FINAL REPORT

ON

THE CREATION OF A NATIONAL SATELLITE LAND REMOTE SENSING DATA ARCHIVE

Prepared For

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL ENVIRONMENTAL SATELLITE, DATA, AND INFORMATION SERVICE
LANDSAT TRANSITION GROUP

Under Contract No. 50-DGNE-6-00149

JULY 15, 1987

EARTH SATELLITE CORPORATION *(EarthSat)*
7222 47th Street
Chevy Chase, Maryland 20815
(301) 951-0104
Telex: 248618 ESCO UR *Telecopier: (301) 951-4077*

FINAL REPORT

ON

THE CREATION OF A
NATIONAL SATELLITE LAND
REMOTE SENSING DATA ARCHIVE

Prepared for

NATIONAL OCEANIC & ATMOSPHERIC ADMINISTRATION
NATIONAL ENVIRONMENTAL SATELLITE, DATA, AND INFORMATION SERVICE
LANDSAT TRANSITION GROUP

Under Contract No. 50-DGNE-6-00149

July 15, 1987

EARTH SATELLITE CORPORATION *(EarthSat)*
7222 47th Street
Chevy Chase, Maryland 20815
(301) 951-0104
Telex: 248618 ESCO UR *Telecopier: (301) 951-4077*

Final Report

on

The Creation of a National Satellite Land Remote Sensing Data Archive

by

Charles Sheffield
Orville R. Russell
James M. Durana
Nancy G. Anderson
Michael C. Place
Roger A. Mitchell
Donald Garofalo

July 15, 1987

ACKNOWLEDGEMENTS

Numerous groups and individuals contributed their time and expertise to assist this project. Special contributions were made by, and special thanks are due to: Peggy Harwood, David Clark, David Hastings, David McGinnis and Howard Warriner, and the staff of NOAA/NESDIS; Nancy Milton, Larry Pettinger and the staff of the U.S. Geological Survey (USGS) at Reston, Virginia; Al Watkins, Bryan Bailey, and the staff of the USGS EROS Data Center (EDC), who provided both invaluable statistical information and a thorough understanding of EDC operations; Bill Cunliffe, Charles Dollar, Ethel Hedlin, Bill Cline, Alan Calmes and the staff of the National Archives and Records Administration; Jim Riordan and the staff of the National Space Science Data Center, Goddard Space Flight Center; Peter Norris and the staff of the Earth Observation Satellite Company; Philippe Delclaux and the staff of SPOT Image Corporation, Reston, Virginia; and Gail Thirlwall-Wilbee, of the Canada Centre for Remote Sensing.

EXECUTIVE SUMMARY

The Land Remote Sensing Commercialization Act of 1984 calls for the Secretary of Commerce, and, through delegation, NOAA, to provide for a U.S. archive of data from U.S. and foreign earth resources satellites, to serve historical, scientific, and technical purposes, including global environmental monitoring. This new archive, named the National Satellite Land Remote Sensing Data Archive (NSLRSDA), is planned to begin operations in 1989. However, the law left the following factors to be determined:

(1) the data to be archived in the NSLRSDA, together with criteria that decide which data should be collected, and in what order of priority;

(2) the satellite remote sensing systems which should provide data to the NSLRSDA;

(3) the location(s) of the archive (subsequently agreed to be the USGS EROS Data Center);

(4) the method of storing and accessing the data; and

(5) the geographic or other frame needed to reference the data, both for acquisition and subsequent recall.

This report addresses the above issues and provides recommendations on each, together with a general methodology for future archive planning and data acquisition by the NSLRSDA. It documents the results of a seven-month analysis performed by Earth Satellite Corporation, with assistance from the National Oceanic and Atmospheric Administration, the U.S. Geological Survey, and the National Archives and Records Administration, and with substantial inputs from several hundred potential users of the archive.

The report contains the following elements:

(1) A review of spaceborne sensors and imaging systems of possible value for land remote sensing.

(2) A review of land remote sensing data bases currently in existence, in this country and abroad.

(3) A survey of archival media, and of the state of the technology of both film and machine-readable storage and retrieval methods.

(4) A review of possible long-term data uses and disciplinary requirements for NSLRSDA archival data, established through interviews and meetings with data users.

(5) The definition of a geographic reference frame for data collection, together with an organizational framework and a systematic procedure for deciding data acquisition priorities in the NSLRSDA.

<u>CONCLUSIONS</u>: The conclusions drawn from these analyses are as follows:

(1) Of the more than two hundred spaceborne imaging systems of possible interest for land remote sensing, thirty-eight are judged of potential value to the NSLRSDA. Final rankings are made among them into three levels:

Level 1: Sensors that can provide data to the NSLRSDA immediately. Initial archival interest should be concentrated on these. They are: the Landsat Thematic Mapper, the Landsat Multispectral Scanner, the Landsat Return Beam Vidicon (no longer functioning), the SPOT High Resolution Visible sensor, the Advanced Very High Resolution Radiometer, and the MOS Multispectral Electronic Self Scanning Radiometer.

Level 2: Sensors not operating today, but which will provide valuable data to the NSLRSDA once they are functioning (8 sensors are in this group).

Level 3: Sensors that are still purely conceptual, or undefined as to design specifics or data availability (24 sensors are in this group).

(2) Significant holdings of valuable data exist and will continue to exist in foreign data banks; however, procedures to ensure the orderly transfer of such holdings, present and future, to meet the needs of the NSLRSDA have not yet been established.

(3) There is great interest on the part of potential users in the NSLRSDA, but also widespread concern that the value of the enterprise will be reduced unless the appropriate data are selected, collected, and stored. Scientific users are seen as an essential part of that continuing selection process.

(4) The same user community feels that the full future uses of archival land satellite data are not yet well understood. The user community is concerned lest the NSLRSDA become too rigid in its structure to accommodate changing priorities.

(5) Although the NSLRSDA is a new enterprise, there already exists in the United States considerable experience in the building and maintenance of data bases for land remote sensing satellite data. This experience should be capitalized upon as much as possible.

(6) At the same time, cost limitations will make it infeasible for the NSLRSDA to encompass or substitute for all remotely sensed satellite existing data bases in this country and abroad; nor will it be possible to impose on those data bases the sampling and copying requirements of a U.S. national archive of permanently valuable data. Thus the continued existence of these uncontrolled data bases must be assumed, while it must further be assumed that their contents may have considerable value for land remote sensing research applications in this country.

(7) Remotely sensed satellite data are increasingly used in digital form. This allows them to be transformed easily to a variety of map projections, and used in conjunction with data bases of other types of information.

(8) The same increased use of digital forms of data makes accurate calibration and processing history of data of increasing importance.

(9) The key technological variable for digitally stored data is not storage medium lifetime; it is the rapid progress in and consequent rapid obsolescence of read-and-write storage technology. Cost per unit data stored is decreasing fast, at the same time as media storage densities increase.

(10) The data volumes under consideration are very large (many trillions of bits of information), and data acquisition strategy, referencing, and retrieval represent a large logistics, data management, and computational problem.

RECOMMENDATIONS: The above conclusions lead to the following recommendations:

(1) The NSLRSDA needs to inventory the holdings of other data bases, particularly those abroad, and to establish, as soon as possible, bilateral agreements with other organizations and nations for the transfer to the NSLRSDA of selected necessary data.

(2) To assure the acquisition of the most useful data by the NSLRSDA, and to give it suitable flexibility of structure, an Archives Steering Group should be established, for the overall guidance of the NSLRSDA. Under this Steering Group an Archive Data Selection Committee (ADSC) should be formed to define data acquisition strategy. The ADSC should comprise principally scientists and data users.

(3) In order to make most efficient use of existing U.S. experience in establishing and maintaining land remote sensing satellite data bases, the NSLRSDA should be established at the EROS Data Center, in Sioux Falls, South Dakota, home of the current Landsat and other space-derived image data bases. Added processing and storage facilities should be provided at that location to permit the needs of the NSLRSDA to be met.

(4) The NSLRSDA should maintain catalogs of other data bases, in this country and abroad; however, it should limit its own scene holdings to the Basic Data Set as defined by the ADSC's data acquisition decisions.

(5) The primary storage medium of the NSLRSDA must be in machine-readable digital form. There must also be stored with the data sufficient ancillary information (documentation of processing performed, calibration testing, evaluation studies, and other significant quantitative descriptions) to permit users to make the best quantitative uses of the data.

(6) Storage in machine-readable form can and should begin at once in the NSLRSDA, recognizing that data copying to other storage media will be inevitable to comply with National Archives' copying requirements. It is expected that such copying will also be economically attractive as storage costs per bit of machine-readable data continue to decline.

(7) Delivery of data from the NSLRSDA should be provided in a few fixed machine-readable and film product formats. The task of integrating these archival data with other data bases, or of otherwise reconfiguring data, will remain with the final users.

(8) A computerized global reference system should be created in the near future. It must permit the efficient generation of data acquisition needs and priorities, by location and by sensor. Control of this system and associated reference data base, and assignment of priorities for entries in it, should be the responsibility of the ADSC.

TABLE OF CONTENTS

FIGURES

TABLES

1.0 INTRODUCTION AND STATEMENT OF OBJECTIVES

Public Law 98-365, Section 602, states the following:

> "(a) It is in the public interest for the United States Government--
>
> (1) to maintain an archive of land remote-sensing data for historical, scientific, and technical purposes, including long-term global environmental monitoring;
>
> (2) to control the content and scope of the archive; and
>
> (3) to assure the quality, integrity, and continuity of the archive.
>
> (b) The Secretary shall provide for long-term storage, maintenance, and upgrading of a basic, global, land remote-sensing data set (hereinafter referred to as the 'basic data set') and shall follow reasonable archival practices to assure proper storage and preservation of the basic data set and timely access for parties requesting data. The basic data set which the Secretary assembles in the Government archive shall remain distinct from any inventory of data which a system operator may maintain for sales and for other purposes."

The law leaves to be separately determined which satellite remote sensing systems will contribute to the "basic data set" for archival maintenance; it also leaves for future definition the method of acquisition for such a data set, and the methods to be used to store or access data, within the guidelines established by law.

This report describes work performed by Earth Satellite Corporation, under contract to NOAA, to recommend for the National Satellite Land Remote Sensing Data Archive (NSLRSDA) (1) data needs, (2) a possible logic for construction of a Basic Data Set, (3) potential and preferred sources of data, (4) suggested archival methods, and (5) possible initial archival acquisitions. Since the archive is explicitly stated to apply to long-term environmental monitoring, the current contract is perceived as part of an ongoing effort by NOAA, which will in the future take account of advances in storage technology and the availability of new satellite data sources.

The objectives of the present contract were as follows:

(1) By direct contact with potential users of archival data, to establish a profile of archival data requirements for land satellite data, recognizing that this can be no more than an initial estimate of needs.

(2) To review spaceborne systems and sensors which are candidates to contribute to the data archive; past, present and future space imaging systems are all to be considered.

(3) To provide a logical hierarchy of data sources, in terms of their relevance to a national satellite land remote sensing data archive, and in terms of data availability, data cost, and perceived assessment by users of the value of each data source.

(4) To examine current data bases of land remotely sensed data, including data types, data volumes, data formats and data storage methods.

(5) To examine present and projected performance of large volume storage devices appropriate to the archiving of remotely sensed satellite data.

(6) To examine facility requirements for such an archive, and review capabilities of existing candidate facilities for maintenance of the national satellite land remote sensing data archive at its designated location.

(7) To identify problem areas in today's collection or storage procedures, and suggest action items to be undertaken by NOAA prior to final establishment of the land satellite data archive.

(8) To make recommendations to NOAA on data types to be included in the national archive, on the relationship of this archive to other data sources, on facility requirements for the national archive, and on the definition of the initial Basic Data Set that the archive should contain.

A summary of objectives, conclusions, and recommendations is given in the separate volume of the Executive Summary. Subsidiary supporting information is provided in the form of separately bound appendices to this report.

2.0 BACKGROUND AND GENERAL DISCUSSION

2.1 History of the U.S. Land Remote Sensing Program

The land satellite remote sensing program of the United States began in 1965 as NASA's Earth Resources Survey Program. It led in July of 1972 to the launch of the first Earth Resources Technology Satellite, ERTS-1 (renamed Landsat-1 in 1975).

Although significant potential benefits were identified early for the Earth Resources Survey Program, it was conceived of, and operated as, an experimental and scientific activity. Thus, for example, no archival procedures following National Archives' requirements for historical data were made; nor was there funding either for the processing to computer compatible tape of all data or for the long-term storage of all received data.

In spite of its experimental status, the Landsat program proved to be highly successful and visible. Other nations installed ground receiving stations for the reception of Landsat data, disciplinary scientists began to use derived information regularly, industrial groups employed Landsat data in survey and exploration programs, and numerous state and local government groups found a place for satellite data in on-going operations.

In fact, despite its experimental status, the Landsat program was increasingly used as though it were an operational program, with assured data continuity and availability.

In 1980, recognizing the difference between Landsat's original status as an experimental program and the data's day-to-day uses, the U.S. Government moved to change the nature of the program; first, to move it to operational status; and second, to perform the transition of the Landsat program to private operation. As originally defined (in Presidential Directive-54, October 1980), the program would be operated within the Commerce Department until 1988, when the transition from public to private ownership and operation would begin. There was still no explicit statement.of the fact that Landsat data would be of permanent national value, and there were no funds for systematic archive development, complete conversion of old data to computer-compatible tape (CCT) form, or periodic inspection and maintenance of stored data.

In 1983, the U.S. Government made the decision to speed up the transfer of the Landsat program from government to industry. Instead of waiting for 1988, in the Fall of 1983 the U.S. Commerce Department issued a Request for Proposal inviting industrial groups to bid on the takeover of the Landsat system. The Land Remote Sensing Commercialization Act (Public Law 98-365, July 17, 1984) specifically called for the provision of a two-satellite system and for the continued operation of Landsat 4 and 5, already in orbit. Although the Request for Proposal did not specify permanent data archiving responsibility for the successful bidder, the Commercialization Act did call explicitly for the definition of a Basic Data Set and the creation of a National Satellite Land Remote Sensing Data Archive to house such a data set. The archive was conceived as a government responsibility, with the industrial owner-operator of the system providing data to the archive, but not maintaining it. The archive would not be restricted to Landsat data alone, but would include as necessary other data sources required for scientific, long-term, global monitoring.

Seven bidders responded to the initial Request for Proposal. Following preliminary evaluation, three bidders were invited to make oral presentations. Following these, two bidders were left. Upon a reduction of proposed government funding for the program, the Eastman Kodak consortium dropped out of the bidding, and a contract was signed on September 27, 1985 with the remaining group, the Earth Observation Satellite Company (EOSAT, a joint venture of Hughes and RCA). EOSAT took over operations of Landsat 4 and 5 in October 1985, and began construction of Landsat 6 and 7.

Since EOSAT lacks data processing, storage and reproduction facilities, customers of Landsat data continue to be served by the U.S. Geological Survey's EROS Data Center in Sioux Falls, South Dakota, who complete the data orders received by EOSAT.

2.2 Data for the National Satellite Land Remote Sensing Data Archive

To provide a clear set of definitions when discussing creation of the NSLRSDA, the following terminology is adopted in this report:

(a) Data bases consist of any collection of satellite remotely sensed image data, in machine-readable or other form, useful in land applications. These data bases may be stored in this country or abroad, in a variety of formats. Such data bases are candidates to provide data for the NSLRSDA, but do not form part of it. Their contents are not subject to the legal requirements for data sampling and copying established by the National Archives for permanently valuable data (see Section 2.3 and Section 6.0).

(b) The Basic Data Set will consist of a defined set of past or future satellite image data, designed to accomplish the stated objectives of the public law as defined in Section 1.0. Existing data defined as part of the Basic Data Set will become part of the NSLRSDA upon physical transfer to national archive storage. This transfer should be done as soon as possible after formal initiation of a national archive of land satellite remotely sensed data. Future Basic Data Set acquisitions enter the NSLRSDA following processing, review, and physical transfer to the archive.

Although the EROS Data Center holds far more Landsat data than any other repository, EDC does not have all data acquired by Landsat. Landsat 1, 2 and 3 carried equipment that permitted both on-board tape recording of images and direct image transmission to ground receiving stations. Landsat 4 and 5 carried equipment that permitted images to be transmitted via the Tracking and Data Relay Satellite System (TDRSS), or directly to ground receiving stations. Landsat 4 was launched in July 1982, and Landsat 5 in March 1984. The first TDRS was launched in April 1983. The second TDRS was on board the Space Shuttle, Challenger, destroyed in January 1986. Therefore, there has been transmission to the United States through TDRS of only the western hemisphere (TDRS-1 is stationed at 41 degrees West).

Many data from Landsat 1, 2 and 3 were received only by foreign ground receiving stations, and the unavailability of TDRS has made this even more relevant for data from Landsat 4 and 5. Thus there exists today no complete repository of Landsat data, here or abroad; there are instead partial data bases in numerous locations. This applies not only to old data, collected 10 or more years ago, but to data being collected today. The Landsat scene collection statistics for 1986 demonstrate the need to consider foreign data bases as potential sources for NSLRSDA (Table 1).

Table 1. Scenes collected by Landsat 4 & 5, 1986
(Source: EOSAT, personal communication, 1986)

	U.S. Reception	Rest of World Reception
MSS Data	158,000 scenes	942,000 scenes
TM Data	118,000 scenes	263,000 scenes

In addition to Landsat, there are other significant sources of satellite land remotely sensed data. These must also be candidates for the Basic Data Set, both as important data sources in their own right, and as alternatives to Landsat where the latter source is unavailable. SPOT and AVHRR data, plus others on systems recently launched (e.g., the Japanese MOS) or soon to be launched, may also contribute to the required Basic Data Set, and therefore be part of the NSLRSDA.

2.3 Project Assumptions and Constraints

The present project was undertaken, beginning in October 1986, with the expectation that the predominant contributor to the NSLRSDA would be Landsat data, past, present and future, as implied in Public Law 98-365. Other data sources were originally perceived as supplementary. During performance of the contract, the status of the Landsat program changed radically. At this writing, there is no funding to build Landsat 6 or Landsat 7. Unless new negotiations are successfully concluded between EOSAT and the U.S. Government, there will be no near-term successor to Landsat 5. If negotiations are successfully completed, and construction begins again on Landsat 6 and 7, the data gap in provision of Landsat data will be at least two years and more likely 3 or 4 years. Therefore, in performing this project, it has been necessary during the last 3 months to rethink the whole question of data sources for the NSLRSDA. Since the imaging systems of alternative sources differ substantially from Landsat, data calibration for comparison of old and new data also takes on an increased importance.

During the course of this project, a communication from the National Archives to NOAA (Appendix 1) also clarified the position of the NSLRSDA vis-a-vis general National Archives' policy. All machine-readable data selected for the Basic Data Set and transferred to the NSLRSDA will be subject to the National Archives' sampling and copying requirements. These requirements are themselves based upon certain experience and assumptions concerning the reliable lifetime of machine-readable storage media. These assumptions and lifetimes are reviewed in Section 6.0. It should be noted that some of the existing Landsat repositories in machine-readable form, at Goddard Space Flight Center and elsewhere, are already approaching the suggested National Archives' limit of such lifetimes, for the earliest data they contain (1972). Thus the early definition and transfer to the NSLRSDA of the older portions of the Basic Data Set from all sources, domestic and foreign, take on special urgency.

Although the definition and creation of a long-term global archive appears at first sight to be a well defined and stable problem, this is not the case. The uncertainties in future satellite collection systems on the one hand, and costs and procedures for storage on the other hand, impose constraints on the possible contents of the archives. Acquisition procedures must be reviewed regularly through the lifetime of the NSLRSDA, storage technology must be monitored, the availability of data reading devices must be assured, and the rationale for basic data set collection must be regularly scrutinized. This archival definition must be viewed as an on-going, dynamic process. The way in which these considerations were taken into account during performance of the project is contained within Section 9.0.

3.0 METHODOLOGY

Since the 1930's, many hundreds of organizations and individuals have been engaged in the use and evaluation of remotely sensed data. Although in recent years some have employed mainly aircraft data and others satellite data, it is fair to say that almost all the emphasis has been placed on near-term applications and scientific analyses. Many workers have not thought in terms of the value or uses of data many years old, nor have they thought in terms of archival data needs.

At the same time, many groups have been concerned with the creation and maintenance of permanent data repositories. However, few if any of those groups have considered the problem of archiving remotely sensed data for long-term historical and scientific purposes; nor have many groups given extensive thought to the long-term archival needs of machine-readable remotely sensed data.

Although there exists much experience in both remote sensing and data archiving, there has been little overlap of these two fields. The key methodological problem of archive definition is to bring together the knowledge of these two distinct disciplines, and use the results to guide the development of the NSLRSDA.

At first sight, two optional courses present themselves: One can seek to bring an archival perspective to the remote sensing community; or one can seek to make specialists in data archives knowledgeable about remote sensing programs. Some combination of these two methods appears preferable to either alone. The methodological procedure employed on this project was as follows:

(1) Through literature review and general surveys, candidate data collection systems were identified and evaluated for the period 1987 to 2005 (see Section 4.0) and existing data bases were reviewed (see Section 5.0).

(2) An overview of storage technology was performed, again using literature reviews and survey methods. Meetings were also held with archiving specialists to determine applicable technological and legal constraints (see Section 6.0).

(3) Using this information, a telephone survey was conducted of satellite remote sensing data users. This was done employing a questionnaire specifically designed to focus respondents' thinking on archival data uses, as opposed to current data uses. It was recognized from the outset that a survey of this type, asking respondents to think of satellite data in a way that was unfamiliar to many of them, was not likely to produce the deepest considerations. Respondents were therefore urged to provide subsequent conclusions or suggestions as they occurred to them (see Section 7.0).

(4) With this information available on data collection systems and suggested archival data uses, a public meeting was held. Participants at this meeting included representatives from both the satellite remote sensing and data archiving communities. The public meeting served to review survey results, and then to provide a

second iteration of the issues by a group including individuals oriented to the traditional problems of archiving (see Section 8.0).

(5) The results of the literature surveys, telephone surveys and public meeting were used to define the strategy appropriate to selecting the initial Basic Data Set, and to develop recommendations for continued data acquisition and Basic Data Set addition. A methodology permitting a flexible approach to different competing data requirements was developed, and is presented in Section 9.0, together with the Basic Data Set.

4.0 SENSORS AND DATA SOURCES

4.1 Introduction and Task Description

A study of potential data sources for the NSLRSDA was undertaken with the following objectives:

- To review spaceborne systems and sensors which are candidates to contribute a permanent data archive; past, present and future space imaging systems were all to be considered.
- To provide a logical hierarchy of data sources, in terms of their relevance to a permanent data archive, and in terms of data availability, data cost, and perceived assessment by users of the value of each data source.

In order to address these objectives, all spaceborne systems that might provide historical data to the archive were reviewed, evaluated and ranked. For sensors not yet flown, this review and evaluation was necessarily speculative and often limited, based on incomplete information. The process started with the review of sensor characteristic information, data format, and volume information on imaging sensors, as well as information on the satellites on which each is, was, or will be flown (see Appendix 2 for a rather complete listing of spaceborne imaging sensors). The sensor list of over 200 sources was narrowed to 65, based on criteria described below, to limit the evaluation only to those sources that might provide data to the archive. The evaluation process was designed and carried out with the purpose of selecting those sensors that are reasonable candidates for contributing land remote sensing data to the archive's Basic Data Set.

Review and evaluation of potential data sources in late 1986, and the subsequent discussions at the February 3, 1987 public meeting, identified three other issues relevant to various data sources and their possible contribution to the NSLRSDA.

(a) The Land Remote Sensing Commercialization Act of 1984, while not explicitly stating the fact, indicates that the Landsat system will be the initial and primary data source for the archive. As of early 1987, this system's future is in doubt. The archive therefore should have contingencies for pursuing its charter without the benefit of post Landsat-5 data. If the system's future becomes more certain, then a review and consideration of past, present and particularly future sensors should be made in terms of their compatibility with Landsat spatial and spectral resolution.

(b) An interrelated issue to the Landsat system as a basic data source is the role of foreign owned and/or operated systems. The Land Remote Sensing Commercialization Act of 1984, which mandates the legal authority to the archive, specifically discusses the inclusion of "...data collected by foreign ground stations or by foreign remote-sensing space systems...."

(c) Any agreement entered into by the NSLRSA, whether with a U.S. or foreign system operator, will serve as a precedent for future agreements. If a fully commercial space remote sensing industry evolves, beyond Landsat 6 and 7 and SPOT 1 and 2*, then the NSLRSDA will become simply a customer to that industry, albeit a particularly influential one. Such a scenario should be considered in the Basic Data Set acquisition strategy.

Foreign sensor data (which must include both the information collected by instruments owned and operated by the countries, and data bases from U.S. sensors maintained outside the United States and not duplicated in domestic data bases) must be evaluated as to their potential contribution to the archive. The evaluation should include and be sensitive to:

- o The prospects of data continuity with policy changes in foreign national governments;
- o The prospects of continued data quality;
- o The likelihood that the data will prove to have scientific value over the long-term; and
- o Guarantees of data availability.

Once any sensor data are acquired by the archive, it should be assumed that the commitment is long-term (greater than 10 years). If not then the value to the archive to pursue part of its charter, i.e., long-term environmental monitoring, is questionable.

4.2 Review and Evaluation Task

First a review of past, present, and future land imaging sensors was conducted, and an evaluation made of each data type for inclusion in the archive. An information search was undertaken to compile a list of

*proposed sensor system

sensors and their associated imaging characteristics, data descriptions, and satellite system characteristics. A data sheet was developed for each sensor (see Figure 4.1, an example sensor data sheet).

4.3 Review and Evaluation Criteria

More than 65 sensor systems were identified from an initial list of roughly 200 candidates by applying the following ranking criteria developed in October-December, 1986, and finalized after the February 3, 1987, public meeting.

1. The sensor (present, proposed or past) must be imaging from a position in space. This was an attempt to separate aircraft sensors from consideration, keeping within the guidelines of Section 101 of the Land Remote Sensing Commercialization Act of 1984 "...the continuous civilian collection and utilization of land remote sensing data from space..."

2. No solely military sensors would be considered. Reference again to Section 101 of the Land Remote Sensing Commercialization Act of 1984 "...the continuous civilian collection...."

3. Soviet sensors will not be considered at this time. Despite the bilateral agreement between the United States and the Soviet Union on cooperation in space, little information is available from the Soviet Union concerning its space based land imaging systems, and US/USSR transfer agreements have been restricted to weather data.

4. Each system selected should have a two kilometer or less resolution field of view (RFOV). This was conceived as a means of separating those sensors that are primarily land remote sensing from those that are meteorologically oriented. This separation does not suggest that an integrated approach with both types of data is not useful or desirable. The restriction was accepted only after acknowledging the practical parameters of cost and size to which the archive will have to adhere. In this regard the NSLRSDA should strive to build institutional flexibility, continually evaluating various sensors as a source for the Basic Data Set.

FIGURE 4.1

Example Sensor Data Sheet

SENSOR
SATELLITE SERIES
COUNTRY
ORGANIZATION
CONTRACTOR
LAUNCH DATE
TERMINAL DATE
PROPOSED OPERATIONAL LIFE
ORBITAL CHARACTERISTICS
- o Synchronization
- o Altitude
- o Inclination
- o Node
- o Repeat Cycle

IMAGING CHARACTERISTICS
- o Type
- o No. of Channels/Frequencies
- o Spectral Range/Freq. Range
- o Resolution (RFOV)
- o Swath Width

MISSION OBJECTIVES

COMMENTS

4.4 Evaluation, Results, and Recommendations

A tri-level sensor hierarchy was applied in order to group the sensors for priority selection as data contributors to the NSLRSDA:

Level I	-	Sensors with the ability to provide data to the NSLRSDA by 1987. (Highest priority)
Level II	-	Sensors not yet operational (by 1987), but upon operation likely to provide data of value to the NSLRSDA .
Level III	-	Sensors that are questionable as to operational date and/or design specifics. This level includes those sensors that are wholly conceptual. (Lowest priority)

A fourth level was recognized but perceived as being outside the conceptual framework, i.e., those data acquired or to be acquired irregularly and with no operational acquisition strategy. Examples include Mercury, Gemini and Apollo mission photography, Large Format Camera photography, and many of those sensors that were launched for research and development or with narrow mission objectives. How these types of sensor data might fit into an acquisition strategy to fulfill the NSLRSDA's mandate needs to be addressed. Selected data from experimental systems, although irregularly acquired, are also of possible value to the NSLRSDA.

The tri-level hierarchy was developed in order to categorize the potential land imaging sensors into subjective groups, based on their predicted relative importance to the archive's Basic Data Set. Data continuity and value to long-term global environmental monitoring were overriding considerations in selecting sensors for inclusion in the tri-level hierarchy.

The results of the sensor evaluation are contained in the following table for Levels I, II and III:

Level I	-	Landsat Return Beam Vidicon (RBV) Camera Landsat Multispectral Scanner (MSS) Landsat Thematic Mapper (TM) SPOT High Resolution Visible (HRV) Advanced Very High Resolution Radiometer (AVHRR) MOS Multispectral Electronic Self Scanning Radiometer (MESSR) Coastal Zone Color Scanner (CZCS) Heat Capacity Mapping Radiometer (HCMR) Seasat
Level II	-	Enhanced Thematic Mapper (ETM) AEROS-A 4-Band Linear Array Space Shuttle Imaging Spectrometer Experiment (SISEX) Radarsat Synthetic Aperture Radar (R-SAR) J-ERS Synthetic Aperture Radar (SAR) J-ERS Visible Near Infrared Radiometer
Level III	-	Landsat-7 Multispectral Linear Array Radarsat R-AVHRR Radarsat R-MOMS STS-7, STS-11, SPAS-01 Modular Opto-Electronic Multispectral Scanner (MOMS) A-EOS Large Microwave Radiometer (LMR) A-EOS Synthetic Aperture Radar (SAR) A-EOS Multispectral Linear Array A-EOS Mid Infrared Imager (MIRI) China Land Satellite Geosynchronous Orbit High Resolution Earth Monitoring Spacelab ESA Optical Imaging Instrument EOS Moderate Resolution Imaging Spectrometer FRG/India Multispectral Electro/Optical Stereo Scanner EOS High Resolution Imaging Spectrometer Chinasat BRESSEX BHASKARA-1 BHASKARA-2 Advanced ESA IRS (Indian) Linear Imaging Self Scan Sensor-1 (LISS I) IRS (Indian) Linear Imaging Self Scan Sensor-2 (LISS II) Space Shuttle Thermal Infrared Multispectral Scanner (STIMS) Tropical Earth Resources Satellite Sensor (TERS)

These categories provide the logical basis for ranking sensors as data sources for Basic Data Set acquisitions. Data sheets on each sensor in Levels I, II and III are given in Appendix 3.

5.0 EXISTING RELEVANT DATA BASES

5.1 Discussion of Data Bases

Collection of land remotely sensed satellite data has been pursued for many years, long before the decision was made to establish the NSLRSDA. A substantial number of land satellite remotely sensed data bases exist, each of them a potential contributor to this permanent archive.

An assessment of the nature, formats, volumes and condition of these data bases was necessary in this project, to establish to what extent, if any, these previously collected data should become part of the Basic Data Set of the NSLRSDA.

The review concentrated on data available from sensors included in Level I of Section 4.3 (except MOS, data from which are not yet available). These consist of:

- o Landsat RBV, MSS and TM;
- o AVHRR; and
- o SPOT HRV

In addition, non-recurring photographic and multispectral data currently stored at the EROS Data Center, while not primary candidates for the Basic Data Set, were reviewed because they are valuable historic data, unavailable from any other imaging source.

The rapid increase of data volumes with increased spatial and spectral resolution is revealed in this review of current data volumes. Noting the trend toward increased resolution, the NSLRSDA must be prepared to accommodate future sensor data having large volumes (see Section 6.0 for a discussion of appropriate large volume storage devices).

5.2 Non-Recurring Photographic, Radar and Multispectral Data

Photographs of the earth were acquired during the early U.S. manned satellite programs. The Gemini missions which orbited the earth in 1965 and 1966 provided hand-held, 70mm camera normal color photography obtained through spacecraft windows. The later Apollo earth-orbiting test missions provided the first multispectral photographs, whose use in arid and forest terrain interpretation helped demonstrate the value of multispectral satellite images.

Skylab, which orbited in 1973 and 1974 providing coverage between 50 degrees North and 50 degrees South, was the first manned mission emphasizing systematic Earth observation. Skylab carried the Earth Resources Experiment Package (EREP), consisting of a multispectral camera (S-190A Experiment), the Earth Terrain Camera (S-190B Experiment), and the S-192 13-channel scanner.

The Space Shuttle has also flown several remote sensing systems. Among these are the Shuttle Imaging Radar (SIR-A and B) Experiments and the Modular Optoelectronic Multispectral Scanner (MOMS). Photography has also been acquired by hand-held 70mm and 140mm cameras, and by the Large Format Camera.

Some of these photographic, radar, and multispectral data from Gemini, Apollo, Skylab and Shuttle missions are currently held at the EROS Data Center; other major repositories are at NASA and the European Space Agency. The extent of the EDC holdings is summarized in Table 5.1, and catalogs showing locations, cloud cover, etc., are maintained at EDC. Data from Seasat are held primarily in NOAA's archives.

Table 5.1: Summary of EDC Early Photographic and Multispectral Data Holdings from Manned Missions
(Source: EROS Data Center)

PLATFORM	DIGITAL SCENES	FILM ROLLS/FRAMES	INDEXING SCHEME
Skylab II, III, IV	1,800	634/44,845	38,765 SCENES IN MIF 36 MICROFICHE INDEXES 31 ROLLS OF BROWSE MICROFILM
Apollo/Gemini	—	127/18,362	478 SCENES IN THE MIF REMAINING COVERAGE IN HARDCOPY CATALOGS AND ON 7 ROLLS OF MICROFILM
Shuttle	—	381/41,025	IBM/PC DATA BASE 29 ROLLS OF BROWSE MICROFILM
TOTAL	1,800	1,142/104,232	

5.3 Landsat Data

The Landsat data base, the U.S. holdings of which are primarily maintained at the EROS Data Center, is continually expanding. International in scope and diverse in data type, this large data base contains a variety of data formats. In order to comprehend these formats, ground processing systems are next discussed.

5.3.1 Ground Processing Systems

Four U.S. ground processing systems have been used to process Landsat data into forms suitable for archiving and image product generation.

Landsat 1-3 RBV and MSS data were originally processed through NASA's Data Processing Facility (NDPF), the original ground system, and output mainly onto 70mm black and white film masters.

Approximately 6,000 of these early MSS scenes were also converted to Computer-Compatible Tape (CCT) form at user request, and stored at EDC. In 1979, when NASA implemented the next generation ground system, NASA also decided to continue generating CCT's for selected 1972-1979 data as a data preservation measure, because of difficulties encountered in converting wide-band video tapes to CCT's as these tapes and the original processing system aged. Approximately 45,000 1972-1979 scenes were systematically selected by EDC as worthy of historical preservation and conversion to CCT formats. Thirty-two thousand of these scenes had been converted by NASA as of March, 1987.

The second ground processing system, the Image Processing Facility (IPF) at the Goddard Space Flight Center (GSFC) and the EDC Digital Image Processing System (EDIPS) began operation on February 1, 1979. Landsat MSS data conversion through this system began in February and Landsat-3 RBV data conversion began in September 1980. High-Density Tapes (HDT) created through the IPF were transmitted to EDC, where EDIPS CCT generation and film processing were performed.

The third and fourth ground systems, the MSS Image Processing System (MIPS) and the TM Image Processing System (TIPS) began operation in July, 1982 with the launch of Landsat-4, as part of the Mission Management Facility (MMF), the current processing system at GSFC. Through this system, MSS HDT data are transmitted via DOMSAT to EDC where EDIPS CCT generation and film processing are performed. TM data are generated at GSFC on black-and-white 241mm film and shipped to EDC in rolls for archiving and product generation; TM CCT's are created at GSFC and shipped to EDC for copying and storage only upon user request.

Table 5.2 summarizes the March 1986 format of the data processed through these four systems. In addition to these U.S. systems, each foreign ground receiving station has its own processing system, data storage method, and CCT format. Like the U.S. systems, the foreign processing systems have changed over the years.

5.3.2 EDC Data Holdings

The EROS Data Center, as the current and primary U.S. Landsat data base, holds almost 800,000 digital scenes as of March 1987, and the number is continually increasing.

Table 5.3 shows the quantity of RBV, MSS, and TM holdings broken down in terms of spacecraft, digital scenes and frames. The total number of scenes broken down by spacecraft and data type is summarized in Table 5.4.

Table 5.2: Archive Medium by Satellite, Date and Sensor
(Source: Landsat Data User Notes Notes (NOAA)
Issue No. 35, March 1986)

Satellite	Date	Sensor	Medium and Location
Landsat 1	Jul 72-Jan 78	MSS	70-mm Film with Selected X-format CCT's at EDC
		RBV	70-mm Film at EDC
Landsat 2	Jan 75-Dec 78	MSS	70-mm Film with Selected X-format CCT's at EDC
		RBV	70-mm Film at EDC
Landsat 2	Jan 79-Jul 83	MSS	HDT-P or A at EDC*
Landsat 3	Mar 78-Dec 78	MSS	70-mm Film with Selected X-format CCT's at EDC
Landsat 3	Mar 78-Aug 80	RBV	70-mm Film at EDC
Landsat 3	Jan 79-Sep 83	MSS	HDT-P or A at EDC*
Landsat 3	Sep 80-Sep 83	RBV	HDT-P or A at EDC*
Landsat 4	Jul 82-Present	MSS	HDT-A at EDC
Landsat 4	Jul 82-Feb 83**	TM	HDT-R at GSFC with Selected 241-mm Film and Customer Requested CCT's at EDC
Landsat 5	Mar 84-Present	MSS	HDT-A at EDC
Landsat 5	Mar 84-Present	TM	HDT-R or A at GSFC with*** Selected 241-mm Film and Customer Requested CCT's at EDC

*HDT-P stands for fully "processed" data with radiometric and geometric correction applied. HDT-A stands for partially processed data, without geometric resampling applied. HDT-R stands for "raw" data.

**The combination of the X-Band transmitter failures and the subsequent solar array power cable failures effectively ended the acquisition of Landsat 4 TM data on February 15, 1983.

***Landsat 5 HDT-R tapes were reused from May 6 to July 27, 1985.

The quantity and data format of Landsat MSS and TM data held at EDC is summarized in Tables 5.5 and 5.6, respectively. Figure 5.1 compares the relative storage requirements of one MSS and one TM scene, illustrating the increase in data volume with increased spatial and spectral resolution.

Table 5.3: Landsat Holdings at EDC
(Source: EROS Data Center, December 1986)

PLATFORM	DIGITAL SCENES	FILM ROLLS/FRAMES	INDEXING SCHEME
Landsat 1-MSS	26,900	--/150,500	ALL SCENES IN THE
Landsat 1-RBV	—	—/1,380	MIF
Landsat 2-MSS	13,000	—/184,700	768,025 SCENES
Landsat 2-RBV	—	--/1,985	16,980 ROLLS
Landsat 3-MSS	86,110	—/230,260	2,433,800 FRAMES
Landsat 3-RBV	150,530	—	SUPPORTED BY
Landsat 4-MSS	40,900	—/38,400	MICROFICHE CATALOGS
Landsat 4-TM	333	—/1,260	550 ROLLS OF MICROFILM
Landsat 5-MSS	85,900	--/82,000	1,360 MICROIMAGE
Landsat 5-TM	3,270		MICROFICHE
TOTAL	407,000	16,978/2,433,754	

Table 5.4: Accession/Sensor Breakdown (Scenes)
(Source: EROS Data Center, December 1986)

	MSS	RBV	TM	TOTAL SCENES
Landsat 1	144,477	1,380	—	145,857
Landsat 2	183,129	1,985	—	185,114
Landsat 3	86,110	150,532	—	236,642
Landsat 4	38,302	—	1,256	39,558
Landsat 5	81,823	--	14,105	95,928
TOTALS	533,841	153,897	15,361*	703,099

*AN ADDITIONAL 50,684 SCENES ARE CATALOGED BUT ARCHIVED BY GSFC

Table 5.5: MSS Holdings at EDC
(Source: EROS Data Center, December 1986)

LANDSAT ARCHIVE SUMMARY

	LS1	LS2	LS3	LS4	LS5	TOTAL
MSS B&W CHIPS	-	41,032	117,528*	38,302	81,823	278,685
MSS B&W ROLLS	150,516	143,698	112,730*	-	-	406,944
MSS ENHANCED B&W'S	219	291	98	211	0	819
MSS STD CCP'S	5,262	6,981	2,700	449	852	16,244
MSS NON-STD CCP's	215	276	33	3	0	527
MSS HDT-P	-	32,525	79,355*	-	-	111,880
MSS HDT-A	-	14,081	61,556*	40,821	85,645	202,103
MSS CCT-X	26,899	13,000	1,250	-	-	41,149

*INCLUDES RBV DATA

CCP = Color Composite Photograph
STD = Standard
HDT-A = High Density Tape; Partially processed data without geometric resampling applied
HDT-P = High Density Tape; Fully processed data
CCT-X = X-format Computer Compatible Tapes; the original NASA processed format

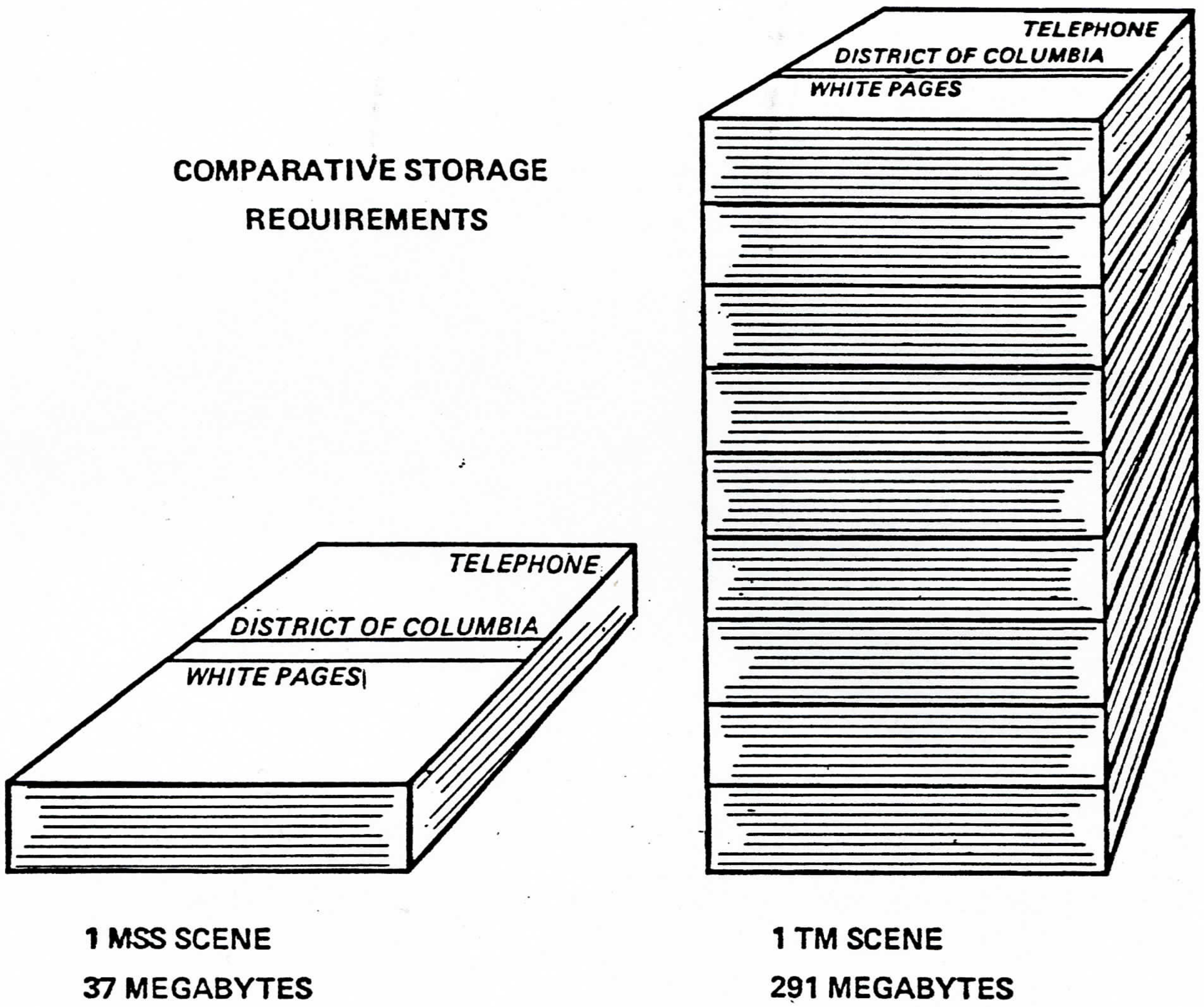
Table 5.6: TM Holdings at EDC
(Source: EROS Data Center, December 1986)

LANDSAT ARCHIVE SUMMARY

	LS1	LS2	LS3	LS4	LS5	TOTAL
TM B&W CHIPS	---	---	---	1,256	14,105	15,361
TM ENHANCED B&W'S	---	---	---	4	3	7
TM STD CCP'S	---	---	---	46	577	623
TM NON-STD CCP'S	---	---	---	47	243	290
TM CORRECTED CCT'S	---	---	---	293	3,055	3,348
TM UNCORRECTED CCT'S	---	---	---	40	215	225
TM-HDT-A REFERENCES	---	---		1,707	64,338	66,045

CCP = Color Composite Photograph
HDT-A = High Density Tape; partially processed data without geometric sampling applied

Figure 5.1: Relative MSS and TM Scene Storage Requirement



5.4 AVHRR Data

The NOAA-8, NOAA-9, and NOAA-10 satellites* carrying Advanced Very High Resolution Radiometer (AVHRR) sensors routinely record and transmit 4-km resolution Global Area Coverage (GAC) data to two stations; one located in Wallops Island, Virginia, and the other in Gilmore Creek, Alaska. The 1 km Limited Area Coverage (LAC) data are collected and transmitted similarly, but only at user request. Some AVHRR 1 km data are transmitted directly to Wallops Island, Gilmore Creek and proposed foreign locations using High Resolution Picture Transmission (HRPT). All data are transmitted via communications satellite to NOAA-Satellite Data Service Division (SDSD) in Suitland, Maryland, for processing. All post-June 1986 AVHRR data are currently archived in Suitland on 160 mb IBM 3480 cartridges (approximately three orbits of GAC data are stored on one cartridge). Some previously acquired data are stored on magnetic tape.

The current volume and growth rate of AVHRR data is shown in Table 5.7. SDSD, as of January 1987, had 3964 GAC tapes of data stored in its active archive library, acquired at the rate of 5 tapes/day; 9154 tapes of LAC data are also currently stored in active archive. 11,694 tapes of GAC data and 371 tapes of LAC data are currently stored in SDSD's inactive archive.

NOAA-SDSD currently does not have a procedure for purging data; all data transmitted are archived. Methods are currently being explored for storing data in deeper archive; 60% of AVHRR data ordered is requested within three days of acquisition.

*TIROS-N, NOAA-6 and NOAA-7, currently inactive, also carried AVHRR sensors.

Table 5.7: Digital Data--NCDC/Satellite Data Services Division (SDSD)

SOURCE/DATA STREAM	DESCRIPTION/PARAMETERS	STATUS	RECEIPT FREQ/MODE		CURRENT VOLUME, GB	GROWTH GB/YR	1998 PROJ. VOLUME, GB
POES AVHRR GAC Level 1b	IR Radiance & Visible Albedo (4km)	A-90	D	H	1,890.0000	459.9000	4,189.5000
POES AVHRR HRPT/LAC L-1b	IR Radiance & Visible Albedo (1km)	A-90	D	H	900.0000	547.5000	3,637.5000
POES K-L-M AVHRR GAC Level 1b	IR Radiance & Visible Albedo (4km)	91-93	-	-	0.0000	459.9000	1,379.7000
POES K-L-M AVHRR HRPT/LAC L-1b	IR Radiance & Visible Albedo (1km)	91-93	-	-	0.0000	547.5000	1,642.5000
Polar Plat. AVHRR GAC L-1b	IR Radiance & Visible Albedo (4km)	94-98	-	-	0.0000	459.9000	2,299.5000
Polar Plat.AVHRR HRPT/LAC L-1b	IR Radiance & Visible Albedo (1km)	94-98	-	-	0.0000	547.5000	2,737.5000

A USGS and NOAA program designed to receive, process and archive AVHRR data, optimized for land earth science applications, begins on May 15, 1987. Data are to be acquired for the coterminous United States at the EROS Data Center. The summary of the operation is shown in Table 5.8, and the circle of coverage in Figure 5.2. However, there are no plans to release these data initially to the general public; rather, data will serve Federal agency research needs, with the possibility of more general distribution if public interest warrants it.

Table 5.8: EDC AVHRR Data Reception and Processing System

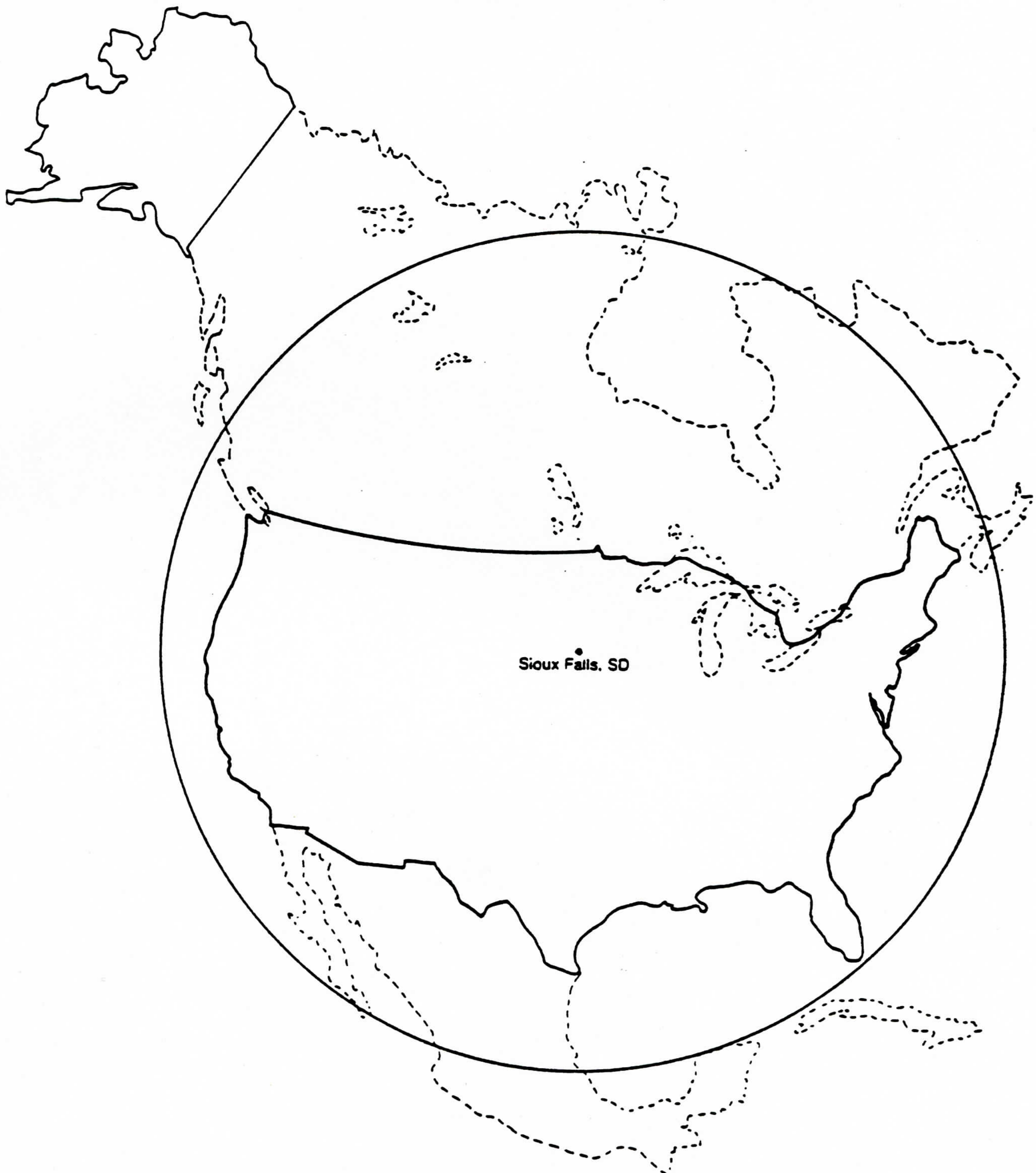
- Joint USGS/NOAA program to receive, process, and archive AVHRR data optimized for land earth science applications
- System Requirements:
  - Acquire 10-bit full resolution data for coterminous U.S. for Federal research purposes
  - Screen data for image quality and cloud cover
  - Extract data subsets
  - Provide radiometrically calibrated and geographically registered data
  - Merge AVHRR data with ancillary map data
  - Provide processed data on CCT, floppy disk, and film
  - Provide processed or raw digital data within 24 hours after data acquisition or product request

5.5 SPOT Data

The French SPOT satellite was launched in February, 1986. Although more than 200,000 scenes had been acquired by SPOT's High Resolution Visible (HRV) sensors, as of February 1987, no formal plans for

FIGURE 5.2

EDC Reception Coverage Circle for AVHRR Data with 5° Antenna Elevation



archiving for long-term use have been announced, nor are the SPOT plans subject to any U.S. National Archives' policies. The following comments therefore reflect no more than current SPOT practices.

Of the scenes so far collected (February, 1987), approximately 48,000 are of "acceptable quality and cloud cover" (a subjective evaluation by SPOT Image Corporation). Only 5,000 scenes have so far been processed to CCT format, and the CCT processing line at Toulouse, France, is only now coming to full production capability. The primary storage medium for SPOT data remains High Density Tape (HDT), with the main archive at Toulouse. Currently, every scene acquired is retained in the data base.

North American image data are received at Gatineau (Quebec) and Prince Albert (Saskatchewan) and recorded there on HDT's. These HDT's are sent to the SPOT Image Corporation facility in Reston, Virginia, for processing to CCT and image products. About 3,000 U.S. scenes with 25% or less cloud cover were stored at Reston by December 1986. SPOT Image Corporation reports that their U.S. sales for the first quarter of 1987 have already exceeded their U.S. sales for the whole of 1986.

SPOT receiving stations, in addition to those in Canada and at Toulouse, are located in Sweden, China, India, Pakistan, Japan and Saudi Arabia. Stations are also planned in Brazil, Australia, South Africa and the Canary Islands.

5.6 Contribution of Existing Landsat Data to the National Satellite Land Remote Sensing Data Archive

It is clear from the foregoing discussion that much of the earlier existing data bases are stored in analog (film) form, and original radiometric values for such data have been lost. At the same time, the early data often constitute a unique historical record of surface conditions as they were in the 60's and early 70's. In their historical content, their value exceeds the value of recent data of comparable quality, since recent data can normally be replaced by new acquisitions. Early historical data should therefore be considered as strong candidates for inclusion in the NSLRSDA. At the same time, it is probably not worth attempting to copy such data, or to seek their conversion to a machine-readable format compatible with newer data sources.

Fortunately, film storage is highly stable, with a long useful lifetime when stored under proper conditions (see Section 6.1). However, there also exist substantial quantities of early digital data. Most of these data are still held on wide band video tapes, and is not being carefully preserved or inspected. In particular, much of the early MSS data collected by Landsat 1, 2, and 3 have never been processed to CCT format, and EOSAT now questions whether many of the wide band video tapes are in good enough condition to be processed. At the very least, the NSLRSDA should perform a systematic evaluation of the early Landsat digital data, with a view to determining how many scenes should be considered candidates for inclusion in the Basic Data Set. A preliminary screening of this kind was performed several years ago, and a set of scenes (about 45,000) selected as worthy of conversion to CCT's for their historical value and quality. This effort was partially completed before it ran out of funds in March 1987, leaving about 13,000 selected scenes still unprocessed to CCT format. At a minimum, that list should be carefully reviewed as potential inputs to the Basic Data Set.

6.0 SUPPORTING TECHNOLOGY

Globally, the bulk of land remotely sensed data available to the public is currently in film format. This might seem appropriate to the casual user, since the primary impression of land remotely sensed data is that of images. However, the serious user, and in particular the scientific user, values these remotely sensed data not only for their appearance, but also for the quantitative information they contain about the Earth's surface.

As the recognition of the quantitative nature of satellite data has increased, there has been a steady shift from qualitative picture analysis to quantitative information extraction. Consequently, many users have moved to machine-readable media such as magnetic tapes and disks.

The stability and lifetime of machine-readable storage has been less well-explored than the properties of film storage. In the next sections, following a brief summary of film storage and retrieval properties, the state-of-the-art of machine-readable storage media is reviewed. The move to machine-readable forms introduces a second important question; namely, the availability and cost of digital data retrieval and analysis systems. These questions are also addressed in the following sections.

6.1 Film Storage and Retrieval

Although a photographic image, particularly one produced on a film base, contains an immense amount of information, today's technology to store and retrieve quantitative information from film is much more limited than for digitally stored images. As a consequence, photographic products are arguably of secondary importance to the NSLRSDA. Nonetheless, existing photographic films and paper will probably constitute a sizable portion of the NSLRSDA's initial holdings. Consequently, an understanding of their archival properties is important.

Black-and-white photographic products consist of photosensitive chemical bonded with a gelatin-type material to a substrate, either paper or polyester film. The character of each of these components can affect the archival nature of the photographic product. The light-sensitive material consists primarily of silver chloride or silver bromide suspended in the bonding agent. These two materials are collectively known as the emulsion. The major film substrate used today is polyethylene terephthalate, usually referred to as polyester, which provides a stable, flexible, and durable base. Photographic prints are composed of similar emulsions, but the substrate is paper. The paper is normally of high quality and acid-free, and when properly processed, has excellent archival properties, with lifetimes of at least a century.

It is difficult to make reliable long-term estimates of the longevity of a material based on short-term, accelerated-aging tests. However, tests on polyester film base-material by Brown, Lowry, and Smith (1984) and Adelstein and McCrea (1981) gave, for properly stored film, a minimum archival life of 1,000 years, with a probable life exceeding 2,000 years.

Determining the lifespan of the image on the film is even more difficult, and depends in part on the stringency of the criteria applied as to what constitutes acceptable limits of deterioration. It is believed, however, that images will be of usable quality after several centuries.

The archival permanence of black-and-white photographic films is influenced by the storage environment. Temperatures should not exceed 68 degrees Fahrenheit (20 degrees Centigrade) for any extended period of time, although short exposures to light in properly constructed enlargers or printers does not appear harmful. Relative humidity (RH) should be maintained between 30 and 40 percent in both storage and when printing. RH values of 60 percent or more invites the invasion of fungus and mold. Contaminants in the atmosphere should be monitored and controlled. Particulate matter should be filtered out of the air and harmful gases removed or neutralized. Harmful gases include nitrogen dioxide, peroxides, ammonia, sulfur dioxide and ozone. Although all of these gases are common in our environment, the technology exists to maintain them at acceptable levels in an archival facility. These problems are discussed in more detail by The National Research Council (1986) and Baer and Banks (1985).

Color films are much more complex than black-and-white films and consist of multiple layers of different emulsion materials designed to perform various functions in addition to those of retaining the dyes that produce the color image. The long-term stability of the various layers in the emulsion is not certain, but the most limiting factor in the longevity of a color image is the permanency of the organic dyes used to produce it.

Eastman Kodak (1985) recommends for color films and negatives that cut film should be stored in a special paper envelope. However, one should not store black-and-white and color materials in the same envelope. Due to the inert character of metal, roll film is best stored in metal containers rather than plastic or wood that may contain volatile substances that could adversely affect the film. Metal canisters should be sealed with a moisture proof tape or other sealing material.

Eastman Kodak (1985) recommends for color film storage a temperature of 0 degrees Fahrenheit (-18 degrees Centigrade) and RH between 30 and 35 percent. RH values of 25% or less can cause brittleness and RH values of 60% encourages the growth of fungus.

Even with these stringent controls, the longevity of color films is not guaranteed. Kodak (1971) recommends for permanent preservation of a color photograph the making of three color-separation negatives on black-and-white film.

6.2 Machine-Readable Archive Media

The advent of electronics and computers in the last 30 years has radically changed the way that remotely sensed images and data are transmitted, recorded, and analyzed. The acquisition of remotely sensed images, particularly from space platforms, requires electronic communications technology which was initially analog in nature, but for the past 10-15 years has been almost exclusively digital. Electronically recorded remotely sensed data in turn spawned the development of image processing of remotely sensed data and geographic information systems, permitting far more useful information to be extracted than is possible with human-readable data such as film. The need for machine-readable data will be sustaining and continue long into the future. However, this change introduces its own problems. While national archives often target a 50-to-

100 year horizon, the rapidly evolving computer industry which developed the machines and media upon which the remotely sensed data are currently stored have an evolutionary cycle decidedly shorter, of the order of 3 to 5 years. This conflict poses significant issues for a permanent archive of machine-readable remotely sensed information.

This section is divided into 6 parts: General Archival Functions, Sensor Source Media, Archive Formats and Subsidiary Data, Machine-Readable Archival Media and Procedures, Machine-Readable Distribution Media, and Findings and Recommendations. The section on Source Media discusses machine-readable media and related equipment issues which the NOAA archive may have to accept. Archive Formats discusses calibration and processing level of stored data. Archival Media and Procedures addresses the long-term preservation issues of machine-readable media. Distribution Media discusses problems and issues related to making copies of the archival media long into the future. The final section attempts to summarize the findings and provide some recommendations.

6.2.1 General Archival Functions

It is presumed that the NSLRSDA will have the responsibility of distributing archival data by making copies onto selected distribution media for the future user. To meet this objective, the archive operation must be able to perform the following basic technical functions, in addition to the normal cataloging functions:

(1) Read the machine-readable remotely sensed data provided by the sources.

(2) Copy the data to an archival medium for long-term storage, and/or to a commonly accepted distribution medium for dissemination.

To perform these basic functions, the NSLRSDA must acquire, operate and maintain the following key components:

(1) Input devices compatible with the source media (i.e., tape drives, etc.)

(2) Software compatible with the data formats utilized on the source media

(3) Computer hardware to support (1) and (2) above

(4) Documentation on the operation of the input devices, media formats and software.

(5) Calibration information on source data.

By and large, today's sensor operating lifetimes in space are the same 3 to 8 year life cycles commonly found in computer hardware. As sensors change, new input devices and software must be acquired to maintain access to the data. As source media input devices become obsolete, they may be retired if the data have been transferred to another archive medium, or the replacement reading devices are already in place. Regardless, the archive much be prepared to maintain compatible input devices for the archive horizon life. An archive horizon life of 50 years poses significant problems for maintaining complex input devices and software. At the same time, converting <u>all</u> machine-readable source media to a common archival medium would significantly raise operation costs.

At issue here is a basic conceptual decision for the NSLRSDA: Should machine-readable source data be converted to a common machine-readable archival medium, or should a diversity of archival media be permitted?

In either case, the NOAA remote sensing archive must maintain a capability of reading the source media to inspect new materials to ensure future readability and proper cataloguing. It is therefore important to review the spectrum of machine-readable source media as is currently being distributed by remote sensing data libraries.

Note that existing data bases complicate the issue further. Regardless of any common machine-readable form agreed to for new data acquisition, old data bases, consisting of hundreds of thousands of scenes, already exist in a variety of machine-readable forms.

6.2.2 Sensor Source Media

Section 4.0 of this report identified the primary sensors, data from which are recommended for inclusion in the NSLRSDA. Table 6.1, Sensor Source Media, identifies these sensors and the media upon which they are currently being distributed. These data were derived by telephone contact with EDC, NOAA, SDSD, and NSSDC. Clearly, today's preferred machine-readable physical medium is 1/2 inch 9-track magnetic tape, although most of these data are also available as photographic products. This is not surprising as this medium has been the industry standard for 10 years as a means of disk back-up. Most sensors are available on the higher density 6250 BPI for those users with more powerful tape drives. Notably missing is the low-density 800 BPI medium and any 7-track tape, common ten years ago. In the past 5 years, these older tape drives have generally been phased out

Table 6.1: SENSOR SOURCE MEDIA

Satellite	Sensor	Physical Medium	Interface	Density	Code	Data Format	Levels	Sq. Scene Dimension	Square Scene Size Mbytes	# of 1600 BPI Tapes/ Scene	# of 6250 BPI Tapes/ Scene	Comments
Landsat-1	MSS	1/2 FeO tape	9 trk	1600/6250	ANSI	X-format	P	185 km^2	40	2(?)	1	
Landsat-2	MSS	" " "	" "	"	ANSI	X-format/BIP2	P	"	40	2	1	
Landsat-3	MSS	" " "	" "	"	ANSI	EDIPS	A and P	"	40	2	1	
Landsat-4	MSS	" " "	" "	"	ANSI	EDIPS	A and P	185 km^2	40	2	1	
Landsat-5	MSS	" " "	" "	"	ANSI	EDIPS	A and P	"	40	2	1	
Landsat-3	RBV	" " "	" "	"	ANSI	RBV	P	"	40	2	1	
Landsat-4	TM	" " "	" "	"	A	LTWG TM Quad-US	P	"	300	12	3	
Landsat-5	TM	" " "	" "	"		LTWG TM Quad-US	P	"	300	12	3	
TIROS-N	AVHRR-LAC	" " "	" "	"	ANSI	AVHRR-format	1B	3000 km^2	40	1	1	1 km resolution
NOAA-6	AVHRR-LAC	" " "	" "	"		AVHRR-format	1B	"	40	1	1	
NOAA-7	AVHRR/2-LAC	" " "	" "	"		" "	1B	"	40	1	1	
NOAA							day/night register pair		40			
HCMM	HCMR	" " "	" "	1600	ANSI	HCMR	day or night	700 km^2	9	1	1	
NIMBUS-7	CZ65	" " "	" "	1600/6250	ANSI	CRT-tape format	1	1623km x 800km	13	1	1	
SPOT-1	HRV					LTWG-SPOT	1	60 km^2	70	2	1	includes pan 10m band
MOS	MESSR											
NOAA-8	AVHRR-LAC	" " "	" "	"		AVHRR	1B	3000 km^2	40	1	1	
NOAA-9	AVHRR-LAC	" " "	" "			"	1B	"	40	1	1	
NOAA-10	AVHRR-LAC	" " "	" "			"	1B	"	40	1	1	

Table 6.1: SENSOR SOURCE MEDIA (Continued)

Satellite	Sensor	Physical Medium	Interface	Density	Code	Data Format	Levels	Sq. Scene Dimension	Square Scene Size Mbytes	# of 1600 BPI Tapes/ Scene	# of 6250 BPI Tapes/ Scene	Comments
Seasat	SAR	" " "	" "	1600/6250		(Also Photo Products)		100km x 100km	38	1	3	
STS 2	SIR-A	5" film				Film Strip		50km x variable				
STS 41G	SIR-B	1/2 FeO tape	" "	6250		(Also Photo Products)		100km x 50km				

of operation, obviating the need to distribute products in these formats, although data in these formats still exist, and must be readable.

The fact that these source media are released in 1/2-inch 9-track media does not mean the original data were gathered or are currently stored on 9-track media. For example, a large portion of the current Landsat inventory is still maintained on 2-inch High Density Tapes (HDT's) in an unprocessed form.

As higher performance media become popular, the data producers are likely to offer source data on these new media. This evolutionary process is largely driven by the general computer hardware market. The evolutionary cycle has been in the range of 5 to 10 years before a new medium has become an established standard for distribution purposes.

The NSLRSDA should plan on acquisition of new peripheral devices to keep pace with accepted standards of the data producers. Currently, this standard is 1/2-inch 9-track 1600 BPI. The evolving standard is 6250 BPI. However, the steps involved in receiving source data for a machine-readable archive goes well beyond the simple capability of reading the media concerned. For example, the U.S. National Archives is currently maintaining an archive of over 9,000 1/2-inch tapes relating to such records as census, military personnel, etc. (interviews with personnel at the Special Archives Branch, National Archives and Records Administration). The following procedure is used in archiving these data tapes:

(1) Make an appraisal decision to retain certain records.

(2) Receive and read the magnetic tape to validate that information is intact.

(3) Using the documentation, create a program that can read the individual records and data fields to ensure the data is accurate.

(4) Create a master copy which is to be stored off-site.

(5) Create a back-up copy from which distribution copies are made.

(6) Identify, mark, index and otherwise catalog the type of information, tape volume and location of each tape.

Considering that the information is being archived for permanent storage, these validation steps seem reasonable. Areas of difficulty include finding current documentation, and developing programs to validate the fields. While not directly applicable to remotely sensed data, the validation aspects of receiving machine-readable sensor data would be similar. For example, instead of developing programs to sample the textual and numerical fields of Census data, the NSLRSDA may acquire programs to read the sensor data and produce a photographic image. Admittedly, this later procedure may be expensive and possibly redundant, because many of the data producers and users are themselves well equipped to produce photo images. A decision to retain in the NSLRSDA any data sent to it would be based on defined data assessment criteria (e.g., cloud cover).

6.2.3 Archive Formats and Subsidiary Data

Locating and archiving adequate documentation concerning sensor data is a possibly formidable task. It is not sufficient to identify the file and field formats necessary for reading the tape. Most sensor-based information is stored in a raster format, in which each pixel on each spectral band has a value from 0 to 128, or 0 to 255, etc. However, these values are relative only to the calibration parameters of that particular sensor. For example, a value of 75 on the visible green band of AVHRR may equate to a value of 132 on Landsat MSS. These relationships are not absolute. Subtler but equally important variations can be caused by the different electro-optic sensing mechanisms, and by "drift" in sensitivity of a single sensor over time. Documenting the radiation sensing properties of the various sensors under consideration has resulted in hundreds of research studies, reports and books. For researchers in future generations to make meaningful use of these data, they must have access to this extensive collection of calibration information keyed to the vintage of individual data sets in the archive.

It is strongly recommended that the NSLRSDA should preserve, in close proximity to the image data, documentation on sensor file and field formats. In addition, the NSLRSDA should preserve, at minimum, all relevant research studies relating to sensor calibration. Without calibration, the sensed images become no more than qualitative views of the Earth's surface. The image library of the NSLRSDA <u>must</u> be supported by an adequate calibration library.

On Table 6.1, Sensor Source Media, the standard processing levels are identified in code form. These levels, whether A or P for Landsat, and 1 or 1B for other sensors, indicate to the informed user the amount of ground processing used to correct for sensor anomalies and geometric distortion. The lower level (A-format) Landsat data has only ephemeris and scanline corrections applied to the individual pixels; the higher level (P-format) Landsat data has geometric corrections applied to create a Universal Transverse Mercator (UTM) map projection. Processing to a higher level causes some (hopefully small) loss of radiometric integrity. Landsat TM is almost always distributed in P-level form, because the A-level data are so difficult to use that only a handful of users have the software capability to process to a usable geometric format. However, future researchers, trying to make meaningful use of P-level data, may need to know the precise mapping algorithms applied to the raw data to produce the final format. This information is part of the ground system processing documentation and software. At issue is whether ground post-processing documentation and/or software should be preserved, and whether such documentation is in fact available for all existing data. In any event, for each data type to be archived, careful research of processing effects must be conducted to decide the best processing level to be archived, recognizing that some users prefer minimally-processed data. Thorough documentation of the algorithms and software used to reach that level will then have to be preserved as well.

The National Archives creates software to read the file and field formats of incoming source tapes, and to validate their contents. This is not a complex programming task, but the variety of different source files used by the National Archives increases the workload substantially. Duplicating this task for the NSLRSDA is simpler, because there should initially be fewer than a dozen different data formats, even including foreign data sources. Software drivers to read these formats are available in the commercial marketplace and may be in the public domain.

6.2.4 Machine-Readable Archival Media and Procedures

The usual perception of "archiving" valuable information is of preserving the original document. While original preservation may be important for historical and symbolic reasons, most documents in an archive are important only for the information they contain. Transferring this information to a more accessible, higher density, or longer lasting medium is a regular practice at all archive institutions. In recent decades, microfilm or microfiche has been the archive medium of choice. The National Archives is constantly investigating new archive media for documents, and has recently begun a demonstration project of optical disk as a replacement for microfilm.

In preserving machine-readable data, the National Archives converts all source media to 9-track 6250 BPI 1/2-inch tape when creating the master and back-up copies. By consolidating to a single archive medium, greater control can be exercised over medium quality, standards, storage, and reader peripheral device maintenance. Because 1/2-inch tape has become, by virtue of its widespread use, a *de facto* machine-readable archive media standard, it is necessary to investigate its archival properties, particularly its useful lifetime.

Standard magnetic tape is composed of three component layers: the substrate or backing, the magnetic particles, and the binder system which binds the magnetic particles to the substrate.

Substrate. On magnetic tape the substrate is polyethylene terephthalate (PET) film. Studies by the National Bureau of Standards (Brown, Lowry and Smith, 1984) have concluded that, in proper storage conditions of 68 to 77 degrees Fahrenheit and 50 percent relative humidity, the lifetime of PET film exceeds 1,000 years, well exceeding an archive horizon of 50 to 100 years.

Magnetic particles. If protected from high heat and other sources of magnetism, the magnetic particles can be expected to retain their polarity indefinitely. In a protected environment, magnetic sources that are hard to eliminate include the Earth's magnetic field and the magnetic polarity of adjacent layers on the tape. The coercive forces of the magnetic particles are all above 300 Oe, which is more than 200 times the Earth's magnetic field (National Research Council, 1986), and 30 times the force from stray fields caused by exposure to electronic equipment common to most computer environments. The effect of magnetic forces on adjacent layers is often called print-through, and is known to have an extremely small effect at room temperature.

Binder. The binder used in most tapes is a polyester-urethane type. It is the weakest link. It is subject to wear when it comes into contact with heads of the tape peripheral readers, and it is especially vulnerable to exposure to water/humidity because the process of hydrolysis begins to occur. Storage at low relative humidity of 40 percent is mandatory.

Over long periods of storage, the PET substrate tension begins to relax. If a tape in this condition is mounted on a normal tape reader, the high acceleration forces encountered during the read process can cinch the layers together, potentially causing significant strain and stretch and eventual data loss. Most tape storage standards suggest retensioning tapes by gently rewinding.

With careful attention to the above deleterious effects, a magnetic tape may last twenty years or more. However, in that time frame, the equipment to read these tapes will almost certainly be obsolete and beyond the economical possibility of repair. As recently as 10 years ago, 7-track 800 BPI tape drives were prevalent; today it is difficult to find a manufacturer who will even maintain them. Today, the standard 1600 BPI 9-track tape drives are rapidly being replaced by 6250 and higher BPI tape drives. Cartridge tape drives are also emerging as a popular replacement for reel to reel tape machines. From an archival standpoint, the progress of technology creates management headaches. But with each succeeding generation of technology, the costs go down, and the storage density increases in substantial measure.

Recognizing the limitations of magnetic tape and the progress of technology, the U.S. Government has promulgated a regulation for the storage of magnetic tapes. Among other tape maintenance regulations, CFR 101-36.1207 "Maintenance of tape file," specifies for inactive storage "a 3 percent statistical sample of all reels of archivally valuable tape files shall be read annually to identify any loss of data and to discover its cause." Furthermore, "to prevent loss of information due to changing technology or to the aging of the storage medium, when appropriate, files shall be written and/or transferred to another machine-sensible medium." At the National Archives, this latter requirement has been interpreted as recopying to another medium every 10 years. Maintaining a static archive of magnetic tape media goes well beyond tape storage costs, and involves an active program of tape maintenance, quality control and continual recopying.

As the volume of tapes in an archive increases, the costs of maintenance rise accordingly. Economies of scale should decrease the annual per unit cost. However, it is likely that recopying to emerging technologies will reduce maintenance and storage costs to a greater extent.

Emerging storage technologies include:

(1) High-density magnetic tape cartridges.

(2) High-density magnetic reel-to-reel 1/2-inch tape

(3) Laser optical disks of various varieties.

(4) Optical tape (not yet viable).

The high-density magnetic tapes offer promising potential but still suffer from the same limitations as current tape technology: namely, wear and hydrolysis on the binder. Not discounting high-density tape, the new optical technologies have generated a high degree of interest and development.

Optical disk technology uses low-powered lasers to sense minute pits or holes in a thin ablative surface. Because the laser can be very narrowly focused, storage densities can be much higher than current magnetic media.

While new high-density magnetic media begin to approach the storage density of optical disks, there are some intrinsic advantages to optical disk media for archival purposes:

(1) In many optical disk technologies, there is a physical change to the ablative surface that cannot be reversed, i.e., pits or holes. If, as in most optical disks, the metal ablative surface is encased in a medium not subject to deterioration, normal oxidation or interaction with the atmosphere will not occur.

(2) Neither the ablative surface nor the encapsuling media comes into physical contact with the sensing mechanism.

Optical disk manufacturers are only now beginning to address the archival lifetime of the storage media. Some manufacturers are beginning to quote an archival life exceeding 30 years, though this is of course unproven. In most cases, it is the <u>chemical stability of the encapsuling media that determines the medium lifetime</u>. As with magnetic tape, the evolutionary lifetime of the disk drive and the related formatting standards is much less than 30 years. Currently, no optical disk formatting standard has been sufficiently accepted to guarantee the 10-year lifetime confidence level currently found in magnetic tape drives. Indeed, many optical disk drives will have a useful life exceeding the life of the format standards, some of which are already obsolete. For the general computer industry, deciding which optical drive technology will become popular and therefore preserve the investment is very difficult at this time. However, in an archival environment, which can be isolated from changes in the source and distribution media, the popular standard criteria may be less important than taking advantage of higher density storage to reduce overall costs. NASA has recognized this phenomenon and is proceeding ahead in at least one area with an optical disk technology. The National Space Science Data Center (NSSDC) of NASA is formulating plans with various methods and technologies to handle a vastly increased input of archived digital data from future sensors. To this end, they have an active program to examine optical disk technology and are in the process of contracting out the complete conversion of the Coastal Zone Color Scanner (CZCS) data to optical disk.

CZCS data captured on magnetic tape in November 1978 is now reaching its maximum archive life. Ninety percent of the data has never been looked at or processed. NSSDC is going to process every scene to the "Level 1 chlorophyll stage." While doing this process, it makes sense to rearchive to a new archive medium. The emerging 3.2 Gigabyte 12-inch double-sided optical disk platter has been chosen. With this technology, 100 Level 1 tapes can be put on one 12-inch platter. There are an estimated 23,000 Level 1 tapes requiring a minimum of 6,000 cubic feet of storage. All of this data, after copying to 250 optical disks, will occupy about 35 cubic feet of storage, an improvement in density of more than two orders of magnitude. From a cost standpoint, 250 optical disk platters @ $300 each is $75,000. 23,000 tapes at the low price of $10 each is $230,000. Using a computer system costing $150,000, including all peripherals, NSSDC is copying CZCS tapes at the rate of 4 per hour to optical disk. A cost comparison estimate of this exercise is shown below. These cost figures are only rough estimates and should not be construed as definitive, or as applying to NSLRSDA operation.

Present 5 Year Archive Costs

Media Costs	23,000 tapes @ $10	$230,000
Storage Costs	23,000 rack-stored tapes @ $5/yr. for 5 years	575,000
		$805,000

Conversion to Optical Disk

Capital Equipment		$150,000
Operator Time at 4 tapes/hour @ $35/hour		200,000
Media Costs	250 platters @ $300	75,000
Storage Costs	250 platters @ $5/yr. for 5 years	6,250
		$431,250

Considering that optical disk technology (and cost) will likely change in five years, a life-cycle of 5 years was used to price this comparison. Including the costs of conversion, the optical disk archive costs, over 5 years, are only half the cost of keeping existing magnetic tape. Should the optical disk technology last longer, these cost savings will significantly improve.

Should the NSLRSDA decide on an archive medium different from the source media, whether it is optical disk or high-density 1/2-inch tape, it is possible to reduce conversion costs further. If the volume were high enough, the distributing agency could copy directly to the archive medium. For example, ordering 10,000 Landsat MSS scenes would today result in 10,000 6250 BPI 9-track tapes. Assuming the archive medium is 3.2 Gigabyte optical disks, the NSLRSDA would have to recopy all 10,000 tapes to optical disk. In such a circumstance, it may be economical to provide the source agency with the optical media and devices, and write directly to optical disk.

In summary, optical disk technology, even in its present uncertain popularity and non-standard environment, can be an economical medium for the near-term archive future, and should certainly be considered as a preferred storage medium for the NSLRSDA.

The cost improvements of current optical disk technology are likely to be repeated with a new storage technology 5 or 10 years hence, as the present pace of technological evolution continues. The price/performance advantage of recopying to this new technology will also be repeated, and the cycle would begin again. Should the evolution of this storage technology cease, then the archive media lifetime will become the driving determinator of when machine-readable data are recopied.

At the beginning of this section, a primary issue to be determined was whether the machine-readable archive medium should be different from the source media. The analysis undertaken in this section suggests that the NSLRSDA will benefit from a single high-density archive medium. The reasons are summarized below:

(1) Taking advantage of high-density media will be justified by price/performance improvements in less than a 5-year cycle time.

(2) Use of a single medium permits selection of the medium best suited for archival purposes rather than distribution purposes.

(3) A conscientious receiving procedure for source sensor data must provide for back-up copies.

(4) While costs per gigabyte of storage will likely decrease, physical storage costs are likely to increase over time.

6.2.5 Machine-Readable Distribution Media

The value of an archive depends in part on the ready accessibility of information to present and future users, and PL 98-365 makes specific reference to "...timely access for parties requesting data." Archives of machine-readable information certainly are not immune from this requirement. Just as human-readable archives maintain facilities, such as copying machines, for reproducing original or microfilm documents, a machine-readable archive must maintain a computer system capable of reproducing archive data on commonly accepted distribution media.

Commonly accepted machine-readable distribution media at the present time include IBM PC-floppy diskettes and 9-track 1600 and 6250 BPI 1/2-inch magnetic tape. The information storage requirements of remotely sensed data precludes extensive use of floppy diskettes. At the present time, 9-track 1600 or 6250 BPI tapes are the universal remotely sensed data distribution medium, and almost all of the current machine-readable data are distributed on magnetic tapes of this type.

9-track tapes are not, however, necessarily the most <u>convenient</u> means currently available for storing and distributing such data. The advent of powerful and inexpensive microcomputers for remote sensing analysis has significantly increased the number of potential users for such data, and distributing agencies have been requested to provide data on more efficient media, such as optical disk. Faced with these requests, a distributing agency assesses the overall demand for that medium, and if economically reasonable begins supplying on that medium. In the case of optical disk, there is currently insufficient popularity of any particular format to justify such offering. The deciding factor in distribution media is popularity across the intended user community. The 9-track magnetic tape is the most popular distribution medium, not only in the remote sensing community but also in the general business and scientific minicomputer and mainframe community. However, it is a sine qua non that the distribution medium of choice will change over a relatively short time frame (5 to 10 years). Emerging distribution media include 9-track 6250 BPI magnetic tape and CD-ROM optical disks.

CD-ROM is a contender to emerge as the standard for published data distribution media ("published data" implies that multiple copies of the same data files are made at the same time from one master). A CD-ROM 5.25 inch optical disk can hold 550 megabytes of data, more than enough to hold one full Landsat TM scene. At about 10 copies, the CD-ROM format becomes economically competitive to magnetic tape.

In the NSLRSDA, what will be the volume of repeat sales of a particular scene? Future demand estimates such as this are most difficult to determine. In the case of Landsat and SPOT at the present time, the overall average number of times a scene is sold, if it is sold at least once, has been estimated to be about 1.5 times.

Interpreting this estimate, a large number of scenes are sold only once, and a relatively small number, estimated to be less than 1%, are sold many times. In this scenario, using magnetic tape to make single copies is more economical than CD-ROM distribution.

In the future, magnetic cartridge tapes, readable by personal computers, may also be an attractive option. They are compact, convenient to handle, and may be easier to handle than reel tapes.

In summary, the machine-readable distribution media of the NSLRSDA should be determined by the popularity of other distribution media used within the user community. What is popular will change over time to more efficient media, and the NSLRSDA must budget sufficient funds to keep pace. The life cycle of machine-readable media technology in the past has been less than 10 years, with serious new standards arising every 5 years. With new technology, such as optical tape currently in the research labs, it is expected this pace will continue for at least several decades.

6.2.6. Findings and Recommendations

At first sight the subject of machine-readable archival storage appears to be dominated by the lifetimes of the storage media, which have horizons somewhere in the 10 to 30 year region. Closer inspection reveals that the driving criterion is not storage media lifetime, it is read-and-write technology development. The cycle time for a new generation of read-and-write equipment, with corresponding increase in storage density, is 5 to 10 years. It is certainly significantly less than storage media lifetimes.

This dictates the strategy of machine-readable archive development. Storage can begin at once, as soon as the NSLRSDA begins to operate. It can employ magnetic tape, or optical disk, or both. However, it must be recognized from the beginning that <u>it will be both necessary and economically attractive to copy stored data every ten years or so</u>; the acquisition cost of new equipment, and the associated copying cost, must not be overlooked in budgeting the operation of the archive.

Note that these comments apply only to <u>machine-readable data</u>. For present film holdings, it makes no sense to think of recopying for the NSLRSDA. The high stability of the storage medium does not justify such a copying operation, even allowing for possible changes in user technology.

6.3 Archival Back-Up

Procedures for protecting valuable data, in machine-readable or other forms, are well-established. They call for storage of a complete copy of all data in a separate physical location, with the same procedures employed for data maintenance and review as in the main site. This has not been done in the past with Landsat data, nor is it done today in Landsat, SPOT or AVHRR repositories. Budget constraints have not permitted redundancy, although everyone recognizes its desirability. Limited redundancy, in the case of Landsat, has been achieved only when other storage, mainly at Goddard Space Flight Center, duplicated data stored at EDC.

The creation of complete back-up copies of the Basic Data Set in the NSLRSDA is regarded as highly desirable. Budgeting requests should include provision for such back-up. At the same time, recognizing probable funding limitations, the possibility of defining a "core data subset" from the Basic Data Set should be considered. Such a core data subset would constitute the minimum acceptable set of back-up data.

If no back-up can be established, the danger of such an omission must be made clear, and the consequences of data loss must be assessed.

7.0 SURVEY OF USERS

7.1 Objectives

The objectives of the user survey were to:

(1) contact the remotely sensed data user community representing Federal and State governments, private industry, and academia in the United States;

(2) inform the user community of the Land Remote Sensing Commercialization Act of 1984 and its language regarding the establishment and maintenance of a satellite land remote sensing data archive in the United States;

(3) obtain information regarding the user community's interest in a data archive and specific requirements for data to be placed in the archive;

(4) elicit general comments and opinions from the community regarding many of the issues associated with archiving satellite data, such as data storage media, desired input/output products, etc., and

(5) seek information on specific applications and geographic areas of long-term scientific interest.

7.2 Approach

A questionnaire (Figure 7.1) comprising ten questions was designed principally for use in a telephone survey, but also was suitable as a guide for face-to-face interviews and as a mail-out. As it turned out, the majority of surveys were done by telephone. Each interviewee was contacted without forewarning, given an introduction as to the purpose of the survey, and asked a series of questions. Because of this somewhat ad hoc approach, interviewees did not have an opportunity to give considered thought to the concept of the NSLRSDA, or their own related archiving requirements. Most interviewees had not given extensive prior thought to long-term archival data uses. However, a follow-up letter was sent out to the survey participants thanking them for their contribution, informing them of the February 3, 1987 Public Meeting at the University of Maryland, and encouraging them to contact EarthSat with additional comments or questions.

7.3 Results

More than eighty groups representing Federal and state government, private industry, and academia were contacted. Table 7.1 lists the groups and each survey participant. Figure 7.2 illustrates the organizational composition of the survey participants. Appendix 4 contains the complete mailing address and telephone number of each contact. The list is numerically coded to the responses to each question, which are given in Appendix 5.

Ten questions were asked dealing with:

- User applications
- Need for historical archive data

FIGURE 7.1

ARCHIVES PROJECT QUESTIONNAIRE

The National Oceanic and Atmospheric Administration (NOAA) has been directed by Congress in Section 602 of the Land Remote Sensing Commercialization Act of 1984 to establish and maintain a data archive of Landsat data for historic, scientific and technical purposes. To further this objective, Congress has directed that all Landsat data acquired by EOSAT will be made available to the Archives 10 years after acquisition. It will not be feasible to save all of the imagery acquired so a selection process must be implemented.

The purpose of this questionnaire is to enlist your assistance in providing information that will help determine what type, how much and for what areas satellite imagery should be preserved for posterity.

Date: ______________

Organization: ____

____

____

____

Interviewee: ________________________ Telephone: ______________

Title: ____

What are the phenomena of interest or applications in your work effort?

Do you visualize that 10 or more years from now you will have a need for satellite imagery acquired today? If so, in what manner, form and how much?

FIGURE 7.1 (continued)

Would your needs be primarily for U.S. coverage or would other parts of the world be required? If other parts of the world are of interest, would you identify the areas of major interest?

Would your imagery needs be satisfied with that provided by U.S. satellites or would you require data from foreign owned systems?

Resolution requirements, both spectral and spatial, of archived data.

Would you have a need for radiometric calibration data for the imagery?

Is season or time of acquisition important; if so, what season or time?

What in your opinion would be adequate frequency of coverage in the archives of the Landsat type data? Other? For the world? U.S.?

Do you have an opinion as to how the data should be archived? Digital tapes, photographic, other? Are there any archival methods that you feel should be considered?

FIGURE 7.1 (continued)

Certain parts of the world are subject to rapid change due both to man's activities and to natural processes. Examples are the southward advance of the Sahara desert in the Sahel and deforestation in the tropical rain forests. Can you recommend other areas that you believe are similarly sensitive and should have more than casual coverage in the Archive?

Catastrophic events, both man-made and natural, occur randomly. Some of these may have some predictability, such as imminent eruption of some volcanoes. Most, however, are known only after the event. Can you suggest types of catastrophic events that are of such significance that the Archives should attempt to acquire comprehensive coverage for preservation?

Other people you think we should interview or poll:

Name	Telephone Number

Comments of the Interviewer:

TABLE 7.1

SURVEY PARTICIPANTS

Organization	Participant
FEDERAL GOVERNMENT AGENCIES	
U.S. Fish and Wildlife Service (Nat. Wetland Inv.)	John Montanari
U.S. Department of Agriculture (FCAD)	Pat Ashburn
U.S. Department of Agriculture	John Price
U.S. Department of Energy (Office of Energy Res.)	Frank Wobber
U.S. Geological Survey (Branch of Geophysics)	Richie Williams
U.S. Geological Survey (Branch of Geophysics)	Mel Podwysocki
U.S. Geological Survey (Branch of Geophysics)	Larry Rowan
U.S.A.I.D. (Off. of Forestry & Natural Resources)	Dan Deely
U.S. Environmental Protection Agency (EPIC)	Tom Osberg
U.S. Bureau of Census (Geography Division)	Robert Durland
National Aeronautics and Space Administration	Nicholas Short
National Aeronautics and Space Administration	Paul Lowman
STATE GOVERNMENT AGENCIES	
Arizona State Lands Department	Bill Bayham
Michigan Resources Inventory Program	Michael Scieszka
Minnesota State Planning Agency	Paul Tessar
New Jersey Dept. of Environmental Protection	Susanne Rohardt
South Carolina Land Resources Commission	Dan Fairey

TABLE 7.1 (continued)

ACADEMIC INSTITUTIONS

University of Alabama (Department of Geography)	Tom Lo
University of Alabama (Department of Geography)	Donald Brandes
University of Arizona (Remote Sensing Center)	Chas. Hutchinson
Boston University Center for Remote Sensing	Farouk El-Baz
Boston University Center for Remote Sensing	Fritz Hemans
University of California (Dept. of Geography)	John Estes
University of Colorado	H. Morrow-Jones
Dartmouth College (Department of Geography)	David Lindgren
University of Delaware (College of Marine Studies)	Vitorio Klemas
University of Florida (Department of Geography)	James Henry
University of Georgia (Department of Geography)	Roy Welsh
George Mason University (Department of Geography)	Barry Haack
University of Hawaii at Manoa (Dept of Geography)	Wingert Everett
Hunter College (Department of Geography)	Allan Strahler
University of Illinois (Department of Geography)	Thomas Frank
University of Kansas (Department of Geography)	Jim Merchant
University of Maine (College of Forest Resources)	Louis Morin
University of Maryland (Remote Sensing Sys. Lab)	Bob Ragan
Michigan State University (Ctr for Remote Sensing)	Bill Enslin
University of Minnesota (College of Forestry)	Doug Meisner
Mississippi State (Dept of Geography and Geology)	Ron Shaklee
University of Montana (Department of Geography)	John Donahue
Murray State College (Department of Geosciences)	Luis Bartolucci
University of Nebraska (Conservation/Survey Div)	Donald Rundquist
University of Nevada (Mackay School of Mines)	David Mouat
University of Nevada (Department of Geography)	C. Exline
University of North Carolina (Dept of Geography)	Arthur Hawley
University of North Dakota	William Dando
University of Oklahoma (Department of Geography)	T. H. Williams
Oregon State University (Environ. R/S Lab)	B. Schrumpf
Penn State University (Earth Sys. Science Center)	Eric Barron
Penn State University (Env Resource Research Inst)	Wayne Meyers
Purdue University (Dept. of Forestry & Nat. Res.)	R. M. Hoffer
Rutgers University (Dept Environmental Resources)	Teuro Airola
South Dakota State University (R/S Center)	K. Dalsted
University of Tennessee (Department of Geography)	John Rehder
Texas A&M University (Ctr Strategic Technology)	Chuck Smith
University of Utah (Department of Geography)	Chung-Myun Lee
University of Vermont (Department of Geography)	Aulis Lind
Virginia Tech	Jim Campbell
University of Washington (Urban Planning Dept.)	Frank Westerlund
University of Washington (Department of Geography)	Tim Nyerges
Washington University (Dept Earth/Planetary Sci)	Edward Guinness
University of Wisconsin (Env Remote Sensing Cntr)	Thomas Lillesand
University of Wyoming (Department of Geology)	Ron Marrs

TABLE 7.1 (continued)

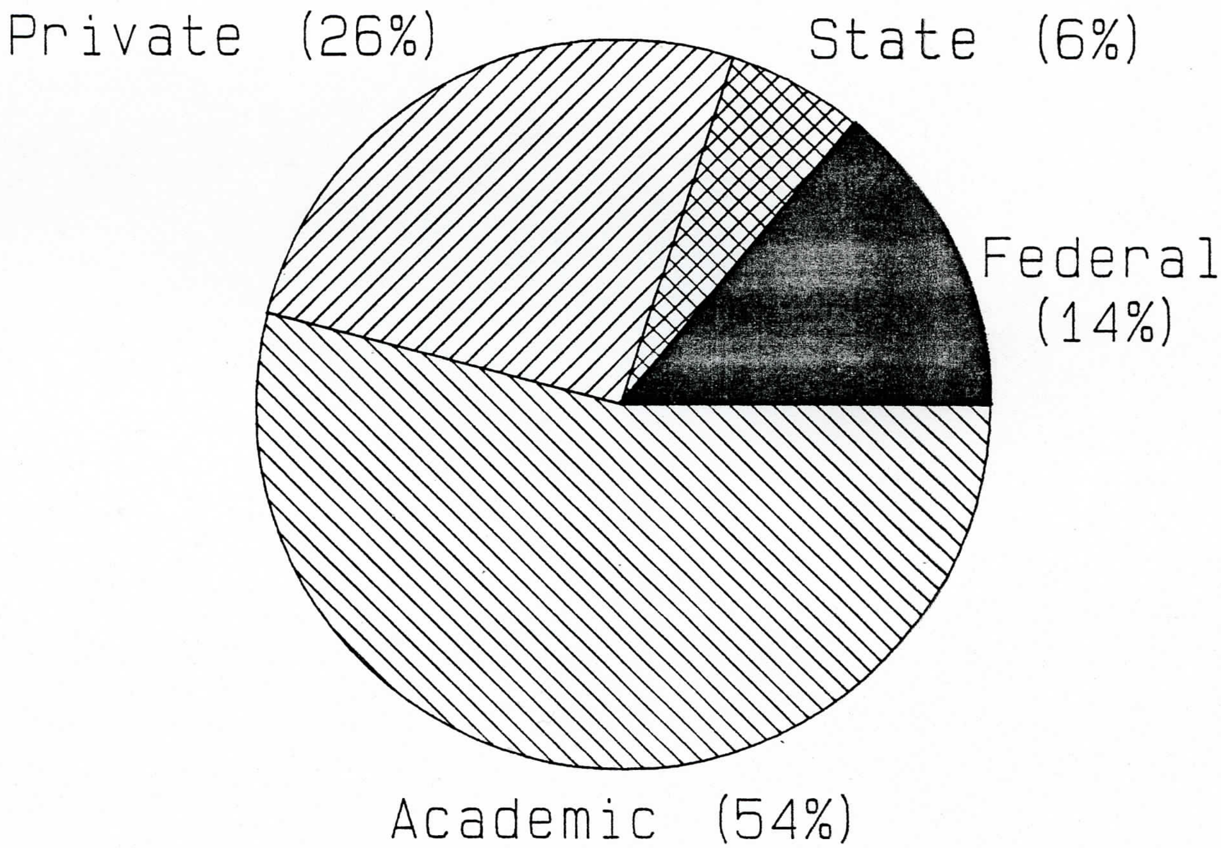
PRIVATE ORGANIZATIONS

Aeroterra, S.A.	Eduardo Viola
AMOCO	Ralph Baker
J. Aron & Company	Ron Levy
Balfour Maclaine International	George Parker
Chevron Overseas	Jim Ellis
Continental Grain Company	Paul Smolen
Earth Satellite Corporation	John Everett
Environmental Research Institute of Michigan	Buzz Sellman
ERDAS Advance Technology Development Center	Bruce Rado
General Buscuit Brands, Inc.	James Carroll
Ray Kreig and Associates	Ray Kreig
Lonray, Inc.	Richard Emanuele
Lunar and Planetary Institute	Kevin Burke
Mobil Research	Bob Borger
Nabisco Brands	Anne Bostick
National Center for Atmospheric Research	F. Bretherton
Private Consultant	Charles Dorigan
Private Consultant	Benedict Levin
Sunshine Biscuits	Peter Faust
Sprague and Rhodes Coffee Trading Corporation	Eugene David
Sun Exploration and Production Company	Stewart Marsh
Westway-Merkuria Corporation	Laureano Suarez

Figure 7.2

ORGANIZATIONAL COMPOSITION

SURVEY PARTICIPANTS



84 PARTICIPANTS

- Coverage areas of interest
- Satellite system requirements
- Temporal data requirements
- Forms for archiving data

7.3.1 User Applications

The first survey question addressed current and future user applications of remotely sensed data. The purpose of this question was to establish a profile of data user application requirements and to confirm that historical archived data would in fact be valuable to users for supporting their future research and operational scientific activities.

A diverse group of remotely sensed data users was surveyed, representing more than twenty-five separate applications of satellite data to the botanical, hydrological, geological, and meteorological sciences. Table 7.2 lists the primary data applications of those surveyed.

7.3.2 Need for Historical Archive Data

Survey participants were asked if they would have a need, ten or more years from now, for satellite data acquired today. In what manner, form, and how much? This was a key question for establishing data user interest in the creation and maintenance of the NSLRSDA. Because of the ad hoc method of contacting data users by phone, very few survey participants had given considered thought to the issue of data volume, i.e., how much data they would anticipate archiving.

TABLE 7.2

SATELLITE LAND REMOTE SENSING DATA APPLICATIONS
By User Type

(Number of Respondents for Each Application)

Application	Federal	State/Local	Private	Academic
Agriculture	2	2	9	8
Alpine Envirn.	0	0	0	1
Archaeology	0	0	0	1
Climatology	0	0	1	2
Coastal Studies	0	0	1	5
Change Detection	0	1	0	2
Desert Envirn.	0	0	0	1
Envirn. Degradation	1	0	1	3
Estuarine Studies	0	0	1	1
Forestry	1	2	1	8
Geology	5	0	7	9
GIS	0	2	1	1
Feature Extraction	0	0	0	1
Ice	1	0	0	0
Land Use/Land Cover	0	4	3	13
Mapping	0	0	0	1
Natural Resources	0	2	2	4
Rangeland	0	0	0	1
Soils	0	1	1	3
Teaching	0	0	0	2
Urban Studies	1	0	2	2
Vegetation	1	3	0	6
Water Resources	1	1	3	4
Wildlife Habitat	0	0	0	1

Seventy-eight percent (78%) of those surveyed said they would have a need for archived data (Figure 7.3). Most said they would use the data for change detection or as a historical record. Twenty-five respondents offered information regarding the form in which they would use the data. Of these, 56% said they would use primarily digital, 8% primarily photographic, and 36% both.

7.3.3 Coverage Areas of Interest

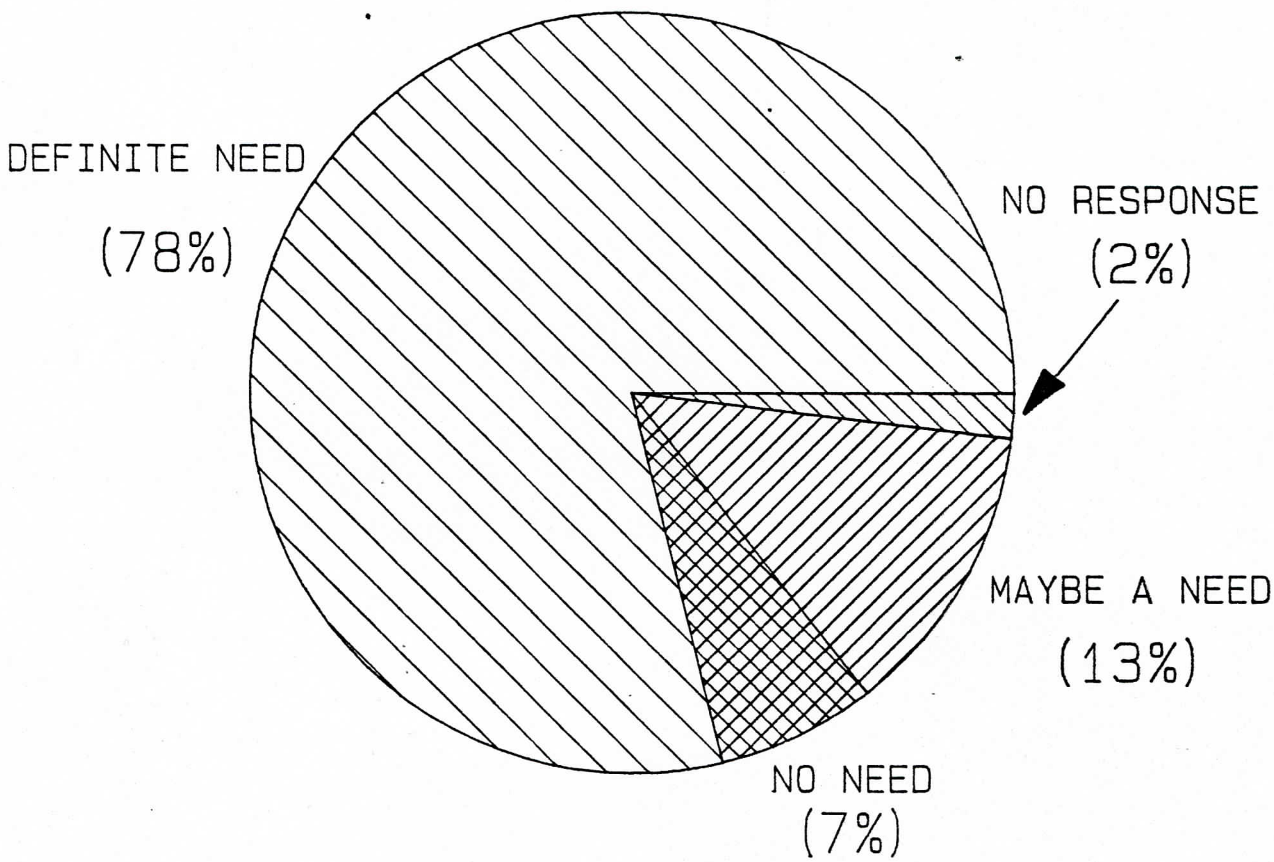
Three of the survey questions were formulated so as to determine data user interest in locational coverage in the United States and abroad. These questions were asked to: (a) establish and confirm the need for worldwide data archive coverage; and (b) determine the feasibility or difficulty of developing an archive data collection plan and procedure.

The majority of those surveyed (68%) wanted both U.S. and non-U.S. data coverage in the archive (Figure 7.4).

Survey participants were also asked to identify those areas and/or phenomena which in their opinion are sensitive, i.e., subject to rapid change due either to man-made activities or to natural processes, and should thus receive particularly complete coverage in the archive. Table 7.3 lists sensitive areas and/or phenomena indicated by the survey participants. They are not listed in any order of importance, because such a ranking would be difficult to make and probably highly controversial.

Figure 7.3

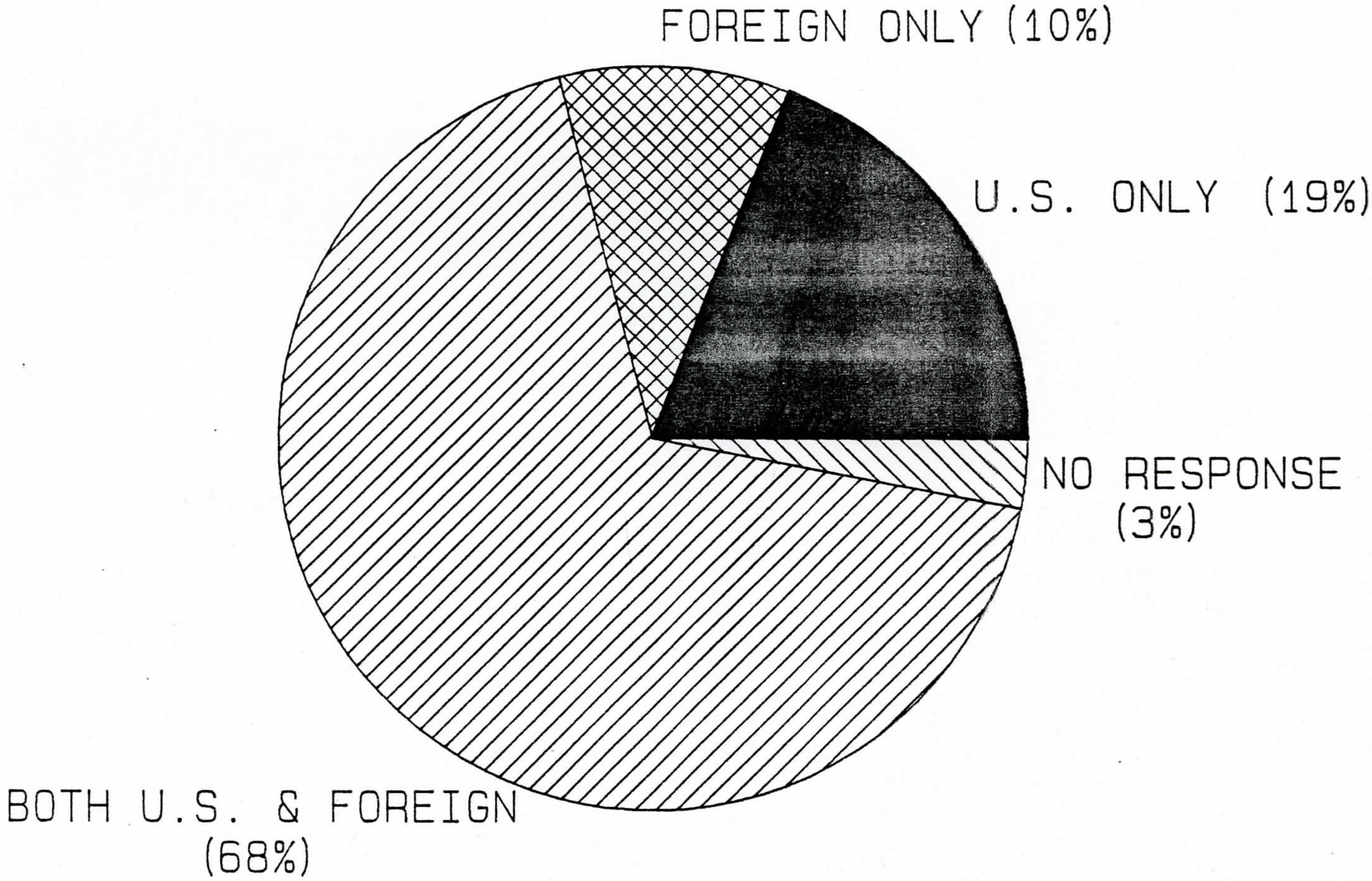
NEED FOR HISTORICAL ARCHIVE DATA



84 RESPONDENTS

Figure 7.4

DATA COVERAGE NEEDS
U.S. VERSUS FOREIGN



84 RESPONDENTS

In addition, each participant was asked to identify those catastrophic events of such significance that the archive should attempt to acquire comprehensive coverage for preservation. Table 7.4 lists these events, also not in order of importance.

7.3.4 Satellite System Requirements

Because this study develops and proposes both an architecture and an initial content for the archive, questions were posed to determine where the data collection efforts should be focused, i.e., what should be the data composition of the archive?

Survey participants were asked if their data needs could be satisfied by data from U.S. satellite systems, or if they would also require data from foreign-owned satellite systems. The majority (more than 74% of those surveyed) indicated that their archive data needs could be served by data from either U.S. or foreign satellite data systems. If data were not available from U.S. satellite systems, many surveyed indicated they would use whatever systems suited their particular project needs. Figure 7.5 summarizes the responses to this question.

Also as part of determining an appropriate archive initial content, survey participants were asked their spatial and spectral data resolution requirements for archived data. Most surveyed cited either Landsat MSS or TM, SPOT, or AVHRR, spatial and spectral resolutions in their responses. Responses were varied and application or discipline-specific, suggesting a need for a range of spatial and

Table 7.3: <u>Sensitive Areas or Phenomena Which Should Receive More than Casual Archive Coverage</u>
(Not listed in order of importance)

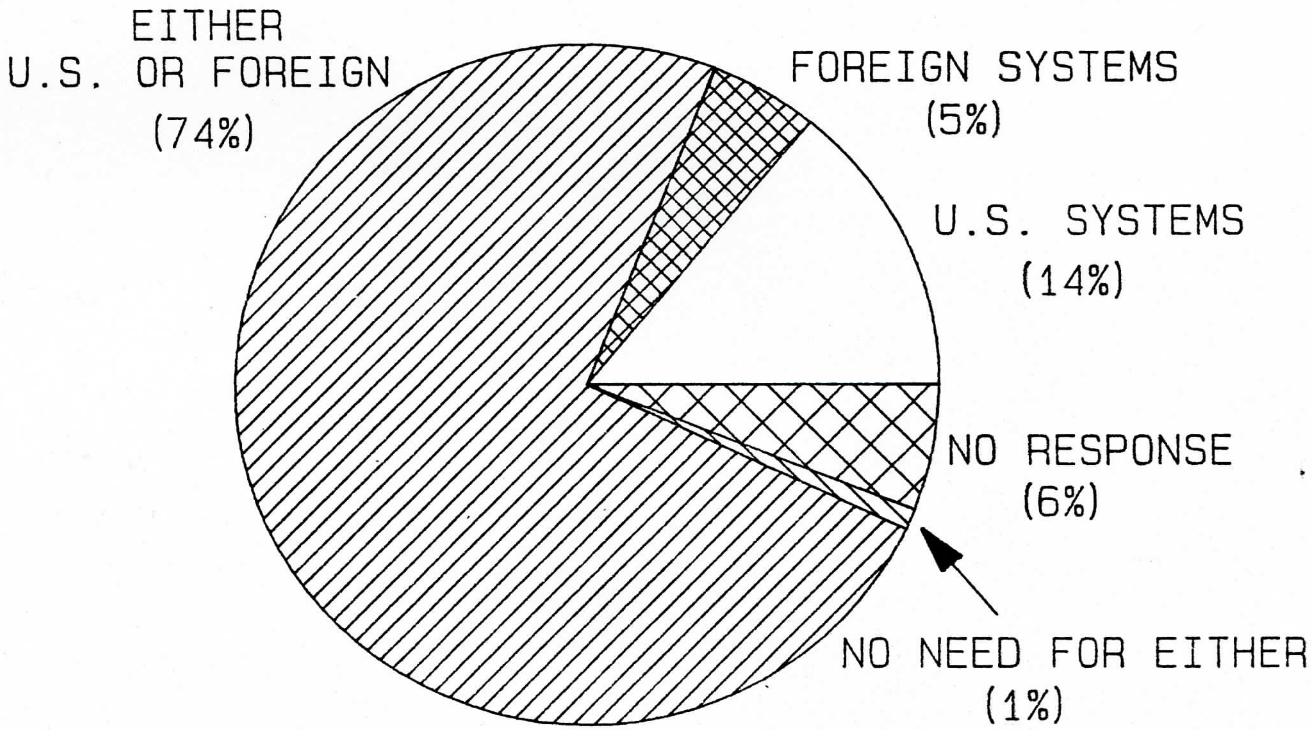
Acid Rain Impact Areas
Major Agricultural Development Areas
Coastal Areas
Desertification Areas
Drought Areas
Flood Zones
Glacial Advances
Insect Damage Areas
Land Use Change
Landslide Areas
River Deltas
Seismic Zones
Snow Cover Changes
Tropical Deforestation
Major Urban Expansion Areas
Water Quality Changes

Table 7.4: <u>Significant Catastrophic Events for Comprehensive Archive Coverage</u>
(Not in order of importance)

Abnormal Crop Production
Coastal Storms
Droughts
Dust and Sand Storms
Floods
Forest Fires
Insect Infestations
Landslides
Land Use Change
Nuclear Events
Large Oil Spills
Rapid Rural Change
Seismic/Volcanic Events
Strip Mining
Vegetation Disease

Figure 7.5

NEED FOR DATA FROM U.S. VS. FOREIGN SATELLITE SYSTEMS



84 RESPONDENTS

spectral resolution data. There is an apparent need in the archive for both low-resolution, broad-region, one-kilometer or even lower spatial resolution data, as well as higher spatial resolution (10-meter or better) coverage of smaller areas. Similarly, spectral resolution needs range from the broader visible to near-infrared bands of MSS and SPOT, to the more narrowly-defined spectral bandwidths of TM and future sensors, required for other types of botanical, hydrological, and geological studies. Figure 7.6 illustrates the range of survey participant preference for various spatial and spectral resolutions for archived satellite land remote sensing data.

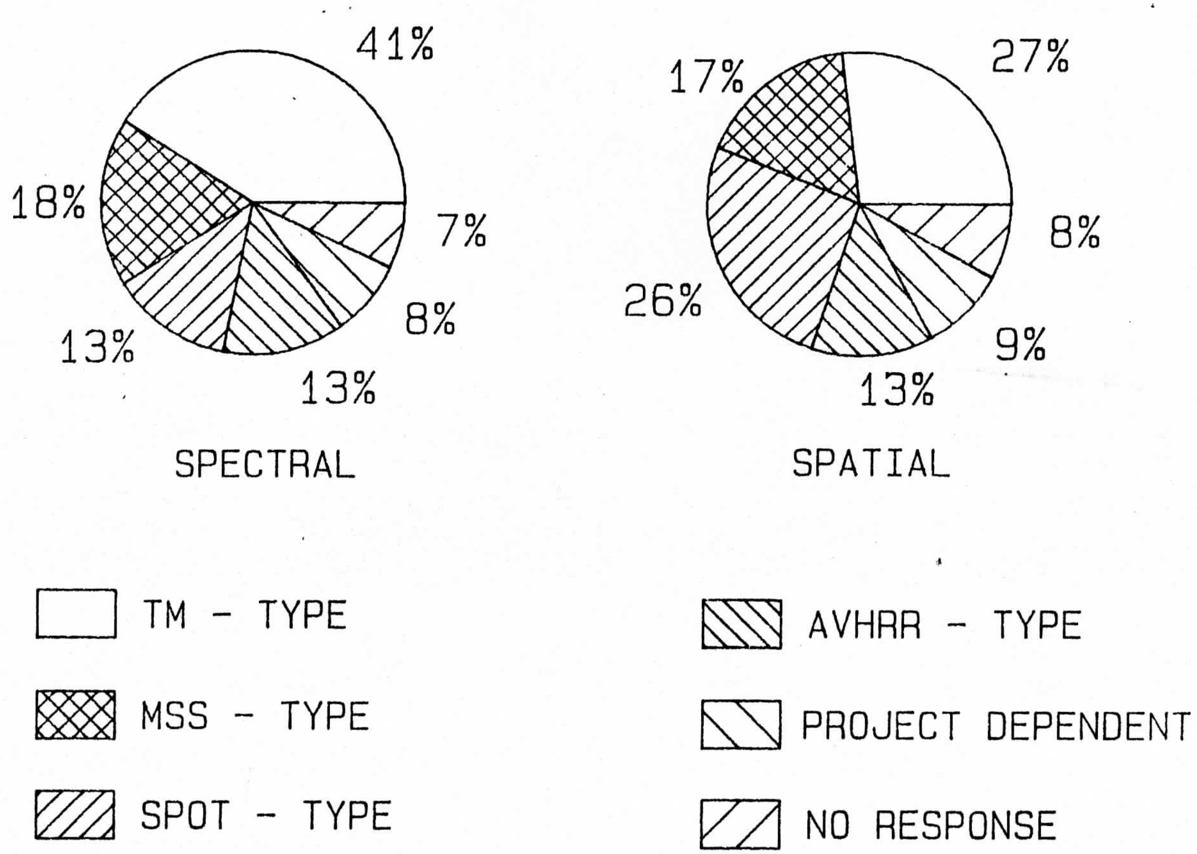
7.3.5 Temporal Data Requirements

In order to develop an effective archive data collection plan and procedure, it is necessary to know the temporal data requirements of land satellite remotely sensed data users. When asked if season or time of data acquisition was important to the data user, and if so, what season or time, the majority of respondents indicated that season and time are important, but highly project-dependent. Many of the respondents stated a need for seasonal data, since different information was extractable from images over their areas of interest during different seasons.

When asked what should be the frequency of data coverage for inclusion in the archive, the most often cited response was "seasonal," with "annual" being the second most requested frequency.

Figure 7.6

SPECTRAL AND SPATIAL RESOLUTION PREFERENCES



7.3.6 Forms for Archiving Data

Survey participants were asked to provide their opinion as to how, or in what form, the data should be archived, i.e., digital, photographic, etc. Fifty-one percent felt that the data should be archived in primarily digital CCT form, two percent thought hardcopy or photographic form should be used as primary storage medium, twenty-nine percent said that both digital and photographic should be archived, and six percent indicated optical disk as a suitable storage medium. Figure 7.7 summarizes this response.

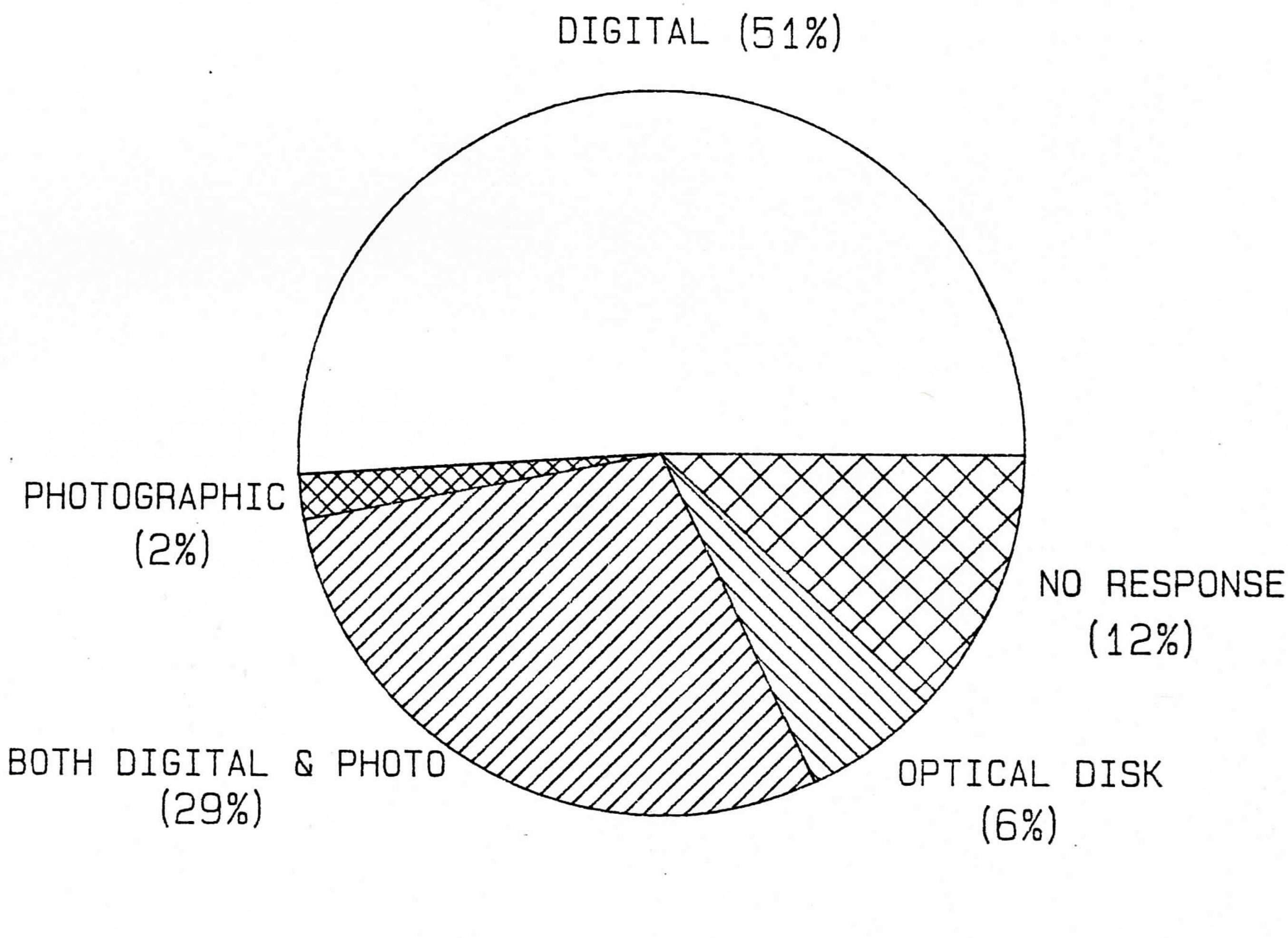
7.3.7 Significant Comments of the Survey Participants

Although the survey was completed on a "no-notice" basis, several of the survey participants had obviously given previous thought to the concept of a NSLRSDA. Their comments are viewed by the project team as significant and worthy of note. Several of these points were also made by individuals who attended the public meeting held at the University of Maryland on February 3, 1987.

Data costs, i.e., the purchase costs for data from the NSLRSDA, was a concern of several of the survey participants. Some indicated that satellite data costs now for Landsat and SPOT data are prohibitively expensive, and that careful consideration should be given to the establishment of more reasonable purchase prices for archived data. A concern was also expressed that commercialization would affect data cost and availability and that establishment of a data archive may help to resolve the data cost and availability issue.

Figure 7.7

HOW DATA SHOULD BE ARCHIVED



84 RESPONDENTS

The importance of collecting and maintaining satellite data in an archive for answering yet-to-be-formulated questions regarding the earth and its resources was stressed, as well as the fact that future methods for processing archived data could make the data even more useful than they are today.

In order to be successful in applying future image processing methods to archived data, it was also emphasized that care should be taken not to overly-process the data placed in the archive. The archived data should be as close to the originally collected signal as possible, and contain only geometric and radiometric corrections, sufficient to make the data useful to users lacking highly specialized processing equipment.

Some of those surveyed recognized that it will be impractical and cost ineffective to archive all collected satellite land remote sensing data. They offered their opinion as to the minimum practical data collection frequency by indicating that seasonal coverage, every five to ten years, would be desirable, with more active areas of interest receiving one- to three-year coverage, and some areas requiring only one-time multi-season coverage.

For determining those sensitive areas around the world which should receive coverage in the archive, it was recommended that the National Science Foundation's Long-Term Ecological Research Sites (LTERS) may provide important input in the data collection decision making process.

It was also recommended by several of those surveyed that a committee should be formed to decide on a regular basis, perhaps semi-annually, what areas should receive more than casual coverage in the archive, and that this method would generate public interest, as well as additional points of contact.

Finally, it was noted that people will be severely critical if this nation does not obtain and preserve land satellite remotely sensed data for looking back in time at resource changes.

7.4 Survey Summary

The survey achieved its principal objectives of confirming satellite data user interest in the establishment and maintenance in the United States of a NSLRSDA, and of eliciting specific comments and opinions on archive data collection requirements, content, storage and retrieval characteristics and products.

A key finding was that, because of the great diversity in satellite data applications requiring a wide range of geographic, temporal, spatial and spectral data characteristics, anything less than seasonal data acquired annually over most of the globe would undoubtedly create data gaps in the archive for many data users. However, the collection cost implications of such a strategy were also recognized.

It was well established by the survey that digital data should be stored in digital form, but should be made available to the archive customer as either a hardcopy photographic product or as a digital product suitable for image processing. The suitability or appropriateness of newer or future data storage media such as optical disk was not established.

It was also recommended that some form of index, catalogue or browse facility be made available to assist potential data purchasers in identifying data held in the NSLRSDA and in assessing its quality. The sophistication of the index or catalogue system could range from a hardcopy listing available as a mail-out or on-line to potential customers, to a pictorial index or catalogue (most likely in microfilm form) which would allow the customer to view cloud cover distribution and areal coverage.

Any photographic browse file tends to be expensive, and slow to generate. There is a real need for fast and inexpensive ways to generate browse file images from digital data.

8.0 PUBLIC MEETING

8.1 Purpose

A one-day meeting was held to:

(1) inform the general public about the legal mandate to establish and maintain a NSLRSDA, as outlined in the Land Remote Sensing Commercialization Act of 1984;

(2) present to the public the goals and objectives of the archive study;

(3) provide an overview of the preliminary results of the study;

(4) encourage an open forum for public discussion of the study goals, objectives, and results; and

(5) elicit the views and opinions of the public regarding the archive, its philosophy, characteristics, and method(s) of operation.

8.2 Registered Attendees

An announcement of the public meeting was published in "The Washington Remote Sensing Newsletter" and the January issue of "Photogrammetric Engineering and Remote Sensing," the journal of the American Society for Photogrammetry and Remote Sensing. In addition, an announcement was mailed out to all participants in the September 10, 1986 workshop for Federal agencies "Defining the Basic Data Set for the National Satellite Land Remote Sensing Archive," and the telephone survey participants as of January 20, 1987. Figure 8.1 is a copy of the public announcement.

The meeting was held at the University of Maryland's Center for Adult Education on February 3, 1987.

Figure 8.1

PUBLIC MEETING

FOR THE NATIONAL SATELLITE LAND REMOTE SENSING ARCHIVE

The Land Remote-Sensing Commercialization Act of 1984 directs the National Oceanic and Atmospheric Administration (NOAA) to establish and maintain a U. S. archive of land remote-sensing satellite data for historical, scientific, and technical purposes, including long-term global, environmental monitoring. The law also specifies that the global, land remote-sensing satellite data holdings planned for this archive ("the basic data set") will consist of data from existing and future, U. S. and foreign satellite systems.

Earth Satellite Corporation, in cooperation with NOAA and the U. S. Geological Survey, is sponsoring this public meeting to exchange information and views about the archive with all interested persons, including academic, private sector, state and local government users and producers of land remote-sensing satellite data. The meeting will be held:

Tuesday, February 3, 1987
9:30 a.m. to 4:00 p.m.

The University of Maryland
Adult Education Center, Room 2100
(SW corner of the campus)
College Park, Maryland

REGISTRATION: Registration is free of charge. A working luncheon will be provided at no cost for attendees who have pre-registered by January 23, 1987. However, public luncheon facilities are available at the Center. For pre-registration and additional information, contact:

Earth Satellite Corporation
7222 47th Street
Chevy Chase, Maryland 20815
(301) 951-0104

A list of pre-registered attendees is presented in Appendix 6 of this report. Actual attendees are listed in Appendix 7. A total of forty-three people attended the meeting. Figure 8.2 illustrates the organizational composition of the group.

8.3 Meeting Agenda

The meeting was held from 9:30 a.m. to 4:00 p.m. The morning was devoted to providing information to the attendees by those conducting the archive study. Presentations were made by NOAA and EarthSat personnel dealing with:

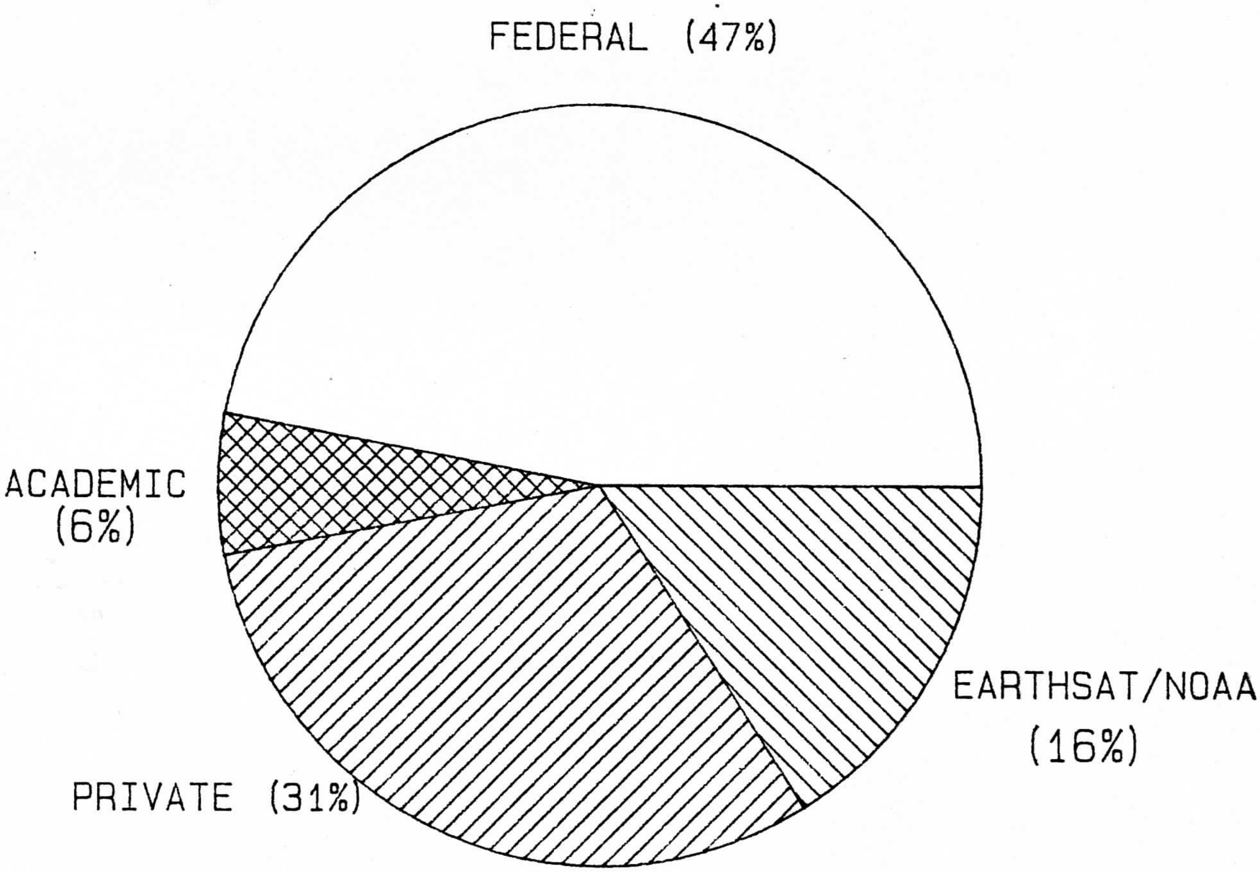
- o Introductions and Review of Project Objectives
- o Review of Data Sources and Data Bases
- o User Survey Results

The afternoon allowed for direct participation by the attendees, by dividing them into disciplinary sub-groups where they could discuss their specific interests in and uses for a NSLRSDA. It was recognized that because of the varied applications of satellite data to the biological, physical, and cultural sciences, archive requirements would also vary, and could best be categorized and evaluated within these disciplinary sub-groups. The sub-groups were:

- o Earth Sciences
- o Renewable Resources
- o Environmental Impact
- o Archival Questions

Figure 8.2

ORGANIZATIONAL COMPOSITION
PUBLIC MEETING ATTENDEES



43 PARTICIPANTS

Upon completion of the sub-group discussions, the entire group of attendees reconvened for a group discussion of data needs, a final review, and a summary of the day's activities.

The complete meeting agenda is presented in Figure 8.3.

8.4 Results

The morning speakers encouraged questions from the audience either during or at the completion of each presentation.

8.4.1 Review of Data Sources: Characteristics of Past, Present and Proposed Spaceborne Imaging Systems

A review of the data sources included the characteristics of past, present and proposed satellite land remote sensing systems. The audience was informed that sensors had been grouped according to a three-tiered hierarchy, as described earlier in this report.

In addition to presentation of this hierarchy of sensors, descriptions of selected Group One and Group Two sensors were also presented using the format presented in Section 4.0.

As a result of this presentation and subsequent comments received from the meeting attendees, a substantial revision was made of their hierarchy, and a complete re-evaluation made of the sensors, their characteristics, and their selection of inclusion in the

FIGURE 8.3: Public Meeting Agenda

AGENDA

9:30 - 10:00	Introductions; Review of Project Objectives	SHEFFIELD/ HARWOOD
	• The purpose, scope and time scale of the National Satellite Land Remotely Sensed Data Archive	
10:00 - 10:40	Review of Data Sources - Part I	
	• Characteristics of past, present and proposed spaceborne imaging systems	DURANA
10:40 - 10:50	Coffee Break	
10:50 - 11:30	Review of Data Sources - Part II	ANDERSON/ RUSSELL
	• Data formats and volumes	
	• Storage Technology	
11:30 - 12:00	User Survey Results	GAROFALO
	• Form of questionnaire employed	
	• Areas of respondent interest	
12:00 - 1:00	Lunch	
1:00 - 2:45	Disciplinary Discussions (in sub-groups)	SHEFFIELD
	Earth Sciences Group Geology Geophysics Soils Geobotany Glacialogy	RUSSELL
	Renewable Resources Group Forestry Agriculture Rangeland Wetlands	GAROFALO
	Environmental Impact Group Water Pollution Landuse Desertification Deforestation	DURANA

	<u>Archival Questions Group</u> Storage Retrieval Cataloging Browse Facility System Costs	**SHEFFIELD**
2:45 - 3:45	Group Discussion of Data Needs 1. Geographical Area 2. Seasons 3. Frequency 4. Sensors 5. Collateral data collection 6. Scientific versus programmatic needs	**SHEFFIELD**
3:45 - 4:00	Final Review and Summary	
4:00	Conclusion	

NSLRSDA data sources. This re-evaluation is presented in Section 4.0 of this report where a summary listing which comprises the final selection of sensors in this three-tiered hierarchy is given.

8.4.2 Review of Data Sources: Data Formats and Volume

A presentation was made of the EROS Data Center's (EDC's) satellite data archive, principally to give participants a feeling for data formats and volumes currently handled by an operational satellite data collection and distribution facility, and to provide an indication of what might be expected in a future data archive. Data volumes from currently archived digital (e.g., Landsat) and non-digital (e.g., Apollo/Gemini) products were presented (see Section 5.0). Also, numbers of scenes acquired per day from Landsat and SPOT were presented, as well as the current EDC Landsat archive volume versus projected yearly acquisition volume and planned SPOT and next-generation systems. Also, data volumes per scene for Landsat MSS and TM, SPOT and a possible next generation system were presented.

8.4.3 Review of Data Sources: Storage Technology

The third presentation of the morning provided the participants with information on various storage media being considered for the data archive. These included: (1) photographic (films or prints); (2) magnetic tape (digital, analog, cartridge); (3) magnetic disk; (4) optical disk; and (5) optical tape. Tables 8.1 through 8.6 list a variety of storage considerations for each of these media.

Tables 8.1–8.6: Media Descriptions Provided for Review at the Public Meeting

Table 8.1: Storage Media Considered

PHOTO (FILM, PRINT)

MAGNETIC TAPE (DIGITAL, ANALOG, CARTRIDGE)

MAGNETIC DISC

OPTICAL DISC

OPTICAL TAPE

Table 8.2: Film

INFORMATION PRESERVING: LOW

STORAGE DENSITY: VERY HIGH

LIFETIME: MORE THAN 100 YEARS

CURRENT USE: EXTENSIVE

IMPROVEMENT POTENTIAL: LOW

READ SYSTEM COST: MINIMAL

READ SYSTEM LIFETIME: UNLIMITED

COST/BIT: VERY LOW

Table 8.3: Magnetic Disc

INFORMATION PRESERVING: VERY GOOD

STORAGE DENSITY: LOW, BUT RAPIDLY IMPROVING

LIFETIME: 20 YEARS OR LESS

CURRENT USE: RARE FOR ARCHIVAL USE

IMPROVEMENT POTENTIAL: GOOD FOR DENSITY
LIFETIME - UNKNOWN

READ SYSTEM COST: MODERATE TO HIGH

READ SYSTEM LIFETIME: LOW (10 YEARS OR LESS)

COST/BIT: HIGH

Table 8.4: Magnetic Tape - Analog

INFORMATION PRESERVING: GOOD, BUT NOT PERFECT

STORAGE DENSITY: VERY HIGH

LIFETIME: 20 YEARS

CURRENT USE: EXTENSIVE

IMPROVEMENT POTENTIAL: LOW

READ SYSTEM COST: HIGH (BECAUSE FEW UNITS)

READ SYSTEM LIFETIME: LOW (10 YEARS)

COST/BIT: LOW

Table 8.5: Magnetic Tape - Digital

INFORMATION PRESERVING: VERY GOOD

STORAGE DENSITY: MODERATE

LIFETIME: 20 YEARS

CURRENT USE: EXTENSIVE

IMPROVEMENT POTENTIAL: GOOD FOR DENSITY
LIFETIME - UNKNOWN

READ SYSTEM COST: MODERATE TO HIGH
DENSITY RELATED

READ SYSTEM LIFETIME: LOW (10 YEARS)

COST/BIT: LOW

Table 8.6: Optical Disc

INFORMATION PRESERVING: VERY GOOD

STORAGE DENSITY: HIGH TO VERY HIGH

LIFETIME: 10 YEARS(??)

CURRENT USE: VERY SMALL

IMPROVEMENT POTENTIAL: VERY HIGH

READ SYSTEM COST: LOW

READ SYSTEM LIFETIME: LOW BECAUSE
NO STANDARDS

COST/BIT: LOW

More details regarding archival storage media are presented in Section 6.0 of this report.

8.4.4 User Survey Results

A presentation was made on the preliminary results of the user survey. The detailed results of the survey are given in Section 6.0 of this report.

8.5 Discipline Sub-Group Discussions

Meeting attendees were asked to sign up for afternoon participation in one of four disciplinary sub-groups listed above. As it turned out, since only one person signed up for the Renewable Resources Sub-Group, that person was asked to join the Environmental Impact Sub-Group. Those who participated in each sub-group are listed in Table 8.7.

Each sub-group was moderated by an EarthSat speaker who had given a morning presentation on the status of the archive study. Sub-group members were asked to provide their views concerning archive data requirements for their particular disciplinary specialty. The same questions asked of the survey participants were presented to the sub-group for discussion and individual response. Comments were varied, but valuable, and some caused the project investigators to go back and give further thought to several issues. The results of this re-thinking have been incorporated into the text, conclusions, and recommendations of this report.

TABLE 8.7

DISCIPLINARY SUB-GROUP PARTICIPANTS

Sub-Group	Topics	Participants
Earth Sciences	Geology Geophysics Soils Geobotany Glaciology	Bryan Bailey Douglas Carter David Hastings Robert Ludwig Nancy Milton Orville Russell
Environmental Impact/ Renewable Resources	Forestry Agriculture Rangeland Wetlands Water Pollution Land Use Desertification Deforestation	Nancy Anderson Jim Durana Donald Garofalo Dean Gesch Graeme McCluggage Ed Partington Joe Sierra Donald Wiesnet
Archival Questions	Storage Retrieval Cataloging Browse Facility	Larry Carver Mike Cosentino William Cunliffe Mike Douglass Allan E. Hinkle Jim Love Mike Miller Christine Myers Ralph Poste Gil Richardson Jim Riorden Charles Sheffield Jay Sircar Locke Stuart Glen Trochelman Vickie Williams Richard Wood

Some of the key concerns/comments resulting from the sub-group discussions were:

- there should be a concentration on critical areas for archive data acquisition, e.g., we might use the National Science Foundation's Long-Term Ecological Research Sites (LTERS) as a guideline; certain sites should be selected and monitored because of their representativeness of a certain problem, e.g., particularly deforestation or desertification areas.

- because there are several data sets in existence or planned for satellite data, we should be looking at the networking of data repositories in order to permit access of such data in a coordinated manner, i.e., how does a potential user identify what is available and access this information? We need a central satellite information service.

- the costs of archiving the data should not be a factor affecting the decision of what should be collected; what should be collected should be determined independently of cost; whether or not it is possible to pay for the archiving is a budget/political issue which should be addressed separately.

- at a minimum we should be concentrating on the collection of at least one-time complete worldwide land coverage; many data gaps still exist for even one-time MSS and TM coverage.

- users should have the opportunity to provide continuing input to the definition of the Basic Data Set.

9.0 STRATEGY FOR DEFINING AND IMPLEMENTING THE BASIC DATA SET

9.1 General Comments

The user survey and the public meeting performed on this project provided the overall logical framework and guidance for the definition of the Basic Data Set used in this study; however, the specific logic for Basic Data Set definition, now and in the future, remains to be defined. That problem is addressed in this section, under the following working assumptions:

(1) The NSLRSDA will have a long lifetime, at least 25 years and probably a century or more. Thus any logic for Basic Data Set definition must not be constrained by today's spaceborne system capabilities, nor, to a lesser extent, by today's computer storage and retrieval systems.

(2) The geographic areas and disciplines of primary archival interest will also change over time, in ways that cannot be predicted today.

(3) The funding levels for data acquisition by the NSLRSDA are unknown.

(4) The NSLRSDA is expected to begin operation in 1989.

Taken together, these four working assumptions make it clear that on the one hand a selection logic for the Basic Data Set is needed that can be applied over a long (tens of years) time frame; while on the other hand, some very specific guidelines must be given for the initial data acquisition strategy.

In Sections 9.2 - 9.5, the general logic for future acquisitions is presented. In Section 9.6, the near-term data acquisitions are specified, reflecting today's perceived priorities, together with the design of a data reference system and logic that will generate the actual locations, timing, and priority of particular scene acquisitions.

We regard the logic for Basic Data Set acquisition as of central importance to the successful functioning of the NSLRSDA, and we urge careful consideration of the procedures suggested in this section.

9.2 Recommendations for Worldwide Coverage

As a result of the telephone survey, face-to-face contacts, and information received during the February 3, 1987 public meeting, it was generally acknowledged that collection of satellite land remote sensing data to meet fully the requirements of all parties would involve almost complete global data coverage, preferably seasonally, and at an annual repeat frequency. If such collection exceeds available budgets, an order of priority of data acquisition must be established.

One of the objectives of this study is to make recommendations for the establishment of a data archive that includes identification of the appropriate sensors for data collection, as well as a data collection strategy based on worldwide coverage. We must make certain assumptions in order to develop a reasonable data collection strategy for the NSLRSDA.

- Assumption 1 - a budget for purchase of data to be included in the archive would be limited, and thus quantities of data to be purchased from various systems for the archive are also limited.

- Assumption 2 - purchase of data for complete global coverage, seasonally, and on an annual basis from current and proposed high-resolution (higher than 80 meters) systems, would exceed any reasonably proposed archive budget.

- Assumption 3 - there are areas throughout the world which, because of pressures from man-made or natural processes, are experiencing changes. These changes may have a significant and possibly adverse effect on both regional and global habitability. Over the long-term operation envisioned for this archive, areas of rapid change will be added to the list and others deleted.

- Assumption 4 - the collection of data for the NSLRSDA from satellite-based sensors should focus first on land areas of the world identified by scientists as experiencing pressures which could affect regional and global habitability, while recognizing that additional new areas will be identified in the future.

- Assumption 5 - there are and will be other phenomena, perhaps not of a critical nature, which have important scientific or economic value, the study of which would be of benefit to mankind. These phenomena should receive data coverage in the NSLRSDA.

- Assumption 6 - complete global coverage with data from existing land satellite systems, identified in this study as appropriate sources for the archive, does not currently exist in the U.S., even if a mix of data types is admitted; initial efforts at archive data collection should strive at a minimum for a complete global data set.

- Assumption 7 - data volume per unit area from higher spatial and spectral resolution systems, as well as data volume for larger area coverage, increases substantially over lower resolution systems; this suggests that collection of archive data from higher resolution systems must be done on a more selective basis than for lower resolution systems in order to meet archive budget limitations.

- Assumption 8 - requirements for archive data will change periodically, and critical areas/phenomena of the world as perceived today will change and new ones will be identified.

These assumptions suggest that a practical way to acquire satellite remote sensing data for the archive is to adopt at the outset a reasonable archive data collection strategy, which takes into consideration data types and volumes of the varying satellite systems and balances these against data requirements. A priority system should be employed which ranks areas to be covered by the system and which allows flexibility within the context of the existing budget for data acquisition. The amount of funds available in the acquisition budget should not be the driving force in deciding what data should and should not be acquired for the NSLRSDA. The priority ranking of data to be collected should be done independently of the budget, although it is recognized that the budget will limit how many of the priorities will be realized during a particular budget cycle.

9.3 Creation of the Archive Data Selection Committee

Selection of data to be incorporated into the archive should be done by an Archive Data Selection Committee (ADSC). The committee should meet at least annually to determine data requirements for the following year, and also have the capability to make immediate decisions on short notice regarding special events. The committee should be multidisciplinary in make-up, comprising physical and biological scientists. It may be worthwhile to have international membership on the committee. There should be an opportunity to "rejuvenate" the committee with new members at periodic intervals so that fresh perspectives are added.

The committee should draw from a variety of sources of information in order to make its selection of archive data. For example, the National Science Foundation's Long Term Ecological Research Sites, or the World Resources Institute's reports which assess the resource base that supports the global economy, should be referred to by the committee. These and other sources identify critical resources throughout the world, and provide useful geographical information in the form of maps and descriptions.

Because the proposed committee will be making current decisions, e.g., based on existing information about known areas of interest around the world, it can be very specific about required data parameters. For example, only one season coverage of a phenomenon may be required in order to provide effective monitoring, or only the most dynamic or representative portion of a larger area of interest may be needed. A committee decision, therefore, has a capacity to reduce large volumes of data which might otherwise be requested by a less flexible, one-time data collection strategy.

Below we describe a sample archive data collection scenario based on the proposed committee approach. First, however, there are elements of the scenario which may be employed using a non-committee approach.

9.4 Non-Committee Activities

The Basic Data Set should at an absolute minimum include a one-time, cloud-free coverage of every land area of the world, preferably close to concurrent in acquisition date. Gaps still exist. The NSLRSDA should strive to fill these gaps with high-resolution, e.g., equal to or greater than 80 meter, data as they become available, and to seek an updated global land coverage when this is feasible, at some suitable time interval (to be determined).

The low-resolution, large-area coverage systems such as AVHRR should collect archive data on a seasonal basis with an annual repeat frequency.

9.5 Committee Activities

Recognizing that there are multiple sources of information to which the data selection committee can refer, we offer as example the annual report of the World Resources Institute, identifying critical resources throughout the world. The areas of critical impact have not been ranked in order of importance by the WRI. However, a methodology can be employed by the Archive Data Selection Committee which looks at multiple overlapping parameters as a way of ranking the most to least important areas for data collection. Parameters might include:

- rate of change occurring
- size of population being impacted (direct/indirect)
- economic value of the resource
- type of natural resource being impacted (trees, people, etc.)
- how resource is being impacted (population, disease, etc.)
- capacity to be monitored using representative sampling
- political volatility of resource or area
- type of impact (immediate, long-term, permanent, temporary)
- other biological/physical resources being impacted
- scientific interest (non-critical)
- impact on adjacent areas
- area of resource being impacted
- location of resource being impacted

If we then employ a rating system whereby each parameter is assigned a rank from 0 to 3, with 0 being the lowest rank (phenomenon not present in this geographic location) and 3 the highest, we can develop a data collection scheme. This is done explicitly in Section 9.6, with some added assumptions about selection logic. In addition to the parameters listed above, data requirements for looking at a specific phenomena must also be included in the selection process, such as:

- o Spatial resolution
- o Spectral resolution
- o Repeat frequency
- o Seasonal
- o Cloud cover

Table 9.1 illustrates how the data selection scheme might look. A numerical ranking is derived. The higher the numerical ranking, the higher the priority for selecting data of a given area or phenomenon.

For areas or phenomena considered to be of significant scientific or economic importance, we recommend that the Archive Data Selection Committee accept nominations for data collection from the data user community, without initially limiting the number of such nominations. The user community would be advised that only a limited number of nominations might be accepted reflecting budget limitations, and the Committee would rank these requests using a similar ranking method as proposed in Table 9.1. The ADSC would be instructed to spread their selection of requests throughout the user community. In this way data users would have an equal opportunity to have their requests accommodated, if not in the current year's quota, then in subsequent years. The process described in Section 9.6 provides an automated method of adjusting acquisition priorities for particular data needs.

9.6 Systematic Approach to Data Acquisition

In order to describe phenomena in unique geographic terms, it is necessary to adopt some standard geographic reference frame. It should be noted at once that the geographic reference frame is the technique by which data of the NSLRSDA are related to each other and to the Earth's

Table 9.1: Proposed Archive Data Selection Scheme for Ranking Key Resource Areas

Phenomenon Being Covered: _

Geographic Location (bracket): lat/long ___/___; lat/long ___/___
lat/long ___/___; lat/long ___/___

Spatial Resolution Required:__

Spectral Resolution Required:__

Repeat Frequency Required:__

Seasonal Requirements:__

Cloud Cover Requirements:__

Circle one value for each parameter:	0*	1	2	3
Rate of Change Occurring	0 low	1	2	3 high
Economic Value of Resource	0 low	1	2	3 high
Suitable for Representative Sampling	0 low	1	2	3 high
Political Volatility of Resource	0 low	1	2	3 high
Political Volatility of Area	0 low	1	2	3 high
Impact on Adjacent Areas	0 low	1	2	3 high
Area of Resource Being Impacted	0 small	1	2	3 large

*A zero value indicates that the phenomenon is not present in this geographic location, or is of negligible importance.

Type of Impact	0 temporary	1	2	3 permanent
Scientific Event Interest	0 low	1	2	3 high
Other Resources Being Impacted	0 n/critical	1	2	3 critical
Location of Impacted Resource	0 n/critical	1	2	3 critical
Type of Resource Impacted	0 n/critical	1	2	3 critical
Impacted Population (direct)	0 low	1	2	3 high
Impacted Population (indirect)	0 low	1	2	3 high

SUBTOTALS =

GRAND TOTAL _____ = IMPORTANCE FACTOR

RANK _____

Data Sources Used:

____

____

Data Selection Committee Members:

_________________________ _________________________

_________________________ _________________________

_________________________ _________________________

_________________________ _________________________

Date of Evaluation: _

Committee Chairman: _

signature

surface. The data files of the NSLRSDA are not converted to the geographic reference frame. They are stored in the most convenient and economical frame, which will usually be the form employed in the individual collection systems.

For example, Landsat data are stored by path, row and date. However, such a reference system is not satisfactory as the general geographic reference frame, for a number of reasons. First, it is not truly global, since Landsat spacecraft do not image the poles, and their path/row system reflects that fact. Second, path/row designations are different for different spacecraft (Landsat-1,- 2 and -3 use a different system than Landsat-4 and -5), and will certainly not apply to SPOT, MOS and other future spacecraft. Third, the Landsat path/rows overlap a little at the equator but a great deal at high latitudes. Thus a particular geographic location can be identified as within several different path/row cells. Finally, path/row cells defined for Landsat may be too large to apply to future generations of spacecraft, with their possible increased spatial and spectral resolution. The Landsat geographic labeling system is therefore not suited to serve as a general geographic reference frame, and some different approach must be adopted.

9.6.1 Geographic Reference Frame

Geographic information systems for the storage and manipulation of map and resource data usually employ one of two different approaches to data management. Either they perceive map data as comprising a set of polygons that cover the whole area; or they divide that area into a set of non-overlapping, rectangular, grid cells of equal area.

Both methods have advantages and disadvantages. The polygon approach permits variable resolution, so that a region with much geographic diversity can be economically stored; on the other hand, combination of many maps is logically difficult using a polygon approach. The grid cell approach takes no advantage of variations in information resolution across a scene, and may therefore require a very large number of cells, corresponding to the highest resolution scene element to be defined. On the other hand, combination of information from many maps is easy with the grid cell approach.

In the following discussion, the grid cell approach is adopted. It permits a simpler approach to the problem of defining a data acquisition logic and for combining information about many phenomena. Most geographic information systems in use today permit polygon-to-cell and cell-to-polygon conversion, thus the actual structure of the reference frame ought to be transparent to the user.

It is proposed that, for NSLRSDA purposes, the surface of the Earth be divided into a set of non-overlapping, approximately rectangular, equal area cells. Although current spaceborne sensors for land remote sensing are limited to the range 81° North to 81° South, for completeness this cell system will be defined over the whole Earth.

An appropriate cell size for the geographic reference frame is suggested as about 100 km x 100 km. (This exceeds the size of a single SPOT HRV image, which is 60 km x 60 km.) Modest increases in spatial resolution (say, to 10 meters) and spectral resolution (say, to 12 bands) imply that a single reference frame cell for such a sensor would correspond to 10 gigabytes of storage in the Basic Data Set.

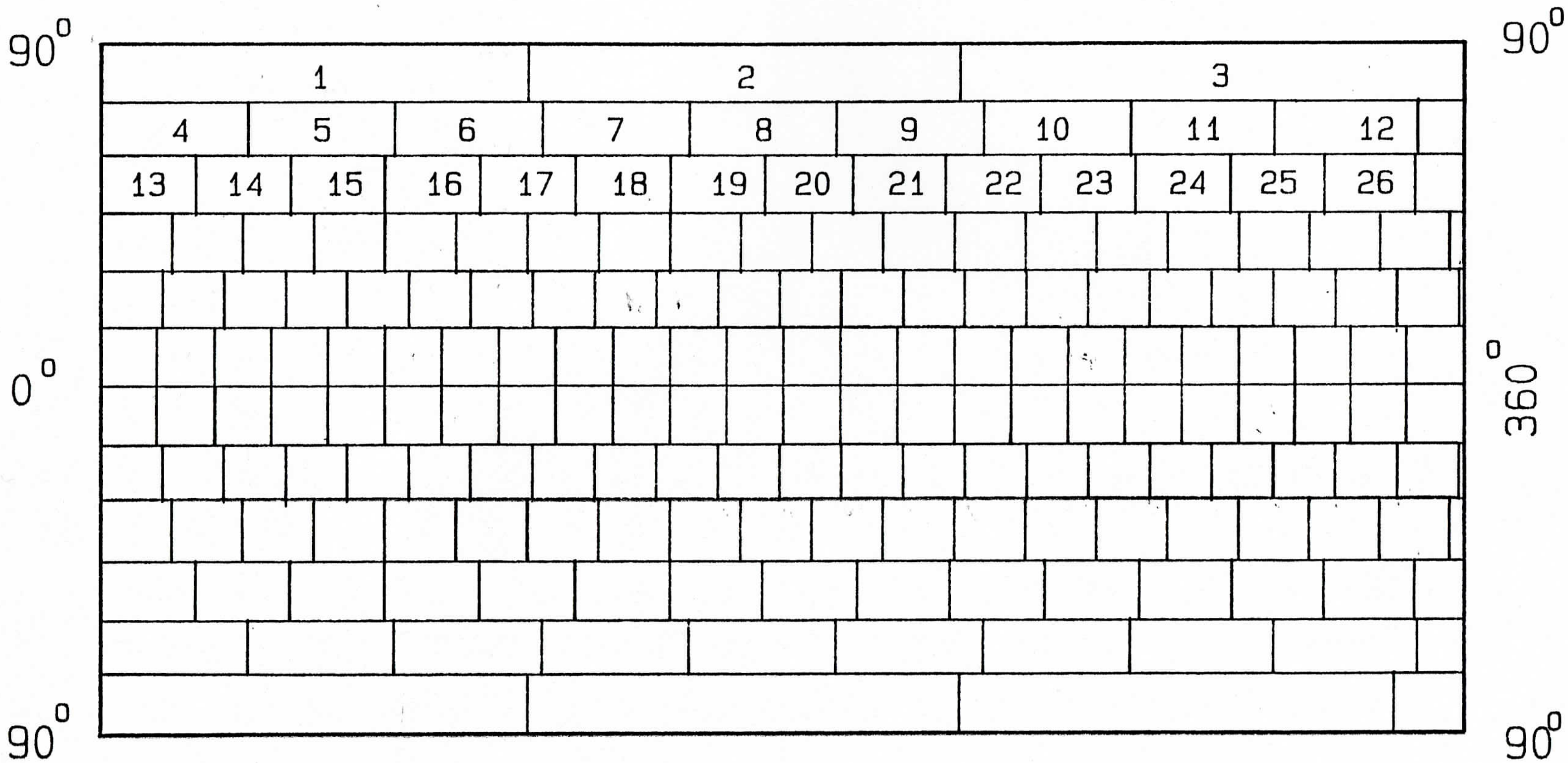
Three factors suggest that it would be unwise to decrease geographic reference frame cell size too far. First, storage densities of machine-readable media continue to increase, so that 10 gigabytes on one tape will not be unreasonable in a few years time. Second, one important characteristic of space remotely sensed data is its broad, synoptic coverage. Third, it is well known that no mapping of the globe to a rectangular array of cells can be exact. There will be distortion at the cell edges.

Although the interest of the NSLRSDA is in land remote sensing, it is simpler to define the geographic reference frame network to cover both land and sea. One simple and unambiguous way of defining a reference frame cell structure would be to proceed westward from the Prime Meridian, cell-by-cell, until the globe has been encompassed. Spacing such rings of cells systematically north and south of the equator leads to a non-overlapping set of cells that completely cover the globe. An individual cell of the geographic reference frame then has a unique latitude and longitude (that of its center point) and can be assigned a unique cell number within the whole reference frame (see illustration of Figure 9.1). The choice of projection used in the geographic reference frame is not important. The illustration shows a Mercator projection. Sinusoid or Eckert equal area map projections are equally good candidates.

Figure 9.1

SAMPLE GRID CELL LAYOUT
AS IT WOULD APPEAR IN A
MERCATOR PROJECTION

(Cell Size Shown Greatly Enlarged)



Note: Proposed grid cell numbering system is "pagewise ordering".

With a cell 100 km x 100 km, about 50,000 cells are employed in defining the geographic reference frame. Latitudes and longitudes are the key variables that relate the cells of the geographic reference frame to the contents of the Basic Data Set.

9.6.2 Defining Overlays for Resource Phenomena

In the NSLRSDA, a phenomenon may be fixed and localized, such as the Hubbard Glacier, or variable and very large in extent, such as acid rain in boreal forests. Many of the phenomena of interest to the NSLRSDA are identified in Section 7.3, but it is certain that others will be added as the archive develops.

Each phenomenon identified as of interest to the NSLRSDA gives rise to its own phenomenological overlay. A phenomenological overlay is a complete set of geographic grid cells, with an importance factor assigned for that particular phenomenon for every cell. The importance factor is that defined in Table 9.1 of Section 9.5. A cell receives a factor of zero if the phenomenon is absent, and a factor from 1 (little importance) to 5 (very important) when the phenomenon is present in that grid cell. A complete overlay for a phenomenon requires that the importance factor be given in every cell for the whole world. However, for any given phenomenon large numbers of zero cell entries are to be expected. Figures 9.2 to 9.4 exemplify the general appearance of sample phenomenological overlays anticipated in the NSLRSDA, shown here in the way it would be addressed in either a cell or polygon approach.

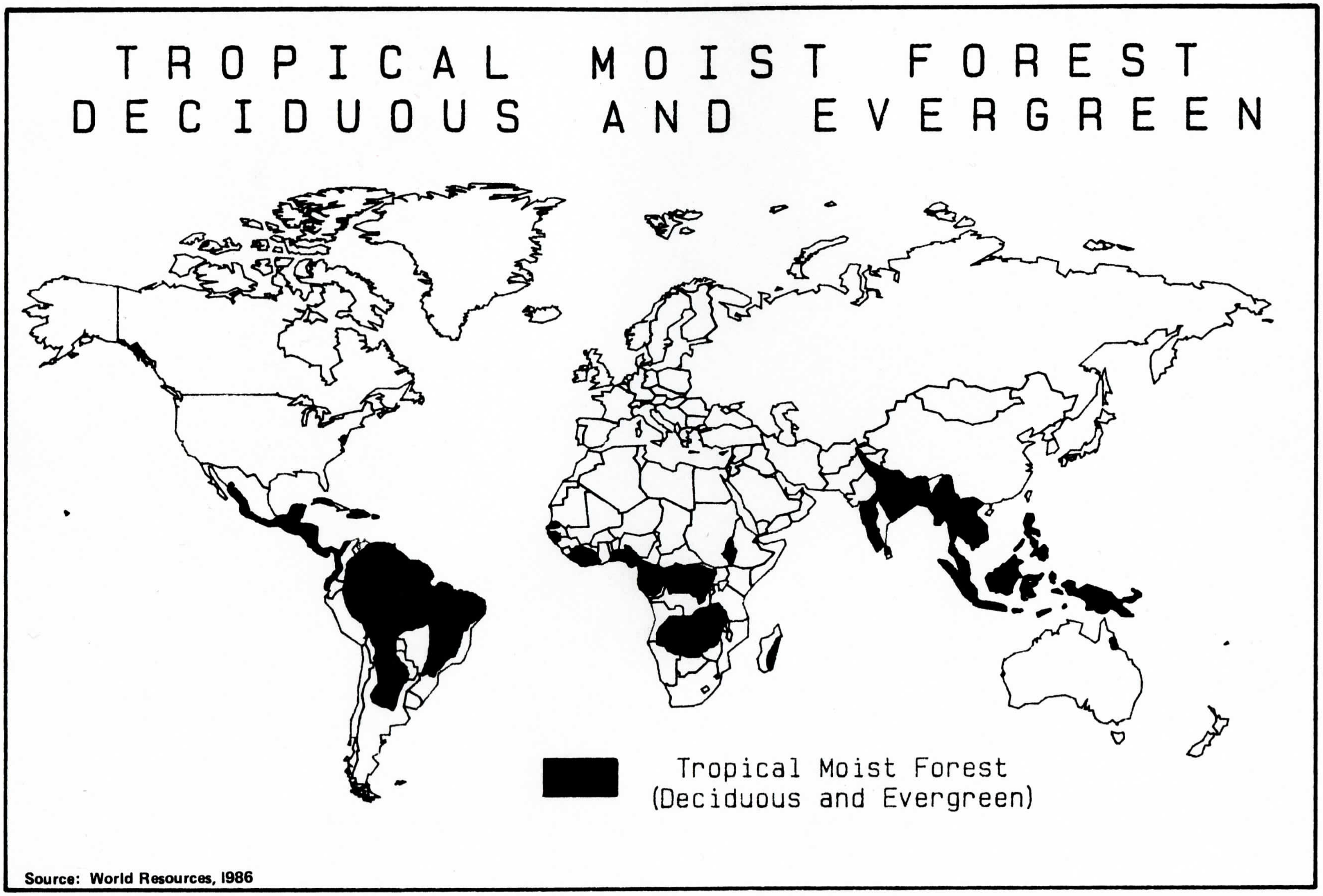
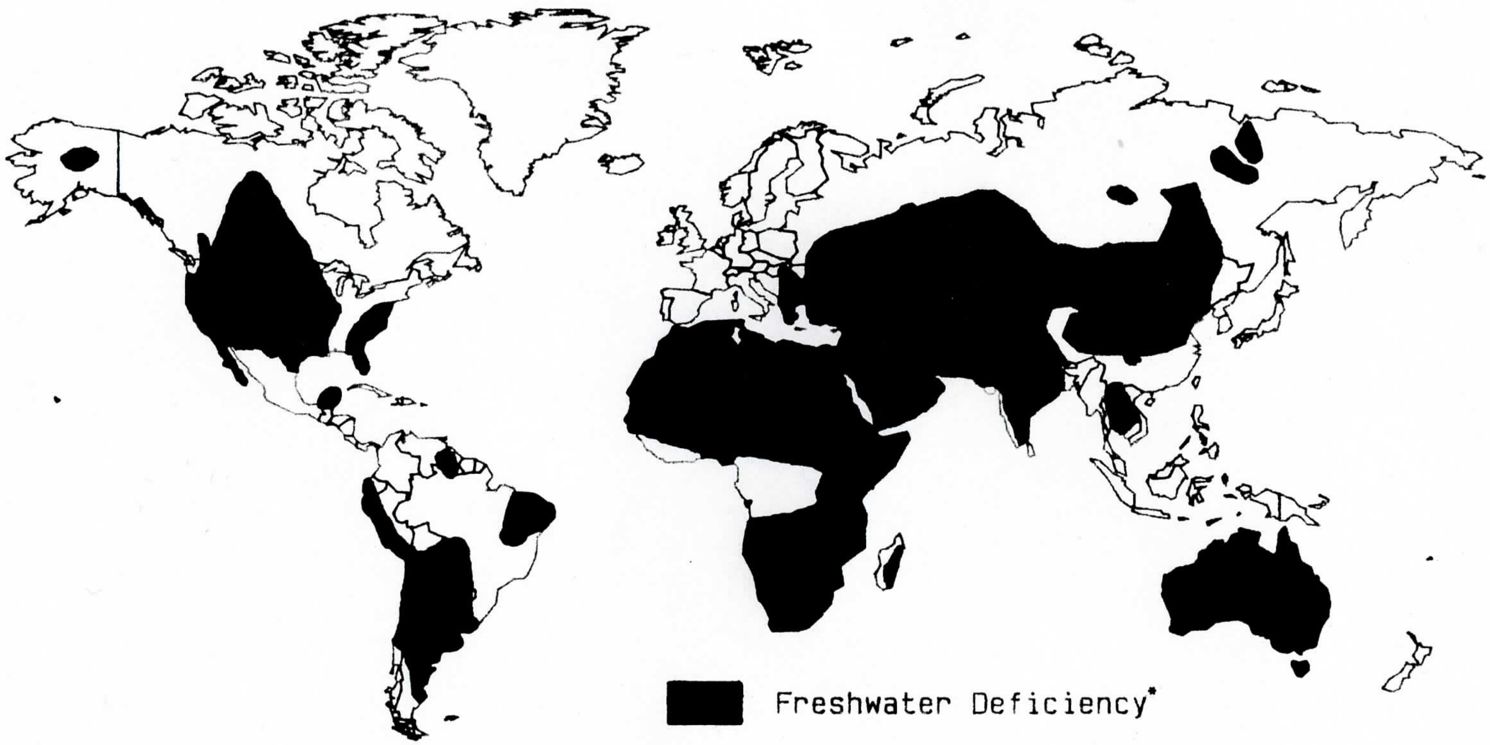

TROPICAL MOIST FOREST
DECIDUOUS AND EVERGREEN
Tropical Moist Forest
(Deciduous and Evergreen)
Source: World Resources, 1986

Figure 9.2

FRESHWATER DEFICIENCY



*Defined as difference between annual precipitation and evapotranspiration. Evapotranspiration corresponds to the water demand from a potential crop not suffering from water deficiency. Source: World Resources, 1986.

Figure 9.3

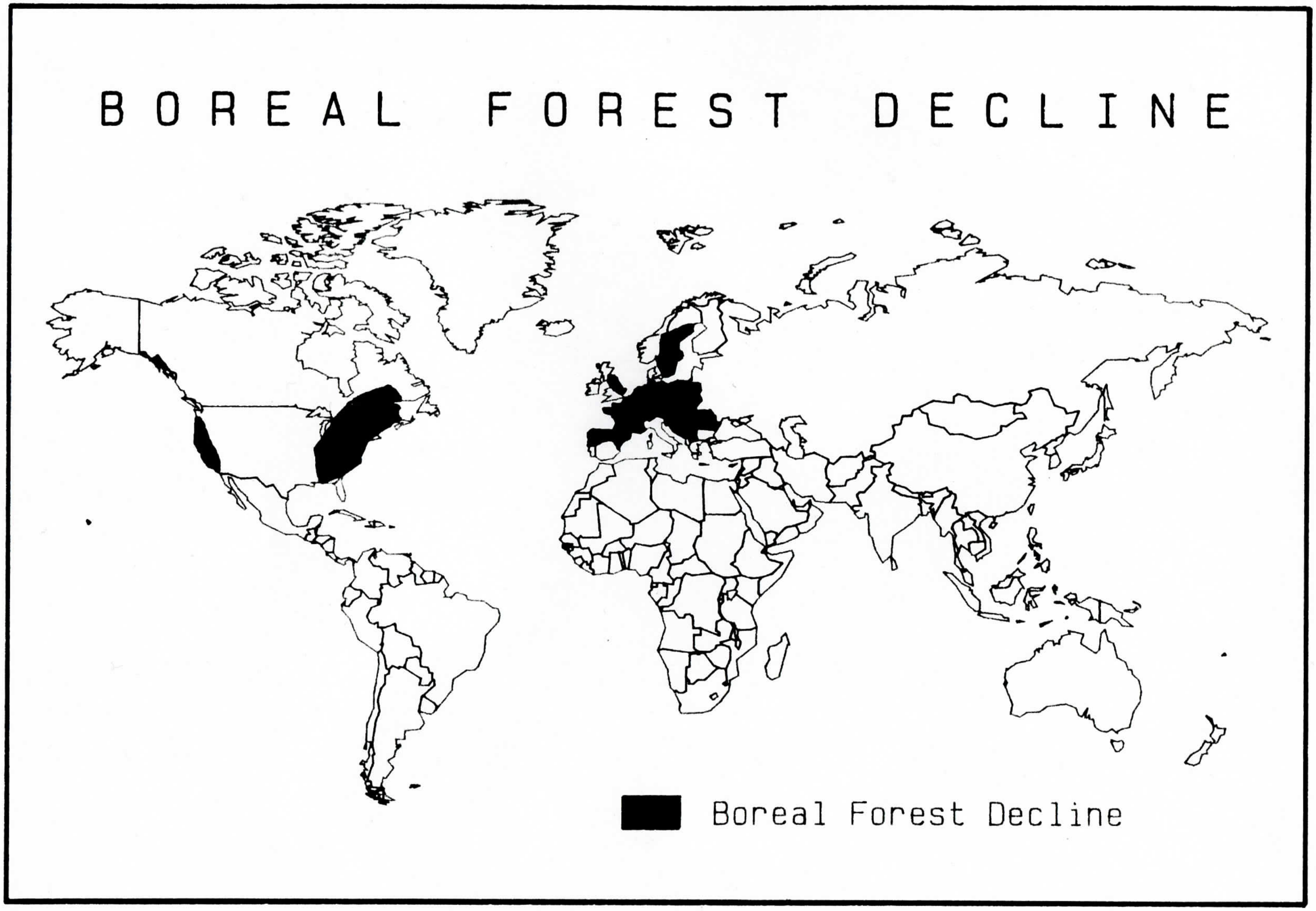


Figure 9.4

If the number of phenomena are of the order of 30, so that there are 30 overlays on the global cell network, about two million entries are needed in the data base; however, these entries can be generated simply and quickly by automatic digitizing from maps.

9.6.3 Defining Data Collection Priorities

Each phenomenon of interest to the NSLRSDA will have associated with it one or more preferred data sources (e.g., TM, SPOT, AVHRR, etc.), together with a desired collection frequency (seasonal, annual, every decade, etc.). For each sensor, we can now, for each grid cell and each satellite data source, calculate a total computed "score." This score is the weighted sum of the assigned importance factors over all phenomena, times a priority factor, that indicates if collection for the phenomenon is overdue. The process of computing the scores for each grid cell is illustrated in Figure 9.5. Note that the analogous procedure, should a polygon-based geographic reference system be used, would become extremely complicated, and perhaps impossible.

The strategy for data acquisition is now a straightforward one: whenever possible, data are acquired first for the cells with the highest scores.

Reference was made in the previous paragraphs to the weighted sum of importance factors. If all phenomena were judged equally important, then all the weights would be equal. By permitting unequal weights to be used, the system allows the Archive Data

Figure 9.5: Collection Priority Calculation: An Example

Data collection system: Landsat TM

Phenomenon	Phenomena Weights*	Cell # i	Cell # j	Cell # k
Acid Rain	10	0** 0#	5 1	3 4
Deforestation	3	5 2	0 0	2 2
Flood Analysis	6	1 3	3 4	3 9
Urban Change	5	0 0	0 0	4 6
Soil Erosion	7	5 10	3 1	2 4
SCORE = Sum of products of importance factor times priority factor times phenomena weights over all phenomena. Highest scores cells are where possible collected first.		398	143	470

*Assigned by Selection Committee.
**Importance factor for each phenomenon in each cell.
#Priority factor for each phenomenon in each cell (reflecting the elapsed time since data last collected versus collection requirement).

Selection Committee to make subjective judgments, varying from year to year, of the relative importance of acquiring information on particular phenomena at particular times. The logical flow of the decision-making process is illustrated in Figure 9.6.

9.6.4 System Use and Update

When a new phenomenon becomes of interest for NSLRSDA acquisition, a new overlay must be generated appropriate to it. Note that the geographic frame (i.e., the grid cell layout) is the same for all overlays, and never varies. The new overlay is created by entering the importance factor for each sensor in each grid cell. Usually, the grid cell entries will be generated by first creating a global map of the phenomenon, and digitizing it through a technique either of raster scans or polygon delineation. Both methods lend themselves well to automated methods, and the addition of new phenomenological overlays is not perceived as a major problem.

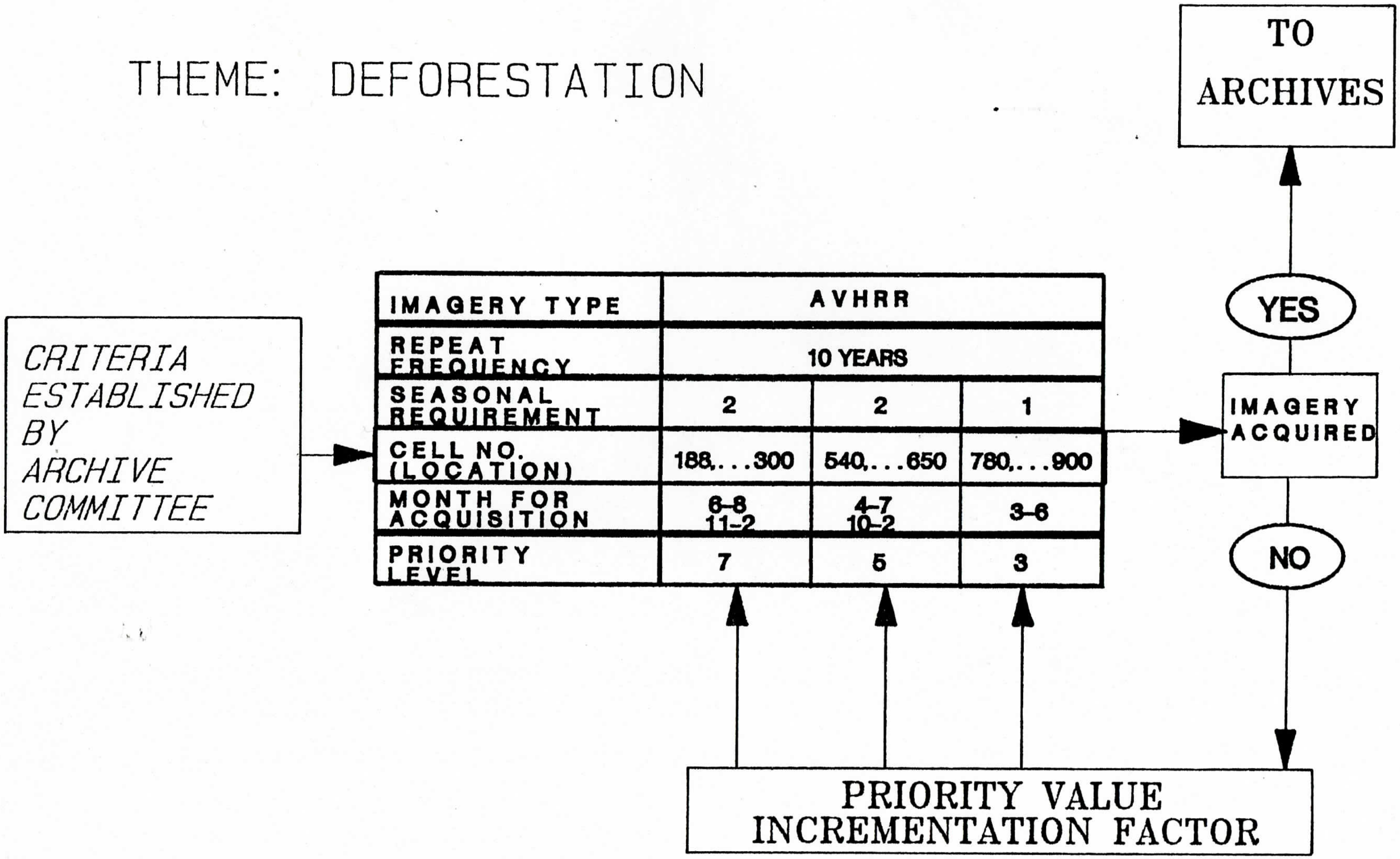
Overlays and overlay weight factors must be periodically reviewed by the Acquisition Selection Committee (at least annually is recommended) and either approved for the next collection period, or updated with changes and additions.

9.6.5 Catastrophic Events

Certain natural events are unpredictable and short-term, but locally can have a catastrophic impact on the physical condition of the earth surface and/or cover. Such events can include major

Figure 9.6

Example of Data Collection Decision Logic



IMAGERY TYPE	AVHRR		
REPEAT FREQUENCY	10 YEARS		
SEASONAL REQUIREMENT	2	2	1
CELL NO. (LOCATION)	188,...300	540,...650	780,...900
MONTH FOR ACQUISITION	6-8 11-2	4-7 10-2	3-6
PRIORITY LEVEL	7	5	3

floods, volcanic eruptions, and perhaps, droughts and fires. Such events cannot be factored into the normal imagery selection process and should be dealt with on a case-by-case basis to judge whether or not the event is of adequate historic or scientific interest to justify acquisition of imagery for archival purposes. Archive management should have the prerogative to make such decisions without the convening of the ADSC group.

10.0 CONCLUSIONS AND RECOMMENDATIONS

The following conclusions and recommendations are divided into two categories: Organizational and Technical. Items which should hold high priority for action are indicated with an asterisk (*).

10.1 Organizational Elements

(1) Conclusion

The NSLRSDA is a new and unique enterprise that must draw on the experience of both the remote sensing community and the data archiving community. Expertise in both areas is essential to the creation and operation of a successful archive.

*Recommendation

An overall Steering Group, with members drawn from the remote sensing and archival communities, and with a leadership and a secretariat from the Commerce Department, should be created. Such a group is a natural outgrowth from the existing ad hoc group formed by NOAA and USGS.

(2) Conclusion

Great interest has been displayed by scientists and technicians involved in remotely sensed data in the NSLRSDA. However, there is concern that the value of such an enterprise will be seriously diminished unless appropriate data are collected and archived.

*Recommendation

The Steering Group should appoint as a sub-group a committee of scientists and data users, with primary responsibility for defining archival data acquisition. This group, which will be termed the Archive Data Selection Committee, should be formed as soon as possible and have permanent existence.

(3) Conclusion

A substantial fraction of the satellite land remotely sensed data acquired in the past and being acquired today exists only in foreign data bases. Systematic procedures to assure the continued existence of such data, or to transfer them to a U.S. archive, have not been defined.

Recommendation

A detailed inventory of the holdings, formats and condition of foreign satellite land remotely sensed data bases should be performed. Where appropriate, bilateral agreements for data transfer to the U.S. archive should be formulated and implemented.

(4) Conclusion

The NSLRSDA is not a static entity. It will be highly dynamic in both data sources and available archive media. Substantial changes in the satellites providing data, anticipated uses of the data, and archiving media, must be anticipated over the next few decades.

Recommendation

The archive structure must be a flexible system, and must include mechanisms to change acquisition priorities, storage technology, and user service technology over time. An Archive's Structure Group (ASG) should be created to serve under the Steering Group. The ASG should be composed of archiving and data base management specialists.

(5) Conclusion

There exists today in the United States considerable expertise and competence in the organization, creation, operation, and maintenance of large data bases of satellite land remotely sensed data. The principal center of expertise and experience in this field since the beginning of the U.S. Landsat program has been the EROS Data Center.

Recommendation

The EROS Data Center at Sioux Falls, South Dakota, has already been designated as the official location and responsible body for the NSLRSDA. However, additional facilities, including processing equipment, data quality evaluation equipment, storage facilities, and personnel, should be added to the EROS Data Center to permit the development of the NSLRSDA, and to allow the provision of back-up data off-site.

10.2 Technical Elements

(1) Conclusion

Other repositories of satellite land remotely sensed data exist today and will continue to exist. It is not feasible for the NSLRSDA to substitute for such data sets, nor is it feasible for the NSLRSDA to duplicate entirely the contents of such data sets, or to impose upon them the sampling and copying requirements of the National Archives.

Recommendation

The NSLRSDA should hold only the Basic Data Set, which will be subject to National Archives' sampling and copying requirements. Other data bases will be referenced in the NSLRSDA at the catalog level, but will not include browse facilities for other data bases.

(2) Conclusion

Satellite land remotely sensed data are used increasingly in digital form. The need to provide accurate radiometric information on stored data, and the need to permit data to be merged with other digital files, is also increasing. At the same time, large quantities of historical land remotely sensed data exist only in film format.

Recommendation

Digital storage of new data acquired for the Basic Data Set is essential, and the primary storage mechanism for the NSLRSDA must be in machinereadable digital form. Data must be available to archive users in machine-readable digital form. However, data available today and stored only on film should be held in this form in the archive, without seeking to redigitize them.

(3) Conclusion

Since a principal use of the NSLRSDA data will be to permit the comparison and evaluation over time of regions of the Earth's surface, information must be available to the archive on the prior processing performed on digital data. In particular, quantitative comparisons should be facilitated.

Recommendation

The NSLRSDA should, in addition to its data files, seek to store complete written descriptions and other documentation of processing performed, calibration testing, evaluation studies, and other significant quantitative descriptions of data stored in the archives. The library of such information should be considered a significant holding of the NSLRSDA.

(4) Conclusion

The problem of assessing priorities for data collection, and determining what data will be acquired by particular sensors for the NSLRSDA, is a major task and one that should be undertaken before the NSLRSDA begins operation.

*Recommendation

A computerized global reference system should be created in the near future, permitting updates of data acquisition needs, from each spaceborne sensor. Control of the reference system, and assignment of priorities for entry to it, will be by the Archive Data Selection Committee.

(5) Conclusion

Data from the NSLRSDA will be used in numerous different applications, and requested in numerous different formats; however, the cost of providing such flexible service would be considerable.

Recommendation

The NSLRSDA should adopt few and fixed formats for provision of data, leaving the problem of integrating archival data with other data types, or otherwise reconfiguring archival data, to the final users. It is suggested that for each data type a single machine-readable output format be provided, with a small number of output array sizes, and that film product outputs be available on a pre-defined and limited number of scales.

11.0 SOURCES AND REFERENCES

Adelstein, P.Z. and McCrea, J.L., 1981. Stability of processed polyester base photographic films. Journal of Appl. Photogr. Eng., 7 (6 December): 160-166.

Baer, N.S. and Banks, P.N., 1985. Particulate Standards for museums, libraries and archives. 75th Ann. Meet. Air Pollut. Control Assoc. Preprint 85-88.

British Petroleum Statistical Review of World Energy, 1985.

Brown, D.W., Lowry, R.F., and Smith, L.E., 1984, Predictions of Long-Term Stability of Polyester Based Recording Media, National Bureau of Standards, NBSIR 84-2988.

Brown, L.R., et al., 1984, 1985, 1986, State of the World, Worldwatch Institute Reports, Norton & Company, New York.

Bulloch, Chris, 1986. Monitoring Global Change: Multidisciplinary Earth Studies to Depend on Spacecraft: SPACE MARKETS, #4, p.196-200.

Caron, Loyola M., Debora S. Steward, 1984, An Inventory of State Natural Resources Information Systems, Final Report, Kansas Applied Remote Sensing Program, University of Kansas.

Carter, V., and Dale, T., 1974, Topsoil and Civilization, University of Oklahoma Press.

Chadwick, M., and Lindman, N., 1982, Environmental Implications of Expanded Coal Utilization, Pergamon Press.

Chen, H.S., 1985, Space Remote Sensing Systems: An Introduction, Academic Press, Inc., New York, NY.

Eastman Kodak Company, 1985. Storage and Care of Kodak Films and Papers - Before and After Processing. Pamphlet E-30.

Eastman Kodak Company, 1971. Kodak Color Films, Data Book No. E-77.

EOSAT, personal communications, 1986; The Earth Observation Satellite Company, Lanham, MD.

EROS Data Center, Meeting in Sioux Falls, December 1986 (personal communications and provided data).

Forest Damage Surveys, 1985, Federal Ministry of Food, Agriculture and Forestry, Bonn, West Germany.

Greenwood, D., 1985, Mapping: The University of Chicago Press, Chicago, 289 p.

Landsat Data Users Notes, Issue No. 55, March 1986, NOAA-Department of Commerce (GPO: 1981-564-078/9).

Light, Donald, 1986, "Mass storage estimates for the digital mapping era." In Photogrammetric Engineering and Remote Sensing, V. 52, No. 3, March 1986, pp. 419-425.

Nare, R.L., 1980, Water of the World, Government Printing Office.

National Research Council, 1977, Resource Sensing from Space: Prospects for Developing Countries: Report of the Ad Hoc Committee on Remote Sensing for Development. National Academy of Sciences, Washington, D.C., 202 p.

National Research Council, 1986, Preservation of Historical Records: National Academy Press, Washington, D.C.

NOAANET Functional Requirements, NOAA-National Environmental Satellite, Data, and Information Service, Data Acquisition Project Office, Washington, D.C., August 1986.

NOAA-Satellite Data Service Division, Suitland, MD, 1986 (personal communications).

Office of Technology Assessment, 1985, Acid Rain and Transported Air Pollutants.

Opitz, B.K. (editor), 1986, Geographic Information Systems in Government, Deepak Publishing, Hampton, VA.

Presidential Directive-54; October, 1980, Public Law 98-365; July 17, 1984.

Robinson, A., Randall, S., and Morrison, J., 1978, Elements of Cartography: 4th Edition, John Wiley & Sons, New York, 448 p.

Slater, Philip N., 1980, Remote Sensing Optics and Optical Systems, Addison-Wesley Publishing Company, Reading, MA.

Slater, Philip N., 1986, Survey of Multispectral Imaging Systems for Earth Observations, Remote Sensing Yearbook.

Spaceflight, International Magazine of Space and Astronautics, Vol. 29, No. 3, March 1987, British Interplanetary Society, 1987, London.

SPOT Image Corporation, Reston, VA, 1986 (personal communications).

SPOTLIGHT, Issues 1 and 4, 1986, SPOT Image Corporation, Reston, VA.

A Study of Alternatives for the Preservation and Reference Handling of the Pension, Bounty-Land, and Compiled Military Service Records in the National Archives, 1984, National Archives and Records Service, internal report.

Strome, M., 1986, Optical Tape as an Archival Medium, personal communication.

Tessav, Paul A., Loyola M. Caron, Jane Germano, Peggy Harwood, Susan Seladones, 1985, A Survey of State Applications of Investments in Landsat Technology: Natural Resource Information Systems Project, National Conference of State Legislatures, Denver, CO, and Earth Resources Data Council, Council of State Planning Agencies, Washington, D.C.

World Resources, 1986, International Institute for Environment and Development, Basic Books, NY.

Yearbook of Forest Products, 1985.