates of hurricane rainfall, and flash flood mitigation. Dr. Laing has a rich history working with operational and research-quality data, and limitations of each. She has been a visiting scientist at NASA Goddard Space Flight Center and is currently a visiting scientist at the National Center for Atmospheric Research (NCAR) Mesoscale and Microscale Meteorology Division. She received the Max Eaton Award from the American Meteorological Society for her paper on the global population of mesoscale convective complexes, and currently serves on its Committee on Satellite Meteorology and Oceanography.

**Roberta Balstad Miller** is a senior research scientist at Columbia University and director of the university's Center for International Earth Science Information Network (CIESIN). Dr. Miller has published extensively on science policy, information technology and scientific research, remote- sensing applications and policy, and the role of the social sciences in understanding global environmental change. She received her Ph.D. from the University of Minnesota, and has been a senior fellow at Oxford University and a guest scholar at the Woodrow Wilson International Center for Scholars. Dr. Miller has also been the director of the Division of Social and Economic Sciences at the National Science Foundation, the founder and first executive director of the Consortium of Social Science Associations, and president and CEO of CIESIN before it joined Columbia University. She has lectured widely both in the United States and abroad. She has been the vice president of the International Social Science Council and has served as chair of the NRC Steering Committee on Space Applications and Commercialization, the NATO Advisory Panel on Advanced Scientific Workshops/Advanced Research Institutes, the American Association for the Advancement of Science's

Committee on Science, Engineering, and Public Policy, and the Advisory Committee of the Luxembourg Income Study.

**Ranga Myneni** is an associate professor with the climate and vegetation group in the Geography Department at Boston University. He received his Ph.D. in biology from the University of Antwerp in Belgium. Dr. Myneni's research examines vegetation cover over Earth as observed from satellites, and he recently has worked extensively with Advanced Very High Resolution Radiometer and Moderate Resolution Imaging Spectroradiometer data.

**Richard Somerville's** major research interest is global climate change. He is a specialist in computer modeling of the climate system. He obtained a Ph.D. in meteorology from New York University and has been a professor at Scripps Institution of Oceanography, University of California, San Diego since 1979. In recognition of his accomplishments in scientific research Dr. Somerville has been elected a fellow of both the American Association for the Advancement of Science and the American Meteorological Society. He is also listed in *Who's Who in America*. The results of his research have been published in more than 100 technical papers. In addition, he has written a nontechnical book explaining topics such as the ozone hole and the greenhouse effect, titled *The Forgiving Air: Understanding Environmental Climate Change*. Among his many honors was his designation as the Walter Orr Roberts Lecturer in Interdisciplinary Sciences for 1999 by the American Meteorological Society "in recognition of significant contributions to the understanding of atmospheric processes derived from multidisciplinary research activities."

**Paul D. Try** is the senior vice-president and program manager at Science and Technology Corporation (STC) and the director of the International Global Energy and Water Cycle Experiment Project Office. He received his Ph.D. in atmospheric sciences from the University of Washington. Dr. Try has expertise in meteorological in situ and remote sensors (satellite and radar), as well as data collection, processing, exchange, and archival. Before joining STC he served in the U.S. Air Force, where he provided oversight management of all DOD research and development in environmental sciences. Dr. Try is a fellow of the American Meteorological Society and was its president in 1996-97.

**Thomas Vonder Haar** is a Distinguished Professor in the Department of Atmospheric Science at Colorado State University (CSU). He received a

Ph.D. in meteorology from the University of Wisconsin. Dr. Vonder Haar's research interests lie in the areas of global energy budget, remote sensing from satellites, local area forecasting, and geosciences. His work has included some of the first results of the direct solar irradiance measurements from satellites and the exchange of energy between Earth and space. Studies on the interaction of clouds and radiation and the general circulation have formed a basis for national and international plans leading to the Global Energy and Water Experiment and programs related to global change. Dr. Vonder Haar developed and directs CSU's Satellite Earth station to support research on storms at all scales. He recently coauthored the new textbook *Satellite Meteorology, an Introduction*, and he is the director of the Cooperative Institute for Research in the Atmosphere. He also is chairman of the World Climate Programme Working Group on Radiation Fluxes, a member of several NASA science teams, and a member of the Science Steering Group for the Global Energy and Water Cycle Experiment. He has received the American Meteorological Society Second Half Century (Charney) Award, the Abell Faculty Research and Graduate Program Support Award, the Engineering Dean's Council Award, and the CSU Distinguished Professor designation. He sits on the Council and Executive Committee of the American Metrological Society and the Board of Trustees of the University Corporation for Atmospheric Research (UCAR). He was recently elected to membership in the National Academy of Engineering.

## STAFF

**Sheldon Drobot** has been a program officer at the Polar Research Board and the Board on Atmospheric Sciences and Climate since December 2002. He received his Ph.D. in geosciences (climatology specialty) from the University of Nebraska, Lincoln. Dr. Drobot has directed NRC studies on *Elements of a Science Plan for the North Pacific Research Board* and *A Vision for the International Polar Year 2007–2008*. His research interests include sea ice-atmosphere interactions, microwave remote sensing, statistics, and long-range climate outlooks. Dr. Drobot currently is researching interannual variability and trends in Arctic sea ice conditions and how low-frequency atmospheric circulation affects sea ice distribution, short-range forecasting of Great Lakes ice conditions, and biological implications of sea ice variability.

**Rob Greenway** has been a project assistant at the National Academies since 1998. He received his M.Ed. in English education and his A.B. in English from the University of Georgia.

# C

# Acronyms and Initialisms

| | |
|---|---|
| ACSYS | Arctic Climate System |
| ADEOS | Advanced Earth Observing Satellite |
| AIRS | Atmospheric Infrared Sounder |
| AMSR | Advanced Microwave Scanning Radiometer |
| AMSU | Advanced Microwave Sounder Unit |
| ARM | Atmospheric Measurement Program |
| ASOS | Automated Surface Observing System |
| ATBD | Algorithm Theoretical Basis Document |
| AVHRR | Advanced Very High Resolution Radiometer |
| BAA | Broad Agency Announcement |
| BS | Bootstrap sea ice algorithm |
| BUFR | Binary Universal Form for the Representation of meteorological data |
| CCRI | Climate Change Research Initiative |
| CCSP | Climate Change Science Plan |
| CDAS | climate data assimilation system |
| CDSC | Climate Data Science Council |
| CDST | Climate Data Science Team |
| CDC | Climate Diagnostic Center |
| CDR | climate data record |
| CEOS | Committee on Earth Observing Satellites |
| CLASS | Comprehensive Large Array-data Stewardship System |
| CliC | Cryosphere and Climate |
| CLIVAR | Climate Variability and Predictability |
| COMET | Cooperative Program for Operational Meteorology, Education and Training Outreach Program |

| | |
|---|---|
| CSTAR | Collaborative Science, Technology, and Applied Research Program |
| CZCS | Coastal Zone Color Scanner |
| | |
| DAAC | Distributed Active Archive Center |
| DLESE | Digital Library for Earth Science Education |
| DLT | Digital Linear Tape |
| DMSP | Defense Meteorological Satellite Program |
| DOD | U.S. Department of Defense |
| DOE | U.S. Department of Energy |
| | |
| ECHO | EOSDIS Clearing House |
| ECMWF | European Centre for Medium-Range Weather Forecasts |
| ECS | EOSDIS Core System |
| EDR | environmental data record |
| EOS | Earth Observing System |
| EOSDIS | EOS Data and Information System |
| ERBE | Earth Radiation Budget Experiment |
| ESA | European Space Agency |
| ESMR | Electrically Scanning Microwave Radiometer |
| EUMETSAT | European Organisation for the Exploitation of Meteorological Satellites |
| | |
| FAQ | Frequently Asked Questions |
| FCDR | fundamental climate data record |
| | |
| GACP | Global Aerosol Climatology Project |
| GCM | General Circulation Model |
| GCOS | Global Climate Observing System |
| GEO | Group on Earth Observations |
| GEWEX | Global Energy and Water Cycle Experiment |
| GIS | Geographic Information System |
| GMS | Geostationary Meteorological Satellite |
| GOES | Geostationary Operational Environmental Satellites |
| GPCP | Global Precipitation Climatology Project |
| GPCC | Global Precipitation Climatology Center |
| GSFC | Goddard Space Flight Center |
| GVaP | Global Water Vapor Project |
| | |
| HIRS | High Infrared Sounder |

| | |
|---|---|
| HSB | Humidity Sounder for Brazil |
| ICSU | International Council of Scientific Unions |
| IDPS | Integrated Data Processing System |
| IGOS | International Global Observing Strategy partnership |
| IMS | Interactive Multisensor Snowmap |
| IOC | International Oceanographic Commission |
| IOOS | Integrated Ocean Observing System |
| IPO | NPOESS Integrated Program Office |
| IR | infrared |
| IRI | International Research Institute |
| ISCCP | International Satellite Cloud Climatology Project |
| ISLSCP | International Satellite Land-Surface Climatology Project |
| ISO | International Organization for Standardization |
| JAXA | Japan Aerospace Exploration Agency |
| LTA | long-term archive |
| MOBY | Marine Optical Buoy |
| MODIS | Moderate Resolution Imaging Spectroradiometer |
| MSU | Microwave Sounder Unit |
| NASA | National Aeronautics and Space Administration |
| NCAR | National Center for Atmospheric Research |
| NCDC | National Climatic Data Center |
| NCEP | National Center for Environmental Prediction |
| NCTP | NOAA Climate Transition Program |
| NDVI | Normalized Difference Vegetation Index |
| NESDIS | National Environmental Satellite, Data, and Information Service |
| NEXRAD | Next Generation Radar |
| NIDIS | National Integrated Drought Information System |
| NGDC | National Geophysical Data Center |
| NOAA | National Oceanic and Atmospheric Administration |
| NODC | National Ocean Data Center |
| NOPP | National Ocean Partnership Program |
| NORLC | National Ocean Research Leadership Council |
| NPOESS | National Polar-Orbiting Operational Satellite System |
| NPP | NPOESS Preparatory Project |

| | |
|---|---|
| NRC | National Research Council |
| NSDL | National Science Digital Library |
| NSF | National Science Foundation |
| NSIDC | National Snow and Ice Data Center |
| NT | NASA Team sea ice algorithm |
| NT2 | NASA Team sea ice algorithm 2 |
| NWP | Numerical Weather Prediction |
| NWS | National Weather Service |
| | |
| OAR | Office of Atmospheric Research |
| OCTS | Ocean Color and Temperature Sensor |
| OFCM | Office of the Federal Coordinator for Meteorological Services and Supporting Research |
| OLR | Outgoing Long-wave Radiation |
| ORAP | Ocean Research Advisory Panel |
| OSC | Orbital Sciences Corporation |
| OSDPD | Office of Satellite Data Processing and Distribution |
| | |
| PODAG | Polar DAAC User Working Group |
| POES | Polar Operational Environmental Satellite |
| PPDS | pilot planetary data system |
| | |
| RFI | Request for Information |
| RSS | Remote Sensing Systems |
| | |
| SBSTA | Subsidiary Board on Scientific and Technical Assessment |
| SDR | sensor data record |
| SeaWiFS | Sea-viewing Wide Field-of-View Sensor |
| SIMBIOS | Satellite Intercomparison and Merger for Biological and Interdisciplinary Ocean Science |
| SIRS A | Space Infra-Red Sounder |
| SMMR | Scanning Multichannel Microwave Radiometer |
| SPARC | Stratospheric Processes and Their Role in Climate |
| SR | Scanning Radiometer |
| SSM/I | Special Sensor Microwave/Imager |
| SST | sea surface temperature |
| SSU | stratospheric sounder unit |
| | |
| TCDR | thematic climate data record |
| THREDDS | Thematic Real-time Environmental Distributed Data Services |

| | |
|---|---|
| TIROS N | Television InfraRed Operational Satellite Next-generation |
| TOA | Top of Atmosphere |
| TOVS | TIROS Operational Vertical Sounder |
| TRMM | Tropical Rainfall Measuring Mission |
| T2 | Global average temperature for the lower troposphere |

| | |
|---|---|
| UAH | University of Alabama, Huntsville |
| UNFCC | United Nations Framework Convention on Climate Change |
| USGCRP | U.S. Global Change Research Program |

| | |
|---|---|
| VIIRS | Visible/Infrared Imager Radiometer Suite |
| VTPR | Vertical Temperature Profile Radiometer |

| | |
|---|---|
| WCRP | World Climate Research Programme |
| WDC | World Data Center |
| WGA | Western Governors Association |
| WMO | World Meteorological Organization |

# D

# EOSDIS Lessons Learned

## HISTORICAL BACKGROUND

In the late 1970s through the mid-1980s NASA began supporting a series of pilot data system studies to develop publicly accessible electronic data systems. These included such discipline-based systems as the Space Physics and Astrophysics Network, the Pilot Climate Data System, the pilot planetary data system (PPDS), the International Satellite Cloud Climatology Project (ISCCP), the Pilot Land Data System, etc. As NASA entered the age of launching great astronomical observatories in the late 1980s, the notion of specialized information processing and distribution centers emerged. These centers were organized around such instrument spectral domains as the visible spectral data at the Hubble Science Telescope Institute at Johns Hopkins University, the Infra Red Center at Cal tech, the High Energy X Ray Institute at the Marshall Space Flight Center, and the Upper Atmosphere Research Satellite instrument-processing teams. As part of the congressional approval of the EOS mission in 1990, the NASA Earth Science Enterprise committed to supporting the development of a long-term comprehensive data and information system (EOSDIS) whose products would be easily accessible both by the science research community and the broader public. Based on the information systems experience gained, the EOSDIS system design would employ a distributed open architecture. In addition to its functional requirements for space operations control and product generation for EOS, the EOSDIS would be responsible for the data archival, management, and distribution of all NASA Earth science mission instrument data (including EOS) during the mission life. The system would be organized as an integrated collection of distributed active archive centers (DAACs) providing the data services and interfaces with the user community. A common and core infrastructure of hardware and software capabilities would constitute the EOSDIS Core System that would be geographically distributed at the eight DAACs.  Each DAAC would be focused mainly on a particular Earth system component or discipline, such as atmospheres, oceans, land, snow and ice, hydrology, radiation and chemistry, and even socio-economic influences.

## NEED FOR THE EOSDIS

Four primary spacecraft make up the long-term measurement system of the EOS mission. They are Terra, launched in December 1999, Aqua launched in May 2002, ICESat launched in January 2003, and Aura to be launched in 2004. In addition to the processing of the instrument data from these satellites, the EOSDIS has responsibility for the archival and management of all NASA Earth science mission data products prior to EOS as well as data from NASA instruments flown on foreign satellites. The NASA Earth Science Enterprise is responsible for assuring the long-term permanent preservation of these data and has negotiated agreements with the operational agencies (NOAA and USGS) for their permanent retention. EOSDIS will support migration of the data to these agencies.

The EOSDIS performs flight operations for the above four EOS spacecraft; processes, archives, and distributes data from 17 instruments on six EOS spacecraft; and archives and distributes data from more than 40 instruments from more than 15 EOS and non-EOS spacecraft. The system today serves approximately 2 million users per year internationally, archiving almost 4 terabytes per day, distributing about 2 terabytes per day and maintaining the current archive, which is larger than 3 petabytes and is growing. In addition, the system supports 1,800 different data types, manages some of the nation's largest spatial databases, interfaces with over 35 external systems and depends on more than a million lines of code, with more than 60 commercial off-the-shelf products integrated with custom code deployed on a variety of vendor servers. This system is unprecedented in scope and scale for a NASA mission, and one of the largest, if not the largest, working civilian scientific data system ever built.

What distinguishes the EOSDIS from any of the above space mission data systems are the massive volumes of data ingested, processed into higher-level standard products and archived within hours to a few days of acquisition, and distributed to a broad user community on a routine basis.

Over the decade-long period of planning and implementation, the architecture and design of the EOSDIS have undergone nearly continuous evolution to incorporate new technologies and changing science requirements. In addition to managing a relatively large number of research instruments and satellites, EOSDIS provides the scientific data products to a broad community of users, including other value-added providers, such as the Federation of Earth Science Information Partners; the partners are responsible for the development and distribution of specialized and enhanced products for small, focused user communities. In terms of space flight operations and control, the EOSDIS has demonstrated that it can manage and execute some of the most intricate orbit executions by aligning these spacecraft into trains of operational and research satellites trailing and underflying each other separated in orbit by mere seconds to minutes in order to provide near simultaneity of observations of complementary instruments on disparate spacecraft. Previously management of such multi-instrument configurations and generation of data products from such measurements were possible only from a single satellite with a very large capacity. This capability has enriched the scientific mission capabilities at significantly reduced costs.

An additional contribution supported by the EOSDIS is the number of Pathfinder climate data studies from similar or nearly identical instruments flown since the inception of high-resolution satellite remote sensing in the early 1970s on multiple spacecraft from operational and research satellites, some spanning decades. The lessons learned provided by the EOSDIS itself as well as through the experience gained by supporting science instrument processing teams, the core DAAC processing capabilities, and the Earth science information partner processing capabilities afford ample examples to evaluate the advantages of and drawbacks to producing various datasets that should prove useful in the design of the NPOESS/NPP operational EDRs and CDRs.

## EOSDIS SYSTEM PERFORMANCE

The EOSDIS has undergone numerous reviews and budget reductions by scientific and data system committees. As a result of implementing their various recommendations, the current system configuration, scope, performance, and services have been dramatically changed from the original concept in terms of functionality, capability, and scale of communities served. The history of the EOSDIS has been stormy in terms of functionality and high expectations of various communities and criticisms have been numerous. One complaint often made concerns the high cost of the system compared with the data systems and operations costs for similar component products on other satellite systems. Another frequent argument made about the EOSDIS is the "one size fits all" customers' approach of the system.

While these arguments certainly have a large degree of validity, the system nevertheless has managed a timely delivery of the exceedingly demanding and scientifically credible data products. Today the system is routinely meeting its requirements supporting the EOS mission goals while reprocessing many of the datasets with improved algorithms and better calibration. It is not yet clear at this relatively early date in the expected lifetime of the EOS missions whether an information system organized around traditional dedicated mission data systems approaches, would have produced comparable performance more cost-effectively.

Another aspect of the system often overlooked are the synergistic capabilities afforded by the infrastructure resources and broader information science capabilities that can be brought to bear during such unanticipated events as fires, volcanoes, hurricanes, and floods. The breadth of the system has made possible a host of services because of its size (e.g., commercial adoption of data format standards for Earth science, specialized tools for geographic systems and visualization, system interoperability, global directories, and high-speed broadband EOS network accessibility for its user community). More recently the introduction of online data pools providing the most popular data products has led to a growing increase in their accessibility. The development of the EOS Clearing House with open applications programmer interfaces enables development of user interfaces tailored to specific communities; for example, MODIS provides L1 processing source code to direct broadcast users and have an open source code policy with respect to science algorithms. There has not yet been much demand for such software other than for direct broadcast stations.

The EOSDIS offers a large target during enterprise budget crunches and flight launch delays over the long term. Budget reductions in planned funding have forced scale-backs in the planned introduction of functionality and technological upgrades in system capabilities. Whether the traditional mission data approach with dedicated instrument or spacecraft systems or some modification of the present system will be more flexible in adapting services under such budget restraints is an open question.

## LESSONS LEARNED FROM EOSDIS

The NASA EOS, and in particular the Aqua satellite and its system of instruments, affords a highly representative collection of instruments with comparable numbers of spectral bands and spatial resolutions very similar to that being planned for the NPP. The class of scientific products produced by the EOSDIS are forerunners of the environmental data records expected from NPP. Thus, the lessons learned from the development and conduct of the EOSDIS as it evolved in the decades leading to the current system offers NESDIS an unparalleled opportunity to benefit from that experience in planning the production of NPOESS/NPP CDRs. The following lessons learned from the EOSDIS illustrate some of the management philosophies that can be used to sustain the system design architectures and the considerations of meeting evolving customer needs over decades as the evolution of technology, scientific requirements, and budgetary constraints change.

- **Science investigator-led processing systems.** A programmatic change in early 1998 transferred the responsibility for most EOS data processing from the DAACs (and the EOSDIS Core System) to EOS science instrument teams and their facilities. This transfer, accomplished through a call for proposals, was a major reason for the success and timely delivery of the EOS standard data products, and accounted for the high degree of scientific community acceptance of these products. One reason for the scientists' willingness to assume day-to-day involvement in operational data processing was that in many cases the principal investigators felt that they would be judged by their peers on the quality and timeliness of delivery of these products. This transfer did have significant implications in terms of budget allocations, delegation of management oversight, creeping science requirements growth, interface coordination, software configuration control, hardware and network resource growth, and security, to name just a few issues that had to be addressed.
- **Planned evolutionary upgrades.** The EOSDIS has changed significantly in architecture, design, and implementation since its original planned configuration. Infusion of more recent but mature information technology has made it possible to support the scope of data products and services without compromising the functionality of the EOSDIS and scope of mission under the ensuing budgetary constraints over the years. If anything, the functionality has expanded to support a much larger community than originally envisaged. In fact, a community of EOS partners has been established by NASA's Earth Science Enterprise to participate in broadening the EOSDIS in many different ways: data and portal providers; algorithm product processors and producers; data services and distribution nodes to research and educational users; value added providers; international and interagency centers; and low cost direct broadcast reception to universities, state and local agencies, and commercial organizations. It is this transforming functional evolution that is changing the popular misconception of the EOSDIS from a highly centralized, inflexible, cost-inefficient data and information

system to one that has a more open and distributed architecture both in terms of science processing and user applications. No longer is the EOSDIS developed, maintained, and restricted by outdated requirements and design processes. It has adopted a limited set of open source architecture concepts to address the current needs and capabilities as a natural evolution and to avoid having to define costly revised system versions or scrapping systems and restarting with a clean slate. While all of the source code is not available for anyone to modify and share, some of the modules developed as a part of EOSDIS have been reused by other organizations.

- **Program and project management.** Creating widely acceptable CDRs from NOAA operational satellites will be as difficult a science challenge as managing a data information system as complex as the EOSDIS. Garnering the full support by a diverse and broad representation of the science community from the initial proposed concept, plan, scope, and implementation for an approach is critical to the success of this NESDIS undertaking. Unfortunately, in developing the EOSDIS the science community was not completely supportive from the start and was unsatisfied with the approach of an EOSDIS Core System with noncompetitively selected DAACs for the scientific processing of higher-level data products. Another concern of the science community was the size of the EOSDIS budget being appropriated to a single large contractor responsible for the system development. The lack of an effective interface between the science community and a large centrally managed science information system developed by a large industrial contractor came as a culture shock. As a result various stakeholders found themselves engaged in conflict over priorities and requirements, with no realistic mechanism to reach closure between information technology system development teams, the science instrument teams, and external science communities. New systems must allow users to gain ownership of requirements through sponsored workshops to reach community consensus, and initiating processes to enable users to prioritize requirements allowed stakeholders to feel more comfortable with the direction the system and project were taking.
- **User working groups.** A valuable component of the DAAC activities has been the critical evaluation and directions provided to each DAAC by its User Working Group, appointed through NASA.
- **Incremental development.** The project could have been more effective if it had pushed for early operational releases with incremental growth in functionality. The first operational release of the EOSDIS Core System (ECS) was delivered to support Landsat 7/Terra and provided a complete, end-to-end capability. It was deployed over six years into the ECS contract

and was subject to many problems. Some of the difficulties were due to ill advised technology choices that would have become readily apparent in an operational environment (e.g., Distributed Computing Environment). Having a stable baseline of some core components would have made it easier to add additional capabilities. As evidence, the releases to support the Aqua and Aura missions have been delivered with progressively fewer problems, while adding new functionality. The early deployment of truly core components would have allowed and even fostered the development of value-added components from the broader environmental science and engineering and external data and information system development community.

- **Technology reuse.** Reuse of independently developed components is possible and has occurred within EOSDIS. Some examples are the EOS Data Gateway (reused from Version 0 EOSDIS), the Simple, Scalable, Script Based Science Processor (GSFC DAAC developed component now in the production system), Land Processes (EDC) DAAC-Billing and Accounting system for Landsat-7 (borrowed from USGS), SeaWiFS processing system adapted to MODIS and the Ozone Monitoring Instrument Science Investigation Processing System, GSFC DAAC-developed Version 0 systems reused in Aura SIPSs at the Jet Propulsion Laboratory. This was enabled by the maturation of a stable, base-lined ECS and the implementation and publication of standard interfaces to its components.

# E

# Previous NRC Recommendations

| **Report** | **Recommendations** |
|---|---|
| *Global Environmental Change: Research Pathways for the Next Decade* (NRC, 1998) | • The strategy for obtaining long-term observations designed to define the magnitude and character of Earth system change must be reassessed. Priority must be given to identifying and obtaining accurate data on key variables carefully selected in view of the most critical science questions and practically feasible measurement capabilities.<br>• The USGCRP must revitalize its strategy for the data systems used for global change research. Emphasis must be placed on designing and selecting flexible and innovative systems that appropriately reflect focused responsibility for data character that provide open access to the scientific community and the public, and that rapidly evolve to exploit technological developments. In particular, the USGCRP must closely monitor the progress of the innovative “federation” concept for data systems. |
| *Adequacy of Climate Observing Systems* (NRC, 1998) | • Stabilize the existing observational capability.<br>• Identify critical variables that are either inadequately or not measured at all.<br>• Build climate observing requirements into the operational programs as a high priority.<br>• Revamp existing climate research programs and some climate-critical parts of operational observing programs through the implementation of the ten principles of climate monitoring.<br>• Establish a funded activity for the development, implementation, and operation of climate-specific observational programs. |

| **Report** | **Recommendations** |
|---|---|
| *Issues in the Integration of Research and Operational Satellite Systems for Climate Research: Part I. Science and Design* (NRC, 2000) | • Climate research and monitoring capabilities should be balanced with the requirements for operational weather observation and forecasting within an overall U.S. strategy for future satellite observing systems.<br>• The Integrated Program Office for NPOESS should give increased consideration to the use of NPOESS for climate research and monitoring.<br>• The NASA Earth Science Enterprise should continue to play an active role in the acquisition and analysis of systematic measurements for climate research as well as in the provision of new technology for NPOESS.<br>• Joint research and operational opportunities such as the NPOESS Preparatory Project (NPP) should become a permanent part of the U.S. Earth observing remote sensing strategy. |
| *Ensuring the Climate Record from the NPP and NPOESS Meteorological Satellites* (NRC, 2000) | • NOAA should begin now to develop and implement the capability to preserve in perpetuity the basic satellite measurements.<br>• NOAA should guarantee climate researchers affordable access to all raw data records in the long-term archive, with an emphasis on large-volume data access.<br>• NASA, in cooperation with NOAA, should support the development and evaluation of CDRs, as well as their refinement through data processing.<br>• NOAA and NASA should define and develop a basic set of user services and tools to meet specific functions for the science community, with NOAA assuming increasing responsibility for this activity as data migrates to the long-term archive.<br>• NASA, in cooperation with the Integrated Program Office, should develop the NPOESS Preparatory Project as an integral component of a climate data system.<br>• NOAA, in cooperation with NASA, should invest in early, limited capability prototypes for both long-term archiving and the NPP data system.<br>• NASA and NOAA should develop and support activities that will enable a blend of distributed and centralized data and information services for climate research. |

| Report | Recommendations |
|---|---|
| *Issues in the Integration of Research and Operational Satellite Systems for Climate Research: Part II. Implementation* (NRC, 2000) | • Data should be supported by metadata that carefully document sensor performance history and data processing algorithms.<br>• The research community and government agencies should take the initiative and begin planning for a research-oriented NPOESS climate data system and the associated science participation.<br>• Quality assessment should be an intrinsic part of operational data production and should be provided in the form of metadata with the data product.<br>• Radiometric characterization of the Moon should be continued and possibly expanded to include measurements made at multiple institutions in order to verify the NASA results. If the new reflectance calibration paradigm is adopted, then the objective of the lunar characterization program should be to measure changes in the relative reflectance as a function of the phase and position of Earth, the Sun, and the Moon rather than absolute spectral radiance.<br>• The system should have the ability to reprocess large data sets as understanding of sensor performance, algorithms, and Earth science improves. Examples of sources of new information that would warrant data reprocessing include the discovery of processing errors, the detection of sensor calibration drift, the availability of better ancillary data sets, and better geophysical models.<br>• The results of sensitivity studies on the parameters in the data product algorithms should be summarized in a requirements document that specifies the characterization measurements for each channel in the sensor. Blanket specifications covering all channels should be avoided unless justified by the sensitivity studies.<br>• Competitive selection of instrument science teams should be adopted to follow the progress of the instrument from design and fabrication through integration, launch, operation, and finally, data archiving, thereby promoting more thorough instrument characterization.<br>• Science teams responsible for algorithm development, data set continuity, and calibration and validation should be selected via an open, peer-reviewed process (in contrast to the approach taken with the operational integrated data processing system [IDPS] and algorithms, which are being developed by sensor contractors for NPOESS).<br>• Validation, an essential part of the information system, should be undertaken for each data product or data record to provide a quantitative estimate of the accuracy of the product over the range of environmental conditions for which the product is provided. |

| Report | Recommendations |
|---|---|
| *Reconciling Observations of Global Temperature Change* (NRC, 2000) | • The nations of the world should implement a substantially improved temperature monitoring system that ensures the continuity and quality of critically important data sets.<br>• The scientific community should perform a more comprehensive analysis of uncertainties inherent in the surface, radiosonde, and satellite data sets.<br>• Natural as well as human-induced changes should be taken into account in climate model simulations of atmospheric temperature variability on decade-to-decade time scales.<br>• The scientific community should explore the possibility of exploiting the sophisticated protocols that are now routinely used to ensure the quality and consistency of the data ingested into the operational numerical weather prediction models to improve the reliability of the data sets used to monitor global climate change. |
| *Improving Atmospheric Temperature Monitoring Capabilities* (NRC, 2000) | • NESDIS should create a web site that includes information on spacecraft and instrument condition and changes that are of interest for the construction of CDRs. In addition to the official NESDIS TIROS Operational Anomaly Reports (TOAR), this site should be interactive to allow climate investigators to communicate their findings and opinions concerning the behavior of specific instruments and/or channels. The site should be well organized, with cross-referencing by category, and should include a good search capability to enable interested parties to find what they want. An attempt should be made to hierarchically construct the site so that issues judged by NESDIS to be of greatest importance to the climate record are most prominently featured. The information contained on this web site would become part of the permanent metadata record for each instrument.<br>• NESDIS should take responsibility for the construction and validation of CDR-quality bulk-layer temperature time series from the SSU and AMSU for the analysis of stratospheric climate variations and trends.<br>• NESDIS should establish for each POES operational instrument a structure by which the communication of information may be assured as CDRs are developed and refined. This could be implemented with the establishment of an ad hoc group of individuals who are involved in some way with the development of the instrument and the CDRs. Sponsored meetings or workshops, possibly with published proceedings, would help ensure that the right mix of people have access to one another. Another approach could be the formal establishment of, for example, an MSU/AMSU Climate Science Team which would afford the members the opportunity to deal with issues of calibration, validation, long-term stability, and future requirements for deep-layer atmospheric temperature (as well as other microwave-based products). The team could also advise NESDIS/NCDC on issues of data storage, data access, and which significant products to archive. |

| Report | Recommendations |
|---|---|
| | • A network performance monitoring system to identify both random errors and time-dependent biases in both space-based and in situ observing systems would enable NOAA and the scientific community to identify and correct errors as soon as possible in these critical observing systems. These diagnostics should become part of the metadata associated with the observations.<br>• NOAA should reinvigorate its efforts to "ensure a long-term climate record" (NOAA, 1995). This perspective should permeate the full range of activities related to climate observation, including instrument design and specification, instrument siting, specification of observing methods, data and metadata archiving, production and validation of CDRs, data analysis, and dissemination of products.<br>• NOAA should take responsibility for identifying proven CDRs and ensuring that the construction of these be maintained. In addition, NOAA should assume responsibility for supporting and developing the required scientific expertise, documenting the CDR construction methodology, and ensuring that the scientific expertise has been institutionalized, rather than merely residing with individual scientists. It is also important that the time series can be reproduced by future investigators.<br>• NOAA should put a high priority on measuring all aspects of the radiometer's system gain function and baseline offset during pre-launch testing. The usual set of thermal-vacuum tests should be expanded and done more rigorously, and the test results should be made readily available to the scientific community for evaluation. Since the calibration drift seems to be related to temperature, a sufficient number of precision thermistors should be mounted on the various radiometer components (antenna, feedhorn, front-end receiver, detector, etc.) for on-orbit monitoring and drift detection. More robust on-board calibration systems (e.g., additional reference loads) should be considered for future missions. |
| *Satellite Observations of the Earth's Environment* (NRC, 2003) | • NASA and NOAA should jointly work toward and should budget for an *adaptive* and *flexible* operational system in order to support the rapid infusion of new satellite observational technologies, the validation of new capabilities, and the implementation of new operational applications.<br>• A strong and effective Interagency Transition Office for the planning and coordination of activities of NASA and NOAA in support of transitioning research to operations should be established by and should report to the highest levels of NASA and NOAA.<br>• All NASA Earth science satellite missions should be formally evaluated in the early stages of the mission planning process for potential applications to operations in the short, medium, or long term, and resources should be planned for and secured to support appropriate mission transition activities.<br>• NOAA and NASA should improve and formalize the process of developing and communicating operational requirements and priorities. |

| **Report** | **Recommendations** |
|---|---|
| *Planning Climate and Global Change Research: A Review of the Draft U.S. Climate Change Science Program Strategic Plan* (NRC, 2003) | • Ensure the existence of a long-term observing system that provides a more definitive observational foundation to evaluate decadal- to century-scale variability and change.<br>• The observation system must include observations of key state variables such as temperature, precipitation, humidity, pressure, clouds, sea ice and snow cover, sea level, SST, carbon fluxes, and soil moisture.<br>• More comprehensive regional measurements of greenhouse gases would provide critical information about their local and regional source strengths. |