

September 1992

EARTH OBSERVING SYSTEM

Information on NASA's Incorporation of Existing Data Into EOSDIS



147776



**RESTRICTED--Not to be released outside the
General Accounting Office unless specifically
approved by the Office of Congressional
Relations.**

555 485

**Information Management and
Technology Division**

B-249927

September 25, 1992

The Honorable George E. Brown, Jr.
Chairman, Committee on Science,
Space, and Technology
House of Representatives

Dear Mr. Chairman:

On March 13, 1992, you requested that we provide information about the National Aeronautics and Space Administration's (NASA) progress in incorporating existing earth science data into the early version of the Earth Observing System Data and Information System (EOSDIS), known as Version 0. The incorporation of these data is important because the success of NASA's evolutionary development approach to EOSDIS depends on the early availability of useful data so that the scientific community at large can test the system's capabilities and provide feedback to NASA. This report discusses NASA's progress to date in this area. Details of our objectives, scope, and methodology are provided in appendix I.

Results in Brief

NASA plans to include a variety of data in the early EOSDIS, including reprocessed versions¹ of certain key large-volume data sets that have been collected over many years. However, few of the data sets earmarked for reprocessing will be sufficiently reworked to be available for users to test early EOSDIS capabilities and provide feedback. Although EOSDIS Version 0 is scheduled to become operational in 1994, none of the six large data sets selected for inclusion by the scientific community is likely to be fully reprocessed by that date. Many projects remain in the preliminary planning phase, with few details worked out regarding how their data sets will be made available through EOSDIS. Others do not even have preliminary plans or milestones because scientists are still debating various alternatives for reprocessing their data to make it useful within EOSDIS.

Background

The Earth Observing System (EOS) is the centerpiece of NASA's Mission to Planet Earth Program, whose goal is to obtain a scientific understanding of the entire earth as an integrated system and to determine the processes that contribute to environmental balance or change. Central to EOS is a

¹"Reprocessing" includes two major activities: (1) recalibrating the original data to eliminate instrument and processing artifacts and (2) creating new, highly refined subsets of these data that focus on specific aspects of earth science.

series of space-based observatories carrying a variety of instruments that will collect data about the earth's atmosphere, biosphere, oceans, and land surfaces as a complete system over a 15-year period, beginning in 1998. Data from these instruments will be processed and distributed to the scientific community through EOSDIS. In addition to the data from the EOS observatories, EOSDIS will include all existing NASA earth science data. EOSDIS is to make this vast wealth of data easily accessible and affordable to a broad range of researchers, thus aiding them in carrying out important new kinds of global change research.

EOSDIS Development Strategy

In keeping with the recommendations of the scientific community and to ensure that EOSDIS will satisfy the data and information needs of such a diverse, multidisciplinary group, NASA has adopted an evolutionary development approach.

Rather than determine all requirements up front, NASA is following a "build a little, test a little" methodology to iteratively define and validate how global change researchers will use the system. Building on existing earth science systems and data, NASA began work on Version 0, an in-house-developed set of early capabilities of the system, in 1990. This initial system is scheduled to be operational by mid-1994. NASA intends to award a contract for future versions in November 1992. Version 1, which will represent the first major version of the contractor-developed system, is scheduled to become operational in 1996.

The success of NASA's evolutionary development approach will require the extensive involvement of the scientific community and its close collaboration with system developers. To encourage these future users to participate in the development process, early versions of EOSDIS must offer easy and convenient access to a wide range of important global data sets so that the research community at large will be motivated to test the system and provide feedback on its usefulness. NASA's EOSDIS development approach is discussed at greater length in a previous GAO report.²

Expanding Availability of Data Is a Major Version 0 Goal

One of Version 0's primary objectives is to expand the amount of earth science data that is readily available to the research community. In order to do this, NASA officials set up scientific advisory groups to identify "pathfinder" data sets for early attention by the EOSDIS Project. The

²Earth Observing System: NASA's EOSDIS Development Approach Is Risky (GAO/IMTEC-92-24, Feb. 25, 1992).

pathfinder data sets are generally large data sets collected over a number of years by a series of operational satellites, such as the National Oceanographic and Atmospheric Administration's (NOAA) polar orbiting and geostationary satellites. These data sets are especially useful to global change researchers because they span enough years for scientists to detect climate trends. However, because these data have been collected over time and from more than one instrument and many satellites, they are not all calibrated and recorded the same way and must be carefully reprocessed before they can be generally useful to a wide range of researchers.

The Pathfinder Project's charter is to reprocess these data, recalibrating and correcting inconsistencies introduced by the different algorithms used in the original processing. The scientific community was engaged to help define, generate, and verify new geophysical data subsets from these basic data sets that would be applicable to the study of global change. These new "pathfinder" data subsets, along with the original data, are to be archived within EOSDIS.

Three pathfinder data sets were originally identified by EOS management in collaboration with the scientific community. These include the Advanced Very High Resolution Radiometer (AVHRR) data sets held by NOAA, TIROS Operational Vertical Sounder (TOVS) data held jointly by NOAA and NASA, and Geostationary Operational Environmental Satellite (GOES) data, held under NOAA contract by the University of Wisconsin. Recently added to the list are the Special Sensor Microwave/Imager (SSM/I) data acquired by NOAA from the Department of Defense, Scanning Multichannel Microwave Radiometer (SMMR) data recorded from the Nimbus-7 satellite, and Landsat data from the U.S. Geological Survey archive at the Earth Resources Observations System Data Center. EOSDIS user working groups may, in the future, identify additional candidate pathfinder data sets. However, their selection is ultimately an EOS management decision. The National Research Council, in its assessment of satellite earth observation programs in 1991 and its interim report on NASA's plans for EOSDIS, recommended that NASA "continue to develop existing 'pathfinder' data sets"³ and that "the Pathfinder data program under way should be accelerated."⁴

³Assessment of Satellite Earth Observation Programs, National Academy of Sciences' National Research Council, National Academy Press, 1991, p. 4.

⁴Interim Report of the National Research Council's Panel to Review Earth Observing Data and Information System (EOSDIS) Plans, April 9, 1992, p. 7.

Scientific working groups, composed of members of the community familiar with the data sets, have been formed and are charged with determining the scientific needs for the data and selecting "community consensus" algorithms for reprocessing them. Each working group is also responsible for recommending what data services will be required by users. So far NASA has spent approximately \$5 million on the pathfinder activities.

Few Pathfinder Data Sets Will Be Available Early Enough for Timely Feedback

Although work on the Pathfinder Project began in 1990, little progress has been made. Only the AVHRR data set, with the exception of the ocean data subset, is likely to be available for use under Version 0 in 1994. The full GOES data set, on the other hand, will not be available until 1995, followed by TOVS in 1997 or 1998. Only one of the three Landsat pathfinder data subsets will be available by 1996. Two of the projects are still in the preliminary planning phase.

The AVHRR pathfinder is a 12-year data set collected from instruments flown on the NOAA-series polar orbiting meteorological satellites. These data provide a continuous record of information about the earth's vegetation, atmosphere, and ocean surface temperature from August 1981 through the present. Reprocessing of AVHRR is farthest along and appears likely to have definitive results in time for users of Version 0 to be able to interact with the data. The 12-year AVHRR data set is currently being copied from magnetic tapes to optical disks at Goddard Space Flight Center. This voluminous data set will then have to be copied to another medium in order to be included in and accessible through the Version 0 system, an activity that will not be completed until 1994. The reprocessing of land and atmosphere data subsets is scheduled to begin in January and June 1993, respectively. Depending on system performance, it is anticipated that the land data subset will be available to Version 0 users by June 1994, and that the atmosphere subset will follow within a year.

Work on the AVHRR ocean data subset has not progressed as rapidly. Unlike the land and atmosphere projects, less consensus exists within the scientific community regarding reprocessing algorithms for the ocean data. As a result, the project charged with deriving the ocean data subset remains in a research phase. Scientists at the University of Miami are currently generating candidate algorithms and trial data products. The results will then be evaluated by a working group and an algorithm selected for generating a 1-year trial data set. No plans have yet been made for reprocessing the data from the remaining 11 years. Because the AVHRR

data is a multi-terabyte⁵ archive, NASA will have to dedicate many resources to reprocessing this data subset.

The TOVS pathfinder data set was collected by instruments also flown on the NOAA-series polar orbiting weather satellites. These data provide a continuous record since 1978 of the temperature and moisture of the earth around the globe and may provide evidence of climate change. However, the completely reprocessed data sets will not be available until 1997 at the earliest, i.e., 1 year after Version 1 of the operational EOSDIS is developed. Sharp disagreements among members of the working group concerning reprocessing methodologies have convinced project officials of the need for a period of research and validation. Accordingly, three groups of researchers are each taking different approaches to product derivation. The results of these three efforts will be assessed on an iterative basis as reprocessing takes place. Scientists may not be able to determine which approach is best until the entire TOVS data set is reprocessed.

The GOES data set is of particular interest to researchers studying regional climate variations. Collected by instruments flown on NOAA's east and west geostationary satellites, this data set offers a unique opportunity to observe the same geographic region of the globe for more than 15 years. Moreover, the GOES archive contains more than 100 terabytes, and reprocessing and managing these data will provide a challenge equivalent to that posed by EOS. The reprocessing of the GOES pathfinder data set is scheduled to begin in January 1993. However, the data set will require considerable cleaning up before useful products can be generated. The complete reprocessed set will be available in July 1995, 1 year after Version 0 becomes operational. It will be archived at the University of Wisconsin.

The Landsat archive contains high-resolution data, which provides capabilities for monitoring detailed aspects of the environment since the launch of the first Landsat satellite in 1972. Information about land cover and land cover change, water resources, oceanography, and marine resources can be derived from this data set.

The Landsat Pathfinder Project comprises three separate activities, two of which have not advanced beyond the planning phase. Science teams have just begun initial reprocessing of the complete data for the third activity, which is to derive deforestation data for the tropics. Final distribution

⁵One terabyte of data is approximately 1 trillion bytes. A byte is equivalent to one character in a text file, such as the letter "a."

products should be available to EOSDIS by 1996, 2 years after Version 0 comes on line and concurrently with the contract-developed Version 1.

The SSM/I and SMMR data sets contain information about atmospheric properties such as water vapor and rain rate, ocean surface properties such as surface wind speed, and land surface temperatures. Sea-ice and snow-cover parameters can also be derived from these data. The SSM/I and SMMR projects are still in the preliminary planning stages.

Observations

Global change research is an emerging and complex discipline. Traditionally, the earth science community has studied satellite data from specific missions using particular sensing techniques over short periods of time. Few researchers have experience working with long-term climate and global change data sets as voluminous as those in the Pathfinder Project, as evidenced by the inability of scientists in the TOVS working group to reach consensus concerning reprocessing of their data set. As a result, program officials were not aware of the complexity of the task when the Version 0 development strategy was formulated and milestones were fixed.

It appears from our review that NASA may be unable to provide researchers with full access to certain important, large data sets during the early Version 0 phase of the EOSDIS project. With only limited portions of these data sets available through EOSDIS, users may have less incentive to thoroughly use and test the early versions of EOSDIS. In developing EOSDIS, NASA has committed itself to a learning program dependent on early user interaction and feedback from the scientific community at large. Given that access to the key pathfinder data sets is likely to be limited, it will be important that NASA make certain that the Version 0 phase of EOSDIS development still provides sufficient experience with a wide range of users to assure that future full-scale versions of the system will indeed offer scientists the capabilities they need to do global change research.

We conducted our review between February and August 1992, in accordance with generally accepted government auditing standards. As requested, we did not provide a draft of this report to NASA for its review and comment. Instead, we discussed the report's contents with NASA officials, who objected to the presentation of some supplementary material. This material has been deleted from the final report. In addition, we have incorporated their other comments as appropriate.

As agreed with your office, unless you publicly announce the contents of this report earlier, we plan no further distribution until 30 days from the date of this letter. We will then send copies to the appropriate House and Senate committees; the Administrator, National Aeronautics and Space Administration; and other interested parties. Copies will also be made available to others upon request.

Please contact me at (202) 512-6240 if you have any questions concerning this report. Major contributors are listed in appendix II.

Sincerely yours,



Samuel W. Bowlin
Director, Defense and Security
Information Systems

Contents

Letter	1
Appendix I Objectives, Scope, and Methodology	10
Appendix II Major Contributors to This Report	11
Related GAO Products	12

Abbreviations

AVHRR	Advanced Very High Resolution Radiometer
EOS	Earth Observing System
EOSDIS	EOS Data and Information System
GAO	General Accounting Office
GOES	Geostationary Operational Environmental Satellite
IMTEC	Information Management and Technology Division
NASA	National Aeronautics and Space Administration
NOAA	National Oceanographic and Atmospheric Administration
SMMR	Scanning Multichannel Microwave Radiometer
SSM/I	Special Sensor Microwave Imager
TOVS	TIROS Operational Vertical Sounder

Objectives, Scope, and Methodology

In March 1992 the House Committee on Science, Space, and Technology requested that we review NASA's progress in incorporating existing earth science data into the early version of EOSDIS. To meet our objectives, we

- interviewed EOSDIS project officials at NASA headquarters, Washington, D.C.; and EOSDIS project officials at NASA's Goddard Space Flight Center, Greenbelt, Maryland;
- reviewed EOSDIS program documentation, discussing its goals and approach to EOSDIS development;
- reviewed the EOSDIS science data plan covering Version 0 and beyond;
- interviewed the managers of the EOSDIS data centers at the Goddard Space Flight Center, the Jet Propulsion Laboratory, and the Alaska Synthetic Aperture Radar Facility about their data plans;
- obtained draft project assessments from scientists involved in reprocessing the pathfinder data sets;
- reviewed reports prepared by the scientific working groups associated with the pathfinder data sets; and
- reviewed previous assessments of EOSDIS prepared by various scientific groups and committees.

We performed our work at NASA headquarters, Washington, D.C., and the Goddard Space Flight Center, Greenbelt, Maryland.

Major Contributors to This Report

Information
Management and
Technology Division,
Washington, D.C.

Ronald W. Beers, Assistant Director
John A. de Ferrari, Assignment Manager
Elizabeth L. Johnston, Evaluator-in-Charge

Related GAO Products

NASA: Changes to the Scope, Schedule, and Estimated Cost of the Earth Observing System (GAO/NSIAD-92-223, July 22, 1992).

EOS Data Policy: Questions Remain About U.S. Commercial Access (GAO/IMTEC-92-44, June 25, 1992).

Earth Observing System: Broader Involvement of the EOSDIS User Community is Needed (GAO/IMTEC-92-40, May 11, 1992).

Earth Observing System: NASA Needs to Reassess Its EOSDIS Development Strategy (GAO/IMTEC-92-7, Feb. 26, 1992).

Earth Observing System: NASA's EOSDIS Development Approach Is Risky (GAO/IMTEC-92-24, Feb. 25, 1992).

Earth Observing System: Information on NASA's Selection of Data Centers (GAO/IMTEC-91-67, Sept. 18, 1991).

Space Data: NASA's Future Data Volumes Create Formidable Challenges (GAO/IMTEC-91-24, Apr. 8, 1991).

Space Data: Information on Data Storage Technologies (GAO/IMTEC-90-88FS, Sept. 12, 1990).

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are \$2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

**U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20877**

Orders may also be placed by calling (202) 275-6241.

**United States
General Accounting Office
Washington, D.C. 20548**

**Official Business
Penalty for Private Use \$300**

**First-Class Mail
Postage & Fees Paid
GAO
Permit No. G100**
