Early EOSDIS:
A Data and Information System for the
Study of Global Change

center,tab(#); l l. Jeff Dozier#University of California #Sequoia 2000 Technical Report 91-8 #Center for Remote Sensing and Environmental Optics #Santa Barbara, CA 93106-3060

r l. Internet:#dozier@crseo.ucsb.edu OMNET:#J.DOZIER Telephone:#805 893 2309 Fax:#805 893 2578

INTRODUCTION

The initial step in the development of the Earth Observing System Data and Information System (EOSDIS) is to deliver a working prototype system for use by the Earth science research community by July 1994. EOSDIS will be NASA's contribution to the confederation of national and international agency data systems to support global change research and other uses of environmental data [Dozier and Ramapriyan, 1991], including GEWEX, the Global Energy and Water Cycle Experiment. This prototype Version 0 EOSDIS will not provide every feature that later systems will provide, so clear choices must be made as to priorities in its implementation. Technical readiness and affordability will determine how much capability the system will actually offer.

EOSDIS Version 0 (V0) is a fusion of data holdings, data services, and research community infrastructure. All must advance to improve the use of data to support global change research. The science priorities will guide the choice of data sets to emphasize and provide to the community and the levels of service to be encompassed within V0.

I. CONTEXT OF EOS AND EOSDIS

EOS, EOSDIS specifically, and the rest of the NASA Earth Science and Applications Program function in the context of the U.S. Global Change Research Program and an emerging web of international ties embracing satellite mission cooperation, major ground-based measurement programs, and the general coordination of research. Many of the satellite data that will be available for use in the EOSDIS V0 time frame will be provided by our foreign partners and other U.S. agencies—the European Space Agency (ESA), Japan, the National Oceanic and Atmospheric Administration (NOAA), and the Defense Meteorological Satellite Program (DMSP). Non-satellite data and basic research support will come from the complete set of U.S. Global Change Research Agencies and from national and international programs outside the U.S.

The strategy for EOSDIS V0 is to focus on the priority needs for reducing uncertainties in our understanding of global change, articulated by the Intergovernmental Panel on Climate Change (IPCC). These needs, generally consistent with the priorities of the U.S. Global Change Research Program, are divided into six areas [IPCC Working Group I, 1990; McBean and McCarthy, 1990]:

a. clouds: primarily cloud formation, dissipation, and radiative properties, which influence the response of the atmosphere to greenhouse forcing;

b. oceans: the exchange of energy between the ocean and the atmosphere, between the upper layers of the ocean and the deep ocean;

c. greenhouse gases: quantification of the uptake and release of greenhouse gases, their chemical reactions in the atmosphere, and how these might be influenced by climate change;

d. polar ice sheets: which affect predictions of sea level;

e. land surface hydrology: improved estimates of precipitation and evaporation and runoff into the oceans;

f. ecosystem processes: as sources and sinks of gases that have radiative properties in the atmosphere, and their role in exchange of moisture and energy between the land surface and the atmosphere.

II. DATA SETS

Progress over the next three years will be constrained by the data sets that exist or will be collected in this time frame. Although EOSDIS V0 must work with a wide variety of sources of data—space measurements, field and laboratory data, model results—NASA has particular responsibilities within the Global Change Research Program emphasizing satellite data. Thus, for EOSDIS V0, the set of satellites flown, flying, or to be launched in the coming years dictate the opportunities for progress in specific areas.

A. Clouds and Radiation

Clouds, aerosols, and greenhouse and sulfur gases all influence the transport of radiation into and out of the Earth's atmosphere. How each will respond to the build up of greenhouse gases in the atmosphere and other natural and anthropogenic changes will play a major role in determining the magnitude of global change and its regional variation in insolation and precipitation.

Radiative fluxes are sensitive to the amount, distribution, and optical properties of water and ice clouds and are central to the problem of greenhouse heating. Low clouds increase the Earth's reflectance while radiating in the thermal infrared at about the same temperature as the Earth's surface, thus cooling the Earth and offsetting warming caused by increased greenhouse gases. High clouds trap outgoing infrared radiation and amplify greenhouse warming. Increased aerosols reduce the net radiative flux at the surface, and net flux to the ocean determines the rate of ocean warming and consequent volume expansion.

EOSDIS V0 will contribute to our understanding of clouds in two major ways:

1. By archiving and distributing data from GOES and NOAA AVHRR, EOSDIS V0 will support GEWEX, which has as one of its objectives the more precise quantitative deduction of all energy fluxes within the atmosphere and at the air-sea and air-land interfaces.

2. Data from previous NOAA AVHRR instruments can be inter-calibrated to help determine whether the Earth’s albedo is changing.

B. Oceans

The thermal response time of the oceans is larger than that of the other components of the climate system. Thus the transport of heat by the oceans is thought to be a major determinant of the timing of greenhouse warming. Some climate models predict little warming of the Southern Hemisphere when thermal effects of the oceans are included. One key need is to determine the quantitative role of the oceans in the horizontal and vertical transport of heat in the Earth system.

In dealing with the Oceans area, V0 EOSDIS will support research to answer the question of horizontal transport of heat through the near-surface circulation of the world oceans. The key data sets will be new ocean surface topography measurements from TOPEX/Poseidon, near-surface wind fields from the ERS-1 AMI scatterometer, and sea surface temperature measurements from the ERS-1 Along-Tract Scanning Radiometer (ATSR).

The second question that is accessible to major progress in the V0 time frame is the quantification of the evaporative flux from the oceans. Key measurements to combine in this study are sea surface temperature and microwave measurements of moisture over the ocean. The needed satellite data are continuing to be collected by AVHRR and SSM/I instruments, and new capabilities to measure sea surface temperature will be obtained from the ERS-1 ATSR.

C. Greenhouse Gases

Greenhouse gases in the atmosphere, such as carbon dioxide and methane, are part of natural cycles. For these gases, anthropogenic rates of release are small percentages of the natural fluxes between the atmosphere, ocean, and ecosystem. Small adjustments in the ocean’s circulation and chemistry or in the life cycle of terrestrial vegetation could affect the stability of carbon dioxide and methane in the atmosphere.

During the V0 time frame, data from the SeaWiFS instrument will be made available to NASA. A follow-on to the Coastal Zone Color Scanner (CZCS) SeaWiFS will provide measurements of ocean productivity for such experiments as the international Joint Global Ocean Flux Study (JGOFS).

D. Polar Ice Sheets and Sea Ice

The polar regions are thought to be the main region of the ocean where there is substantial downward transport of water—bottom water formation. Through this process water which is in thermal and chemical equilibrium with today's atmosphere is transported to depths where the existing water was last in contact with the atmosphere before the industrial revolution. This offers a key buffer to the build up of surface heat and soluble greenhouse gases. One of the key uncertainties in predicting global change is whether some climate change could trigger a cut-off in this process.

This is an area where the V0 time can be one of major progress. The key question is to characterize the amount and distribution of bottom water formation and the related evaporative transport of heat and moisture to the atmosphere in the polar regions. The SSM/I instrument is currently providing wide area surveys of sea ice type, concentration, and extent. The ERS-1 and J-ERS-1 SARs will provide high resolution imagery of the sea ice on a frequent basis enabling determination of ice circulation as well as the concentration of leads and other features. Other sources of data are available to use in inferring the atmospheric surface temperature over polar seas and the calculation of surface winds over the ice. Together, these data sets should enable major progress in quantification of deep water formation during the V0 time frame.

Another aspect of polar processes is the growth and retreat of glaciers and ice sheets. The question is whether glaciers are reliable indicators of global warming on the 10 year time scale and whether they are currently growing or receding. During V0, data will be available to describe the surface proflies of glaciers and ice sheets in regions where their slopes are gentle—primarily the inland areas. Some scenarios for global warming predict that the interiors of ice sheets will initially grow due to increased precipitation resulting from increased evaporation from a warmer sea surface. Actual determination of whether glaciers are growing or receding will require characterization of changes at their mouths or edges. This will require ground surveys or instruments planned for the next century.

In the V0 time frame, altimeter data from historical and current satellites can be used to profile glaciers and ice sheets. These data will form the baseline for future studies.

E. Land Surface Hydrology

In moving through the Earth system, water plays a central role in many regulating physical, chemical, and biological processes. Evaporation, condensation, and runoff drive parts of the atmospheric and oceanic circulation and redistribute absorbed solar energy. Water vapor is the primary greenhouse gas. Water is the universal solvent and the medium for most biogeochemical cycling.

Despite its importance, present quantitative knowledge of the large-scale hydrologic cycle is poor. GEWEX addresses four scientific objectives:

- to determine water and energy fluxes by global measurements of observable atmospheric and surface properties;
- to model the hydrologic cycle and its effects on the atmosphere and ocean;
- to develop the ability to predict variations of global and regional hydrologic processes and water resources and their response to climate change;
- to foster development of observing techniques and data management and assimilation systems suitable for hydrologic and climatic prediction.

EOSDIS V0 will support studies of land hydrology by providing data from the SAR missions on ERS-1 and JERS-1, the passive microwave data from SSM/I, access to Landsat and SPOT catalogs, and access through NOAA to NEXRAD.

F. Ecosystem Processes

Changes in terrestrial ecosystems are a critical element in determining the regional impacts of global change, but in the near term what can be determined is the effect of current changes on the carbon budget of the atmosphere. The role of deforestation and regrowth in augmenting or buffering the fossil fuel driven increase in atmospheric carbon dioxide must be quantified. In addition, the sources of the increase in atmospheric methane must be determined both to enable prediction of future trends in this greenhouse gas and to enable any emission trading approach to regulating human emissions of these gases.

In the V0 time frame, a key answerable question is quantification of tropical deforestation on the carbon flux to the atmosphere. High spatial resolution land imagery collected by the MSS or TM instruments since 1972 offer the possibility of extending this determination back over as much as two decades where the early data can be recovered. The data can determine the area affected and the pace of deforestation; research using this data will have to interpret this information into net carbon fluxes.

The other key question to be answered is the role of boreal forests and their environs as a potential or actual source of methane and the current state of selected examples of this ecosystem. Here, the key data will arise from the combination of intense field measurements including the use of aircraft imaging spectrometers and synthetic aperture radars (SARs). These data will be complemented by AVHRR, TM, MSS, and SAR data from satellites.

III. DATA MANAGEMENT NEEDS FOR GEWEX

Attaining the GEWEX objectives will require a programme of extensive meteorological and hydrological observations including implementation of new remote sensing systems, and will include analysis and reprocessing of historical data. Assembly of historical databases for large areas is a major task and will be a primary focus of the Global Change International Project (GCIP). The GCIP data and information management system will be a significant test and milestone for the full GEWEX and

EOS-era data management implementations. The approach used must be flexible enough to accommodate all types of data needed for the study of energy flux and hydrologic processes at small through large scales and for evaluation of the water and energy balances of the global earth/atmosphere system.

In addition to historical data, the GCIP database will include special products derived from these data and a component that is continually being updated with new data in real time. This real time component may include data sets having limited historical record or historical data too expensive to retrieve over the entire study area. Examples include some of the satellite data, NEXRAD precipitation estimates, and data fields produced by data assimilation techniques to be developed as part of GEWEX. An organizational structure must be created to develop the historical database and to maintain the real time additions, including quality control functions.

The data and information management approach for GCIP must address the requirements, policies, and procedures for processing, archiving, documenting, and distributing data. Characteristics that will influence the development of the GCIP data system include the diversity of types of data and information to be acquired and managed, the capability to perform processing of special purpose data sets and data assimilations, the need for low-cost, timely, and open access to data, the geographical diversity of users, and the need for adequate quality control of the data. The data management approach will be designed around the requirements, sampling strategies, and instrument deployments specified by well-defined science objectives, and thus will focus on implementing a data delivery system that can meet the GCIP science and program goals. Emphasis will be placed on management of data in the 1993-1998 time frame. Due regard must be paid to data assimilation requirements, i.e. the need to prepare special merged data sets for model initial and boundary conditions, for archiving model output data, and for evaluating model results. Special links to centers such as ECMWF, NMC, and others involved in large-scale modeling and near real-time data assimilation will be required.

A. GCIP Database

Detailed data requirements will be developed from the GCIP science objectives. The database will include all information necessary to evaluate the water and energy balances of the atmosphere and the earth over the study area as a whole, and over its various parts. Included also will be data needed to study the behaviour of hydrologic processes at scales beginning with small watersheds, through the scale of atmospheric model grid squares, up to the full study area. The following categories of data will be included.

in situ: Radiosonde measurements and surface measurements of hydrological and meteorological parameters.

Surface-Based Radar: Precipitation measurements from the NWS NEXRAD system.

Aircraft: Visible, infrared, and microwave sensors for surface and atmospheric measurements.

Satellite: Atmospheric and surface measurements from geostationary and polar orbiters (GOES, NOAA, DMSP, ERS-1, TRMM, etc.)

Derived Fields/Model Output: Gridded fields obtained from analyses of satellite data (e.g. ISCCP products), interpolated satellite and in situ data for model initialization, and computed model output data.

Geographic Data: Including hydrologic boundaries, soils, land use, vegetation, and geologic data, stream locations, and digital terrain data.

B. Approach

The GCIP data management system (GCIP/DMS) will build on existing academic, government agency, and international data systems and activities. The unique science objectives of GCIP will require the provision of special GCIP/DMS services. These will include:

- production of special data sets such as blended fields of satellite and in situ data and reprocessed historical data using model data assimilation;
- quality control and documentation services to ensure the usefulness of GCIP data;
- means to utilize advanced electronic networks and standardized digital media for data access, data distribution, and information exchange;
- an information center to access project information/status, searchable data catalogs, bulletin boards, data ordering procedures, and on-line documentation;
- a focal point for development and implementation of a GCIP data policy covering issues such as data exchange, standards, data costs, data acquisition priorities, etc.

GCIP/DMS services will be provided by augmenting existing agency processing centers and archives. A central GCIP/DMS facility may be developed for those services that cannot be provided through existing facilities.

1. Data Centers and Archives

A substantial fraction of the historical data of interest to the GCIP resides in established data centers. In the U.S. these data centers include NOAA data centers (e.g. NCDC, NGDC, NSIDC), NASA data centers (e.g. NSSDC, PLDS, NCDS, NODS) the USGS EROS Data Center (and other geophysical databases), and other data centers and services operated or supported by, for example, the Departments of Energy and Agriculture. The National Center for Atmospheric Research maintains a data archive, and several universities and research institutions also maintain local archives of specialized data sets. Specialized data collections result from field experiments and projects such as FIFE, FIRE, HAPEX, STORM, TOGA, and others, most of which maintain their own short-term archives.

Large-scale modeling and data assimilation activities are carried out at centers with supercomputer facilities including some universities and meteorological centers such as the ECMWF, NMC, and UK Met Office. Smaller-scale modeling and process studies are carried out at a large number of distributed university departments and research institutions. These modeling activities drive the requirements for data and data integration, and provide output data that must be archived, validated, documented, and in some cases reprocessed in a continuing cycle. The documentation must contain information on the data calibration, algorithms and models used in the processing, and feedback from users on the quality and error characteristics of the data.

A number of satellite data reprocessing efforts are currently underway (or proposed) such as the NOAA/NASA EOS pathfinder data sets (CZCS, AVHRR, GOES, HIRS, and SSM/I). The NASA/Wetnet project is an example of a prototype EOS processing and distribution system for satellite meteorological data. Early planning for the EOS Data and Information System (EOSDIS) includes the development of Distributed Active Archive Centers (DAACs) which will build on existing elements within established NASA, NOAA, USGS, and NSF data centers (e.g. GSFC, JPL, MSFC, LaRC, EDC, NSIDC, ASF, and others). These (DAACs) will be foci for early testing and evaluation of EOS data processing algorithms and for generation of products and model output using existing satellite and in situ data.

Several international data centers also maintain data of interest to GCIP. In addition to the national meteorological data centers these include the Global Precipitation Data Center (GPDC) at Offenbach and the Global Runoff Data Center (GRDC) at Koblenz. Space agency data centers such as ESA/Earthnet, NASDA, and other national centers, and commercial providers such as EOSAT, SPOT, and others, will be the primary sources of satellite data.

It is a prime requirement that GCIP researchers have rapid, easy, and minimal-cost access to data archived by these centers.

2. Data Management and Quality Control

Historical data and new data acquired or generated for the GCIP must be carefully quality controlled and documented as part of the data management process. Scientists and data system managers must develop policies for quality control and documentation, and for maintenance of processing history and calibration information. Location of information and access to data in the archives will be aided by an on-line directory and data inventories to expedite data searches and to provide information on experiment progress, data availability, documentation, and references. This may be accomplished by utilization of an existing directory such as the Global Change Master Directory and development of a project information database and electronic network services.

3. Data Policy

The success of the GCIP will depend on scientists and agency participants sharing their data with each other. The timely archival of data collected or processed by investigators, and mechanisms to ensure open and minimal-cost distribution to all researchers requires a clearly stated and implementable data policy. The Interagency Working Group on Data Management for Global Change (IWGDMGC) has drafted a Data Policy for use by the Global Change Research Program. This policy is being reviewed by the Committee on Earth and Environmental Sciences (CEES) for adoption by U.S. Federal Agencies, and its elements are currently under discussion internationally as the basis for agreements related to acquisition and exchange of data. The policy should be supported and adopted by the GCIP as part of the global change effort. The WMO operates a well-established policy for free and open exchange of real-time meteorological data. However, many governments and agencies place restrictions on release of some other types of data for proprietary economic and political reasons. GCIP will provide a test of whether an open data policy can be successfully implemented internationally. The draft Data Policy with its current wording is included in Appendix B.

C. System Design

In developing the data and information management concept for GCIP the following issues will be addressed:

1. Centralized vs. Distributed

There are merits in having data sets locally archived and managed by investigators most closely involved with the data collection and utilization. This often leads to better quality control and maintenance of the data. It also integrates participating institutions and scientists more closely with the overall project. Disadvantages of this highly distributed approach are that it is often difficult to locate and obtain data sets held in diverse locations, there is little standardization in media, format, or content, and it is difficult for program management to ensure adherence to standard data policies (such as quality control, documentation, release of data, and timely distribution to other users). A compromise may involve a centralized management/directory system of services with a limited number of distributed processing and archival centers having responsibilities for specifc data and model types.

2. GCIP Data Center

The GCIP may fund the acquisition and archival of new data and generation of new analyses and model output products. Thus it is logical to consider housing some of the processing and archiving functions in a GCIP Data Center (or distributed system of centers). This would allow more focused management of the data including availability of in-house and visiting science staff to perform special processing and data assimilations, assist in validation and quality control, and to ensure that the data archives and inventories are maintained appropriately. A preferable alternative (from the viewpoint of cost and efficiency) may be to host GCIP data foci or projects within existing established data centers (such as NCDS, NCAR or NCDC, for example). This would have the long-term benefits that participating agencies would have an incentive to upgrade/expand their in-house facilities/capabilities and would be left with lasting improvements to benefit future programs.

3. Duplication of Data Holdings

GCIP investigators may acquire and integrate data sets from many sources. Rather than subsequently archive and maintain these data in a GCIP data center, thereby duplicating the original archive, it may be appropriate to archive these data separately only if there is a particular value-added aspect, e.g. an integrated data set for model assimilation or a more efficient time-space media storage scheme to facilitate rapid access to a particular data item.

4. Communications

Scientific research is enhanced by the ability to communicate rapidly and effectively. Electronic data networks should reach a point in their development during the next few years to enable GCIP investigators and management to exchange project information and status reports, query on-line directories and inventories remotely, and exchange data and model results through transparent network interconnections. The GCIP data system should take advantage of the developing national and international research network infrastructure to link its various components and integrate its management.

5. Other Activities

In developing a data management concept for the GCIP, note should be taken of ongoing data management activities of related projects and programs such as WOCE/TOGA, IGBP, UNEP/GRID, the World Data Centers, STORM, EOSDIS, and others. Data sets and data management infrastructure under development for these programs should be utilized by GCIP/DMS to the fullest extent possible.

D. Implementation Strategy

The implementation strategy for GCIP will likely follow those being developed for other recent programs and projects with similar objectives. Examples are the STORM Project and the Earth Observing System (EOS). Significant efforts by scientists and data managers alike has gone into development of the data management plans for these programs. In its data management plan, STORM emphasises the development of a distributed national meteorological data base using existing data centers. Each center would be linked into the existing national research computer network with automatic routing of requests to the appropriate data center. A single catalog/clearinghouse for all data of interest would be established. The EOS Data and Information System (EOSDIS) follows the same philosophy with its Distributed Active Archive Centers (DAACs) and Information Management Service (IMS). However the DAACs, located for the most part at existing data centers, will require major improvements in technology and capacity to cope with the large quantities of EOS-era satellite data. The DAACs include data centers of several U.S. agencies (e.g. NASA, NOAA, USGS, NSF), universities, and international organizations (ESA, NASDA), and will thus require international network links, standards, data policies, and cooperative agreements. Some of the DAACs will include or be affiliated with large-scale modeling and forecasting centers.

Given the critical nature of EOSDIS in support of the international GEWEX program, it would be highly advantageous to use GCIP as a prototyping activity for EOSDIS. EOSDIS has a goal of putting in place a fully functional data management system by 1994. This pre-launch system will demonstrate new technologies for archival and manipulation of large data sets. GCIP will provide a focus permitting an early evaluation of the EOSDIS capability for supporting a major scientific project and for interacting with a diverse, international scientific community. One of the EOSDIS DAACs would serve as the focal point for GCIP data management with access links to all other centers containing GCIP-related data.

Steps should begin as early as possible to build the GCIP data management requirements into the various agency program plans for EOSDIS in order that these may be implemented in a timely manner.

E. Data Policy Statement

The Interagency Working Group on Data Management for Global Change (IWGDMGC) has drafted a Data Policy statement in consonance with the goals of the U.S. Global Change Research Program [Committee on Earth and Environmental Sciences, 1991]. The draft policy statement is in the process of approval at the CEES level and is intended be used as a basis for establishing national and international policy. The policy statement is as follows:

1. The Global Change Research Program requires an early and continuing commitment to the establishment, maintenance, validation, description, accessibility, and distribution of high-quality, long-term data sets.

2. Full and open sharing of the full suite of global data sets for all global change researchers is a fundamental objective.

3. Preservation of all data needed for long-term global change research is required. For each and every global change data parameter there should be at least one explicitly designated archive. Procedures and criteria for setting priorities for data acquisition, retention, and purging should be developed by participating agencies, both nationally and internationally. A clearinghouse process should be established to prevent the purging and loss of important data sets.

4. Data archives must include easily accessible information about the data holdings, including quality assessments, supporting ancillary information, and guidance and aids for locating and obtaining the data.

5. National and international standards should be used to the greatest extent possible for media and for processing and communication of global data sets.

6. Data should be provided at the lowest possible cost to global change researchers in the interest of full and open access to data. This cost should, as a first principle, be no more than the cost of reproduction and distribution. Agencies should act to streamline administrative arrangements for exchanging data among researchers.

7. For those programs in which selected principal investigators have initial periods of exclusive data use, data should be made openly available as soon as they become widely useful. In each case the funding agency should explicitly define the duration of any exclusive use period.

IV. SERVICES NEEDED

Study of global change requires a data and information system that will provide the infrastructure to enable scientific interaction between researchers of different disciplines, and researchers employing different methods. It must be an **information** system, not just a **data** system, providing geophysical and biological information as well as raw data from space-borne instruments. It must allow researchers to collate and cross-correlate data sets and to access data about the Earth through searching, ordering, transmitting, analyzing, and visualizing.

Because scientific expertise is necessary to organize data and create geophysical and biological information, EOSDIS V0 is co-located at 7 Distributed Active Archive Centers (DAACs): Alaska SAR Facility, Jet Propulsion Laboratory, NASA Goddard Space Flight Center, NASA Langley Research Center, NASA Marshall Space Flight Center, National Snow and Ice Data Center, U.S.G.S. EROS Data Center. Affiliated Data Centers (ADCs) include: Consortium for International Earth Science Information Network (CIESIN) in Michigan, University of Wisconsin (GOES data), and NOAA. These centers all analyze and provide access to existing data used in global change research. Services required at the DAACs and ADCs are discussed in this section.

A. Information Management

Among the obstacles to using data from satellites for global change research, poor management of the data ranks near the top. A first, crucial step is to develop a way to search the inventories among the DAAC holdings, in a way that is transparent to the user signed on to one DAAC. Thus the data base management systems (DBMS) at each DAAC need to understand the same queries, so each must be implemented in a standard query language and each must understand the same data dictionary. Moreover, it is desirable to use a DBMS rather than the file system to manage collections of large objects.

Improvement in DBMSs is needed in the long run, although perhaps not in the V0 time frame [Stonebraker and Dozier, 1991]. Earth scientists require access to the following disparate kinds of data, which current standard query languages like SQL are inadequate to support:

- point data for specific locations;
- vector data, organized by polygons or lines;
- raster data, a regular array of point measurements typical of most remote sensing images;

- text data—programs, descriptions of data sets and results of simulations, technical papers, etc.

Current commercial relational data base systems (e.g. DB 2, RDB, ORACLE, INGRES, etc.) are not good at managing the these diverse kinds of data. Next-generation DBMSs (e.g. IRIS, ORION, POSTGRES, Starburst, Gemstone) have some of the attributes needed, but none are adequate for the full range of capabilities.

B. Visualization

Scientific visualization in the 1990s cannot be restricted to the resources available on a single scientist's workstation, or within a single processing system. The visualization environment of the future is one of heterogeneous machines on networks. Improved visualization is needed for interactive science with the data in EOSDIS. Visualization must be incorporated into the DBMS, so that the **data base** can be visualized, in addition to the **data sets** in it. Input through a ‘‘visual query language’’ will be needed.

C. Networks

Networking technologies must support data transfer rates needed for browsing and visualization. Not only is a high bandwidth required, but the network must guarantee delivery of required data without cumbersome pauses. A common scenario during V0 and beyond will be interactive visualization at a scientist's workstation of data from various DAACs. Bandwidth limitations must be overcome by staging of data in response to queries, and by image compression before transmission.

REFERENCES

Committee on Earth and Environmental Sciences, *Our Changing Planet: The FY 1992 U.S. Global Change Research Program,* 90 pp. Office of Science and Technology Policy, Washington, D.C. 1991.

Dozier, J. and H. K. Ramapriyan, Planning for the EOS Data and Information System (EOSDIS), in *Global Environmental Change,* edited by R. W. Corell and P. A. Anderson, NATO ASI Series I: Global Environmental Change Vol. 1, pp. 155-180, Springer-Verlag, Berlin, 1991.

IPCC Working Group I, Policymakers' summary, in *Climate Change: the IPCC Scientific Assessment,* edited by J. T. Houghton, G. J. Jenkins, and J. J. Ephraums, pp. vii-xxxiv, Cambridge University Press, Cambridge, 1990.

McBean, G. and J. McCarthy, Narrowing the uncertainties: a scientific action plan for improved prediction of global climate change, in *Climate Change: the IPCC Scientific Assessment,* edited by J. T. Houghton, G. J. Jenkins, and J. J. Ephraums, pp. 315-328, Cambridge University Press, Cambridge, 1990.

Stonebraker, M. and J. Dozier, Large capacity object servers to support global change research, *Sequoia 2000 Report No. 91/1,* University of California, Berkeley, 1991.