Managing Remotely Sensed Data and Information 2

By the end of this century, U.S. and foreign civilian remote sensing satellite systems will begin to generate huge volumes of data about the Earth on a daily basis. In order for weather and climate forecasters, researchers, resource managers, and other users to make the most efficient use of these data, the U.S. government will have to invest in new technologies for collecting, storing, distributing, and analyzing remotely sensed data. Private industry and government have already invested billions of dollars in a robust data and information infrastructure that can support government efforts to cope with the new data sources. **The information industry will greatly facilitate the rapid growth of a commercial and governmental market for information produced from satellite data.**

In the 1970s, when the operational environmental satellite systems and the Landsat system were first developed, data users had to rely on large, expensive mainframe computers to analyze the data. They further had to depend on the creation and delivery of data tapes and/or "hard copy" images from the central processing facilities. If users wished to browse through data files to select the best quality data, they would either have to depend on the judgement of personnel at the facility, or travel to the facility themselves and examine the data directly. Storage and archive facilities were highly limited and usually relied for data storage on thousands of paper copies, photographic images, or magnetic tapes. Working with the data meant physically retrieving archived data from storage by hand and copying them for each different user.

Since the 1970s, the rapid development of computer processing, storage, and data communications technologies has revolutionized the way data and information, including remotely sensed data, are treated. Today, for example, it is possible to examine sample data delivered online from the archive to the user's computer work station or PC, select the required data, and have the data subsequently delivered over the same data transmission lines. In many cases, billing for the data can be accomplished online as well.[1] These capabilities have come about as a result of the dramatic improvements that have occurred over the last decade in information technologies, changes that have led to the development and vigorous growth of a broad-based information industry.

This chapter provides an overview of the U.S. information industry and the technologies that support it. It explores the role these technologies play in the management and application of remotely sensed data. The chapter also summarizes Federal programs currently in place for archiving and distributing remotely sensed data and examines options for improving archive services.

THE INFORMATION INDUSTRY

Virtually all segments of modern industrial society use some form of data and information technology to improve efficiency and capability. Research, manufacturing, service industries, financial markets, and governance have all been affected by the growth of the electronic information industry. Modem computers and allied technologies make it possible to acquire, organize, store, update, and distribute large amounts of data and information for a wide variety of tasks. In large part, the information industry, which consists of manufacturers and sellers of computer hardware and software[2] and data storage and transmission equipment, as well as information services, will determine the nation's ability to manage and process remotely sensed data. Although information technologies are increasingly capable, the requirements for making large quantities of remotely sensed data available to diverse users will stress the capabilities of existing storage, processing, and transmission systems. Improving the ability to deliver remotely sensed data quickly and accurately will also require improved institutional arrangements.

Processing Data and Information

With the development of high-speed, integrated computer chips and other innovations, the ability to manipulate and analyze large amounts of data conveniently has improved dramatically over the last decade and a half. Personal computers and workstations with fast processing speeds, large amounts of storage, and random access memory adequate for rapid image processing have become standard.

Computer technologies are now much more broadly distributed now than they were just 10 years ago. Rather than working in a mainframe environment, most computer users now use personal computers and/or workstations, which may be linked electronically to mainframe databases, but are capable of running their own software applications independently.

Over the last three decades, computer hardware performance per unit cost has increased by a factor of a million,[3] and has led to rapid growth of digital processing capabilities. During the 1980s and early 1990s, the cost of computing capability dropped substantially—it costs less today *in* 1994 *dollars* to purchase a personal computer based on a 486 processing chip operating at processing speeds of 50 Megahertz (MHz) than it did in 1985 to purchase a system based on a 286 chip operating at 8 MHz.[4] Similar improvements have been

[1] Data fees vary according to the service provided, but for data supplied by government agencies, prices charged, if any, generally reflect the marginal cost of fulfilling a user's need.

[2] Computer hardware and peripherals accounted for sales of at least $65 billion in 1993. See U.S. Department Of Commerce, international Trade Administration, *U.S. Industrial Outlook 1993* (Washington, DC: U.S. Government Printing Office, 1993), p.26 -1.

[3] Malcolm Brown, "The March of the Mighty Chip," *Management Today*, quoting Andrew Sayer of Sussex University, UK, 1991, pp. 26-36.

[4] In this comparison, the 486 chip not only operates about *6* times faster, it is also much more capable.

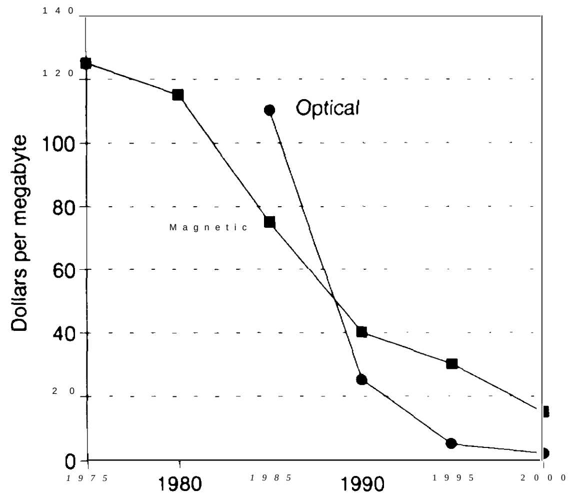
realized in workstation processors, data storage technology, and display technology.

These changes in the information industry have revolutionized the ability to process remotely sensed data. As recently as five years ago, digitally manipulating a remotely sensed image required expensive software and a high capacity workstation. Since then, advances in personal computer capability and software, and reductions in costs have led to much broader access to technology capable of processing remotely sensed data. Many image processing software packages now exist, ranging from "freeware" products to comprehensive professional systems that cost several thousand dollars.[5] Current systems also incorporate far greater amounts of data storage than did their predecessors. Data storage technologies, which include both magnetic and optical media, have improved dramatically while costs have plummeted (figure 2-1). In addition to making the data sets more usable, improvements in storage technologies have had a major impact on the ability of satellite operators to collect and store large amounts of data.

Accessing and Using Data

In order to be beneficial to a variety of users, data must be transportable. Data can be hand carried on portable media like magnetic tapes and disks or optical disks; they can also be transmitted over standard telephone lines, fiber optic cables, or relayed by satellite. Increasingly, individuals and institutions transmit and receive data and information over computer networks similar to telephone lines, but capable of transmitting data at higher rates.[6]

FIGURE 2-1: Cost (in then-year dollars) Per Megabyte of Storage Has Dropped Considerably



SOURCE National Media Laboratory, 1993

The distribution of significant processing power in personal computers and workstations has solidified the interest in online access to data and information. Most large companies (box 2-1), universities, research organizations, and state, local, and Federal government agencies now rely on computer local area networks (LANs)[7] to help workers organize and use a wide variety of data and information. The available software allows users to transmit and receive messages, operate software programs, and access data and information stored in central locations. U.S. industry has spent over $30 billion on local area networks to link workstations.[8] Additionally, users can also be linked together through wide-area networks

---

[5] Many of the commercial packages developed exclusively for remote sensing/geographic information systems applications are priced competitively at $500-$2,5(X).

[6] Because cable television cables can transmit data at higher rates than existing telephone lines, they provide one possible avenue for linking computers. Recently, Continental Cablevision, Inc., and Performance Systems International, Inc., introduced a service to link home computers with the Internet, making it possible to transmit voice and full feature graphics quickly. See Jared Sandberg, "Cable That Ties PCs to Internet to be Unveiled," *Wall Street Journal*, Feb. 8, 1994, p. B10.

[7] Local area networks (LANs), consist of a series of computer workstations linked together through a network server that provides each workstation with operating software and the ability to share electronic mail, data, and information.

[8] Information Industry Association, Digital Information Group, Link Resources 1993.

BOX 2-1: Information Management in the Retail Industry

U.S. industry depends on the ability to manage, interpret, and transmit data quickly and efficiently. For example, retailers like K-Mart and Wal-Mart use real-time electronic data interchange to control inventory, meet customer requests, and handle payroll and scheduling. Efficient transfer of large data sets can cut costs to the retailer, and allow transfer of point-of-sale data to others on the computer network, 1

For example, Wal-Mart's operating costs are low relative to its competitors, in part because the company dedicates only about 10 percent of its stores' square footage to inventory, compared to an industry average of 25 percent. Because sales data are tabulated immediately, the company is able to inform suppliers in a timely manner, and use the information to negotiate better prices from suppliers. Wal-Mart uses satellite links to provide this electronic data interchange, as do K-Mart, Home Depot, and others.

[1] See Lucie Juneau, "Luring Consumers With Conspicuous Efficiency," *ComputerWorld*, September 14, 1992

SOURCE Off Ice of Technology Assessment, 1994

(WANS) that operate over a wider geographic area. A WAN might be made up of one or more LANs and a number of individual computers. In order for distributed information systems to be effective, they must allow easy access to multiple users, segment information in searchable fields, and generally increase the efficiency of those who use the system.

This amalgamation of technologies and businesses has led to important synergies among technologies: technology developments in one sector necessitate and encourage technology development in other sectors. For example, the recent dramatic growth in the availability of multimedia CD-ROM[9] readers, driven by market demand for entertainment and educational CD-ROMS , has led to increased use of the CD-ROM for storing and distributing large amounts of digital data. The development of smaller, more powerful computer processors has led to an explosive growth in cellular telecommunications (currently an industry worth over $7 billion annually).[10] This development, in turn, has made the concept of handheld computers more viable, since cellular links will eventually make wireless computer networking practical in the near future.[11]

The growth of the online information industry (table 2-1) reflects increasing demand for databases, analysis, and information products. Although online access to data and information is

TABLE 2-1: Projected Growth of Online Information Industry (in billions of U.S. dollars)

Market segment	1990	1995
Financial	2,3	3.4
Travel	1.7	2.7
Marketing	1.5	2.9
Credit	1.5	2.0
Legal/regulatory/scienti fic*	1.0	1 9
Real estate	0.3	0 4
Insurance	0.3	0 5
News	0.3	0.4
Other	0.6	1,3
Total	9 5	15,5

* Including patient imformation

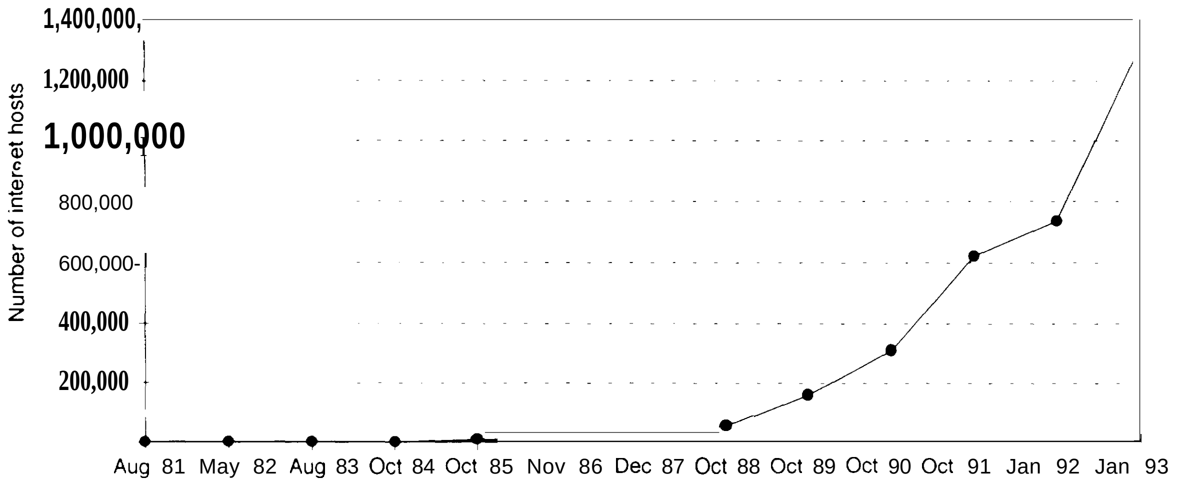
SOURCE. Information Industry Association, Digital information Group, LinkResources, 1993

[9] CD-ROM stands for compact disk with read-only memory, A CD-ROM is technically identical to the compact disk of the music industry.

[10] U.S. Department of Commerce, International Trade Administration, *U.S. Industrial Outlook 1993* (Washington, DC: U.S. Government Printing Office, 1993).

[11] Nine major Careers have teamed with IBM to develop a cellular data standard, and another, Cellular Data Inc., received an experimental license from the FCC to begin testing its technology in 1992.

FIGURE 2-2: Number of Host Computers Tied to Internet



SOURCE SRI International as cited in inforworld, April 12, 1993, p 38

only part of the information industry as a whole, the growth of online systems provides insight to future possibilities for remotely sensed data. The consumer of remotely sensed data can expect, eventually, to be able to tap into a wide variety of online databases.

Although U.S. industry and government have extensive experience using and transferring large tabular data sets among users, fews ystems require the amount of image storage and manipulation that data from some remote sensing systems, such as Landsat and SPOT, require. Processing and storing these geospatial data, which contain several levels of data about each geographic point, provide challenges that sales data do not. In addition, most applications of sales data are well known, making the selection of database formats relatively straightforward. By contrast, applications of satellite geospatial data are continually evolving as data users gain experience with the data. Thus, data and information systems for geospatial data must be flexible and easy to use.

The online distribution of remotely sensed data has been enhanced by the availability of the Internet, a wide-area information system funded in part by the National Science Foundation. Because of its ability to connect individuals with other Internet users and with widely scattered databases, Internet has grown rapidly over the past few years (figure 2-2).[13] The rapid growth of Internet has come about, in part, because many users have access to Internet through commercial data networks, provided by a host of commercial suppliers.

Growth of independent commercial systems such as Prodigy, America On-line, or CompuServe has occurred at rates similar to the growth of Inter-

[12] Internet began in the 1970s as a Department of Defense experimental project to connect computer systems dispersed around the country. The success of this system led the National Science Foundation to fund the development of similar technology to allow scientists and government employees to communicate electronically.

[13] See U.S. Congress, Office of Technology Assessment, *Advanced Network Technology*, OTA-BP-TCT-101 (Washington, DC: U.S. Government Printing office, 1993) for a summary of issues concerning the improvement of Internet and other computer netork technologies.

net. Commercial and public information networks now link millions of people.[14] Customers of such systems can exchange files, electronic mail, and obtain news and other current information. These commercial systems also provide links to the Internet. New remote sensing information systems will make use of Internet or Internet-like systems to serve consumers who want to browse data system holdings, but do not require rapid transfer of large data files.

Because of increased traffic, Internet needs to be upgraded continuously. Despite its value, the fate of the National Research and Education Network (NREN), which would provide a significant increase in communication capacity, is uncertain.[15] Because providers of remotely sensed data and information tentative y plan to use such a distribution system,[16] the Internet will have a Significant impact on remotely sensed data systems. Should a successor(s) to Internet be developed, it will place government in the position of providing services to customers who may want to use Internet for commercial traffic.[17]

Access to online information also provides quicker access to information than most other forms of distribution, and for some applications, timeliness is a key to effective data use. As data transmission techniques and capacity improve, more rapid delivery of data is likely to result in greater numbers of network users. The distribution of large data sets, such as remotely sensed Earth data, will require increasingly powerful communications networks (box 2-2).

Because upgrades to new transmission technologies are expensive, many data communication systems still operate at low data rates. Most dedicated data networks have a data transfer rate that ranges from 64,000 bits per second (bps) to 1.544 million bps. To see what these rates mean compared to a common storage medium consider a standard CD-ROM, which can store 5.4 gigabits of data, or approximately 680 megabytes. Current communication networks have data rates that range from slower integrated signal digital network (ISDN) standard lines[18] that would require nearly an entire day to transmit the equivalent of a CD-ROM to high speed, high capacity T-3 lines that can transmit this amount of data (table 2-2) in about a minute.

COLLECTION AND PROCESSING OF REMOTELY SENSED DATA

Remotely sensed data are acquired by a sensor, then are either transmitted to Earth or stored on board for transmission at a later time.[19] If stored, data are eventually transmitted to a data relay satellite or directly to a ground station when the satellite's orbit takes it within line of sight of the station.[20] The amount of data generated by a sensor depends on several variables: resolution, swath width, and the number of spectral bands included in the sensor. As the resolution of a sensor im-

---

[14] For example, as of August, 1993, CompuServe, a provider of online services, had 1.4 million paying customers. The company adds approximately 10,000 online subscribers (on average) per month. America Online's subscriber base rose from 300,000 in July 1993 to 600,000 in January 1994. See Michael Dresser, "Getting on Line," *The Sun*. Mar. 6, 1994, pp. 1 D, 4D.

[15] NREN would result in a significant upgrade of the government-operated part of the Internet. See Off Ice of Science and Technology Policy, *The National Research and Education Network Program*, A Report to Congress in response to a requirement of the High Performance Computing Act of 1991 (P.L. 102-194), December 1992.

[16] NASA briefings. See also Committee on Earth and Environmental Sciences, *The U.S. Global Change Data and Information Management Plan* (Washington, DC: National Science Foundation, 1992) pp. 81-83.

[17] U.S. Congress, Office of Technology Assessment, *Advanced Network Technology*, op. cit., p. 19.

[18] The ISDN standard is currently met by phone lines in many parts of the country.

[19] Data would be stored if no appropriate ground station or relay satellite are within range of the remote sensing satellite.

[20] Currently, the Tracking Data and Relay Satellite System (TDRSS), a set of three satellites, receive data from proper'] y equipped satellites for retransmission to ground stations.

BOX 2-2: Network Data Transmission

Data can be transmitted in either digital or analog mode, each of which has technical advantages. Current communications networks generally rely on a mixture of old and new communications media that support both analog and digital transmission. An analog signal is a continuously varying electromagnetic wave that can be propagated over a variety of media. A digital signal is a sequence of electrical, radio, or optical pulses that represent (binary) 1s and 0s (each 1 or O is referred to as a "bit" of data), Either analog or digital signals can be sent over wire or optical fiber transmission lines, A signal attenuates (e. g., becomes weaker) the further it travels from its source. Hence, some type of amplification is used to boost the energy in the signal. Unfortunately, amplification also increases the amount of noise mixed in with the signal, To prevent data errors digital transmission can use regeneration the use of repeaters to recover the bits (the pattern of 1s and 0s), and retransmit the signal This procedure preserves the integrity of the data,l The error rate can be made as small as desired (but cannot be made zero) by placing repeaters sufficiently close together

Transmission media (the physical path between the transmitter and receiver) range from insulated copper wires, known as twisted pairs, to optical fibers made from silica or high-grade plastics Twisted pairs carry most analog and digital transmissions. For analog signals, twisted pair transmission lines require amplifiers every 5 to 6 km, digital signals require repeaters every 2 to 3 km. Twisted pairs can accommodate data rates as high as 4 megabits (mega=million) per second (Mbps). Coaxial cable also uses two conductors, but is constructed differently than twisted pair to enable transmission over a wider range of frequencies Cable is used to transmit telephone and television signals and for local area computer networks. A data rate of 500 Mbps makes cable a versatile medium and, because it is better shielded than twisted pair, cable is less susceptible to external interference. Fiber optic cable, which transmits an encoded beam of light by reflecting it at shallow angles through the fiber at data rates of up to two gigablts (giga=billion) per second (Gbps) was one of the most significant technological breakthroughs in data transmission of the 1980s. Low attenuation and the need for fewer repeaters, in addition to light weight and small size, make fiber highly attractive.[2] However, the cost of fiber remains prohibitive for many applications, For instance, lease fees for a fully switched optical network would cost between $5,000 and $10,000 per month, depending on capacity, length of line, and the indvidual carrier's fee structure.[3]

[1] William Stallings, *Data and Computer Communications* (New York, NY: MacMillan Publishing, 1991) pp. 40-59

[2] Stall ings, op cit , pp 59-72

[3] AT&T to Slash T1 Line prices, " *Communications Week,* Aug 10, 1992

SOURCE Office of Technology Assessment, 1994

proves, the resulting data rate increases as the square of the resolution, all other factors being equal. For instance, a sensor with a ground resolution of 10 meters has 4 times the data rate as a sensor with 20 meters resolution viewing the same area. As higher resolution sensors become a reality, the data handling problems become more severe. Higher rates of data collection also require ground stations capable of receiving more bits of data per second.[21]

[21] Or, for systems that transmit data through another satellite such as one of NASA's Tracking and Data Relay Satellites (TDRS), the relay satellite needs to have the capability of high-capacity transmission.

TABLE 2-2: Time Required to Transmit Information Equivalent to CD Storage

Transmission medium	Data rate (bits per second)	Approximate time Required
1,200 bps modem	1,200	1 month
9,600 bps modem	9,600	1 week
ISDN	64,000	1 day
T1 Fiber	1.544 million	1 hour
T3 Fiber	45 million	1 minute
OC-48 Fiber	2.488 billion	1 second

SOURCE U S Congress, Office of Technology Assessment, *Making Government* Work. New Directions for Electronic Service Delivery, OTA-TCT-578 (Washington, DC U S Government Printing Office, September 1993), p 40

Once received at the ground station, data undergo an initial stage of processing. Ground stations apply some calibration and geometric corrections to the sensed measurements. The data may also be geocoded—registered so that each data pixel corresponds to a known point on Earth.** In addition, data may be enhanced for visual presentation, analyzed for information content, and archived according to date, area of coverage, etc. These steps are important in transforming raw data into useful information.

As noted earlier, remotely sensed data can be delivered to the customer in a variety of ways—magnetic tape, photographic prints and transparencies, optical disk, CD-ROMs, and by online electronic transmission. Remotely sensed data present special problems for information systems because of their relatively high demands for storage, processing, and transmission capacity. Most land remote sensing scenes, which are typically 100 megabytes or more, have been transfered on magnetic tape or photographic media.[23] For customers who request delivery of a data product immediately, electronic delivery over commercial telephone lines or dedicated communications lines is possible. In order to transfer the large data sets represented by remotely sensed data, data providers must maximize the data flow rate by using high capacity data lines (box 2-2) and by employing compression techniques to condense the data files.

For example, a typical (multispectral) SPOT scene of 60x60 km is a digital file that requires about 100 megabytes of storage.[24] Hence, transferring an uncompressed SPOT scene of 100 megabytes at a rate of 64,000 bps would require nearly four hours. A single, 7-band Landsat Thematic Mapper image[25] of 185 km by 170 km contains about 400 megabytes of data, therefore taking about four times longer to transfer over an ISDN line than the SPOT scene. The amount of time required for transfer is also influenced by the method of connection. For example, users can download data faster from a database via Internet than over a modem[26] and phone line. This is because telephone line bandwidth is lower than network lines.[27]

Telecommunications companies are rapidly increasing the capacity of their networks. However, the average users' limited access to high capability T-3 lines often requires offline data distribution methods, particularly for a series of large data files. Most data from U.S. remote sensing satellites are available online only for preview; actual scenes are mailed to customers on tape or disc.

22SPOT, Image Corp. for example, now markets data from the SPOT satellites that are corrected for terrain distortions and geocoded. These SPOT view image data are available in several different sizes, including the standard 7.5 minute quadrangles of the U.S. Geological Survey (USGS).

23SPOT Image Corp. has begun to sell data in a wide variety of formats and along geographic lines customized for customer needs. For example, it can put up to 16 standard scenes on a CD-ROM or sell data by the square mile or by the linear mile (with a minimum of 2,500 square miles or minimum length of 100 miles). EOSAT, too, has broadened its range of data delivery media and products.

24A byte is 8 bits of information, the number required to form an ASCII character. Hence, each character in this file requires l byte Of storage. A megabyte is one million bytes.

25 Six of the seven spectral bands have a spatial resolution of 30 meters. Band 6, the thermal band, has a resolution of 120 meters.

26Modulate-demodulate (essentially a digital-to-analog converter).

27&cause of this, many large data and information systems are not accessible via modem.

While the delay in transmitting data in this fashion may seem insignificant in most applications, more timely delivery is needed for uses such as disaster monitoring, agricultural production, or ship routing. In its Earth Observing System Data and Information Systems (EOSDIS)[28], NASA expects to provide the ability for scientists at widely dispersed sites to use networks to conduct research together on large remotely sensed data sets. Many telecommunications experts expect the costs associated with data transmission over high-speed lines (Tl and T3) to fall as capacity increases. Yet despite falling data transmission costs, minimizing the size of files transmitted will ensure further cost savings as data transmission needs increase.

In order to speed transmission and reduce the storage requirements, computer experts have devised a variety of data compression schemes to condense the amount of data into files of manageable size (box 2-3).

The development of compression techniques is critical to the development of large data systems and archives. Compression techniques can be "lossy" or "lossless," that is, the data can absorb some acceptable error level through compression

BOX 2-3: Data Compression

Data compression is the process of condensing, or compressing, the amount of data that must be transmitted from point to point. For example, a single, black and white typed page (like this one) requires about 3.74 million bits (468,000 bytes) of storage when scanned at 200 pixels [1] per inch, and 132,000 bits (1 6,500 bytes) when compressed, Compression ratios (the ratio of uncompressed to compressed data files) range from 71 to 30:1, most remote sensing applications need compression schemes that have near zero loss and achieve 101 to 20:1 compression,[2]

One of the challenges for future remote sensing satellites will be the application of reliable data compression schemes to minimize the transmission time required for large amounts of data. Compression schemes for remote sensing work in two general ways. First, an image or other data set has a great deal of repetition For example, a printed page has many blank spaces that can be condensed for transmission by inserting a single symbol that indicates the length of the blank spaces. A half-tone photograph has fewer blank spaces, but many contiguous areas of equal density that can be indicated in a similar manner A measurement of ocean temperature represented by a color image Will similarly contain large areas of the same color A compression scheme will represent that area with a short instruction to the image processing software to recreate the area with the correct color. When the image is decompressed, the processor fills in the color appropriately,

A second compression technique involves reducing the changes required between scenes, Using the same example, if the temperature profile is updated, or a new area in the same region is measured, only the changes to the original image will be transmitted. New generations of remote sensors will rely on Improved compression techniques. At present, 10:1 compression is achievable for most Imaging applications.

1 Picture elements

[2] Don M Avedon, *Introduction to Electronic Imaging,* (Silver Spring, MD Association for Information and Image Management, 1992)

SOURCE Office of Technology Assessment, 1994

[28] See ch. 3.

or can emerge from a low rate compression scheme with no errors. For high-precision measurements, lossless compression is required.

Archiving data ordinarily requires lossless data compression because of the need for an accurate and complete record of the data. The use of lossy compression methods is acceptable when the user can tolerate some loss of precision. For example, users who wish to browse through low-resolution versions of data in order to check the appropriateness and quality of particular data sets do not require high precision.[29]

DATA ARCHIVES

In order for data and information to be useful to a wide range of users over many years, they must be stored under archival conditions and made easily accessible. The federal government archives an astounding variety of data and information generated by the various departments and independent agencies. Because the government continues to make considerable investment in the acquisition of remotely sensed data, and since many of these data are crucial to environmental studies, especially studies of environmental change, protecting that investment by ensuring that the data are accurate, standardized, and of high quality has become increasingly important.

The U.S. government maintains several data archives, located around the country, which store, protect, and distribute climatic, hydrologic, geophysical, and other environmental data (table 1-2). Many of these archives currently have significant holdings that are used by government, academia, and industry. The two largest archives provide a glimpse of the challenges facing the archiving and distribution of satellite and related data:

U.S. Geological Survey EROS Data Center

Established in 1972, the Earth Resources Observation Systems (EROS) Data Center is the primary archive for land remote sensing data collected by the U.S. government. As the National Satellite Land Remote Sensing Archive,[30] it archives digital data totaling nearly 80 terabytes[31] (figure **2-3) 53** terabytes of which are data from the Landsat system collected between 1972 and 1978.[32] The EROS Data Center also maintains an archive of data collected by National Oceanic and Atmospheric Administration's (NOAA) advanced very high resolution radiometer (AVHRR) carried by the polar orbiting satellites,[33] aerial photographs collected by the U. S. Geological Survey (USGS) National Aerial Photography Program, and USGS airborne radar data. It also contains a variety of Earth science, cartographic, and geographic data.

In 1993, the EROS Data Center distributed nearly $5.7 million worth of USGS and Landsat[34] data products and services. EROS Data Center provides data on a repay basis to government

---

[29] Online data services commonly offer images for customers to browse through, from which Up to 90 percent of the information has been removed in order to handle them quickly and easily. Nahum Gershon and Jeff Dozier, "The Difficulty With Data," *Byte*, April 1993, pp. 143-147.

[30] Provided for in *the Land Remote Sensing Policy Act of* 1992, Public Law 102-555, Section 50*:

"The Secretary of the Interior, in consultation with the Landsat Program Management, 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 *Land Remote Sensing Commercialization Act of* 1984, Public Law *98-365, sec. 602* contained a nearly identical provision relating to a data archive.

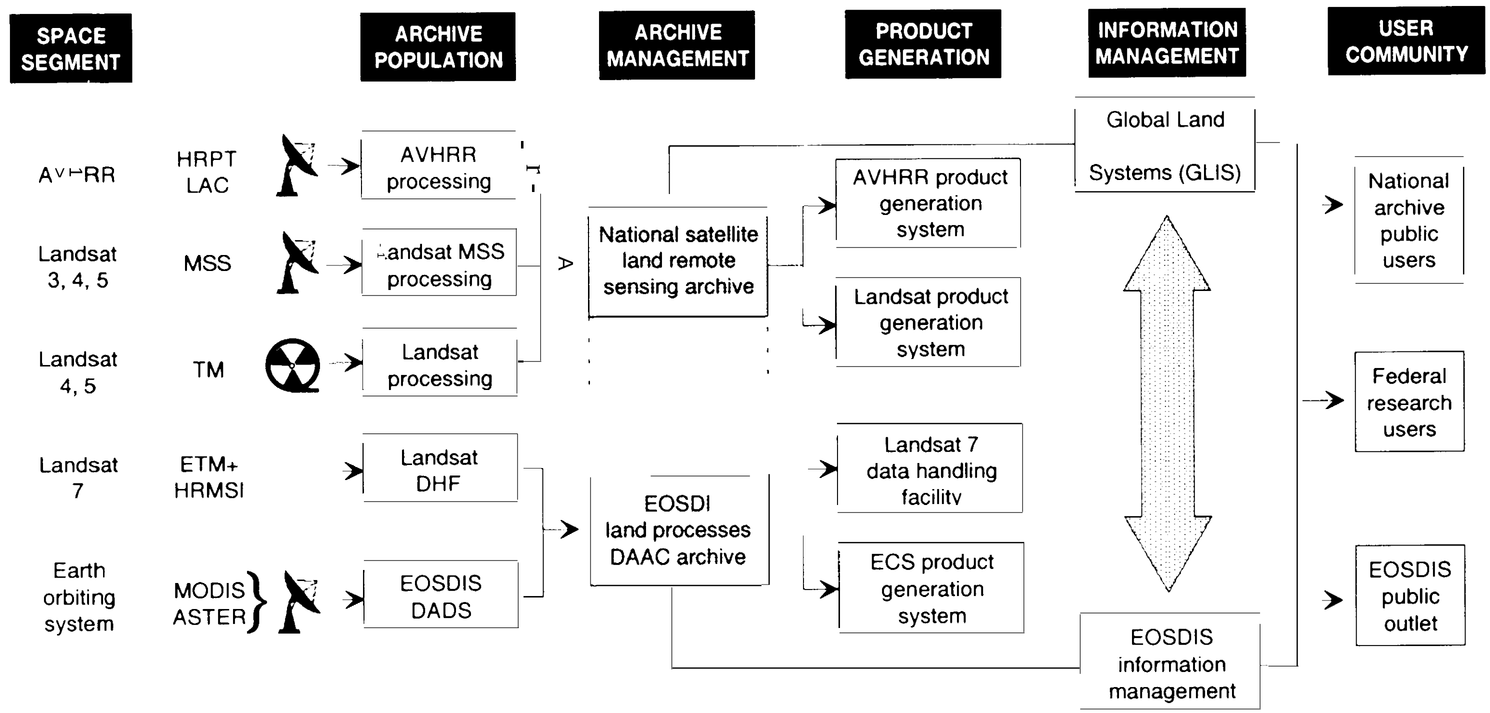
[31] A terrabyte equals 1,000 gigabytes; a gigabyte equals 1,000 megabytes.

[32] U.S. Department of the Interior, U.S. Geological Survey, EROS Data Center, National Satellite Land Remote Sensing Data Archive, Conference Aug. 26, 1993, Sioux Falls SD.

[33] EROS Data Center's AVHRR holdings equal about 3 terrabytes.

[34] Approximately $1 million of the Landsat data were produced at EOSAT, but billed by the EROS data center.

FIGURE 2-3: Satellite Land Remote Sensing Data Managment



SOURCE: EROS Data Center, 1994.

agencies and on a direct payment basis to foreign users and the public. It sells all data at the cost of reproduction, following the guidelines of the Office of Management and Budget's Circular A- 130, which sets out the terms by which government data and information are made available to the public.

In order to inform data customers about its holdings of satellite data, EROS Data Center has developed a "metadata" system that provides data about its digital data holdings, called the global land information system (GLIS). GLIS, which can be accessed through Internet, allows customers to determine what Landsat and other digital data the EROS Data Center holds. In addition, GLIS allows potential customers to examine data on-line [35] to check for extent of cloud cover and other features before placing an order. GLIS does not allow for direct digital downloading of data, primarily because officials of the EROS Data Center did not consider the investment in a billing system to be cost effective.[36] In addition, as described earlier, existing transmission rates are too low for efficient data transfer. Multispectral Sensor digital data on tape cost $200 per Landsat scene. Thematic Mapper data, when they first become available later this year, are likely to cost between $300 and $500 per scene on tape.

The EROS Data Center maintains a staff with expertise in the Earth sciences, such as geology, hydrology, cartography, geography, agronomy, soils science, and forestry. These scientists work on scientific problems of local, regional, and global change and assist EROS Data Center customers in making the best use of their data. The center's staff also includes experts in systems development, telecommunications, and computer sciences, which are needed to improve the center's ability to archive and deliver data more efficiently.

The EROS Data Center will serve as a distributed active archive center (DAAC) for NASA's EOSDIS,[37] adding archival responsibilities for land processes data from new NASA satellites to its current functions. It will archive data from several Earth Observation System sensors, including the advanced spaceborne thermal emission and reflection radiometer (ASTER) and multiangle imaging spectroradiometer (MISR). The process of archiving data from both sensors will require EROS Data Center to install new hardware and software to handle the additional archiving and distribution load.

| NOAA's National Climatic Data Center

The Federal Records Act of 1950 originally established the National Weather Records Center in New Orleans. In 1951, it was moved to Asheville, North Carolina and renamed the National Climatic Data Center (NCDC). NCDC is currently the world's largest active archive of weather and climate data, containing about 30 terrabytes of weather and climate data (figure 2-4).[38] It serves as the archive for data from the National Weather Service, military services, Federal Aviation Administration, and the Coast Guard. NCDC also accepts weather data from foreign sources (figure 2-5). The center archives 99 percent of all NOAA data, including satellite weather images back to 1960. The center archives data collected by satellites, radar, aircraft, ships, radiosonde, and National Weather Service stations. It archives about 55 gigabytes of new data each day. NCDC operates the World Data Center-A for Meteorology, and both gathers and shares data internationally.

---

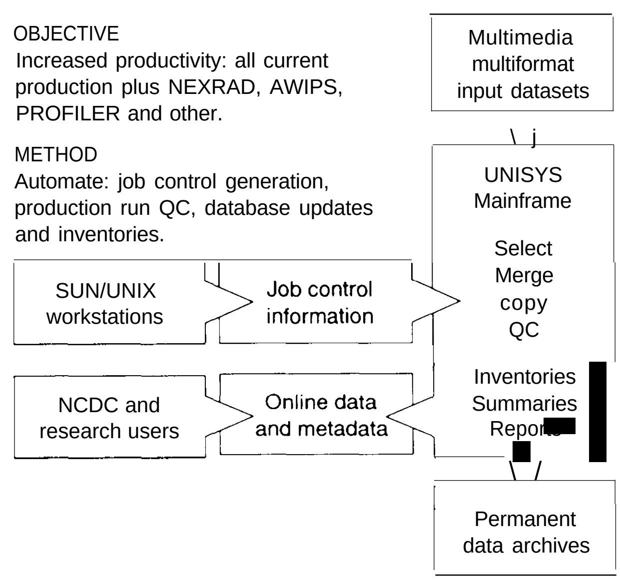
35 These "sampled" scenes, Which are generated by extracting about 1O percent of the original data, are not of sufficient quality to use. but provide customers with an excellent visual tool to determine whether the data can be used for their project.

36 Data Center staff expressed concerns about maintaining credit card numbers and other confidential information in a data system that would allow virtually unlimited front-end access.

37 See ch. 3 for an extensive discussion of EOSDIS.

38 Source: NOAA, NESDIS, briefing, 1993.

FIGURE 2-4: National Climatic Data Center Comprehensive Archive Management System



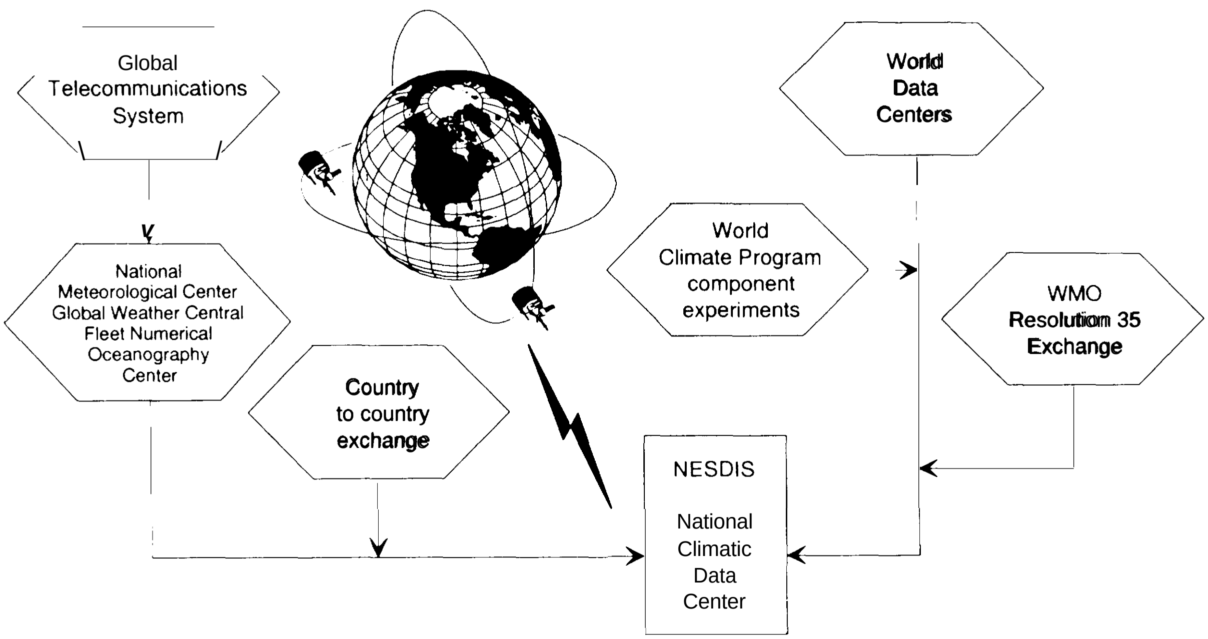
SOURCE National Cilmatic Data Center, 1994

Data from NOAA's Polar-orbiting Operational Environmental Satellite (POES) system are collected and stored at NCDC'S facility in Suitland, Maryland. Backup copies of data tapes are stored in Asheville.

NCDC maintains 455 different data sets and responds to about 90,000 data requests each year, supporting many forms of data dissemination: paper, photographs, magnetic tape, floppy disks, CD-ROM, electronic mail, online dial-up, telephone, and facsimile. Data costs vary according to the amount of effort NCDC personnel expend in providing the information. For example, NCDC charges an additional fee for certifying the authenticity y of climate data, often needed to support legal proceedings. In order to deliver data to data users as quickly as possible, NCDC has developed the OASIS online services (table 2-3).

In addition to publishing their own research on climate, center staff also provide historical perspectives on climate vital to studies of climate change and the environment. For example, in September 1993, the

FIGURE 2-5: National Climatic Data Center: Data Received From International Sources



SOURCE National Climatic Data Center, 1994

TABLE 2-3: National Climatic Data Center (NCDC) Online Access and Service Information System (OASIS)

Data	Period of record	Process time before online availability
Wind Profiler		
Winds -60	31 days	1 1/2 -2 hours
Surface -60	31 days	1 1/2 -2 hours
Moments -60	7 days	1 1/2 -2 hours
Moments -60	7 days	1 1/2 -2 hours
National Weather Service, surface hourly	2 year	2-3 weeks
National Weather Service, rawinsonde	2 year	3-4 months
National Weather Service,		
Hourly	2 year	8-9 months
15 minutes	2 year	8-9 months
N American rawinsonde	from Jan 1, 1992	24 hours
Cooperative Summary of the day (TD 3200) state files	1992	as required
General Circulation Model carbon dioxide	100 year	as required
Field Experiment		
CaPE	06- 07/ 1991	as required
STORM -FEST	02- 03/1 992	as required
GCIP	02- 04/ 1992	as required
NEXRAD Level II		
Inventories	**03 / 91-**	2-4 months

NCDC has online data and metadata available by file transfer protocal (FTP) computer access Data are placed online as soon as possible after receipt or processing, Those data are available without charge via FTP for immediate downloading (up to 50 MB) or users can order data for offline delivery (standard NCDC charges). In addition to data, important metadata are Included with the online data Station histories, data dictionaries, field experiment information, and data inventories are Included,

SOURCE National Climatic Data Center, 1994

center published a technical report on the special weather stresses experienced in the United States during the summer of 1993 and their effects on the U.S. population.[39] The center also supports a wide variety of industrial, agricultural, and engineering applications for historical climate data (table 2-4).

TECHNOLOGY FOR ARCHIVING DATA

Data archiving requires use of high-density storage media. Current media used by some of the larger data bases include magnetic tapes (1/2-inch tape, 8mm tape, 4mm tape), and optical disks (CD-ROM, CD-WORM,[40] and larger optical disks). Manufacturers now routinely develop and market new magnetic and optical media, forcing difficult choices among the government agencies and other data providers about how to improve their archives and distribution system. The development of storage standards can be a significant problem. If data are stored on media supported by only a few suppliers, they could become expensive and difficult to replace as they wear out. In the worst case, they could become unreadable as the

[39] Neal Lott, "The Summer of 1993: Flooding in the Midwest and Drought in the Southeast," NCDC Research Customer Service Group, Technical Report 93-04.

[40] CD-WORM stands for compact disk, write-once, read-many times memory.

TABLE 2-4: Applications of Historical Climate Data

Customer type	Data used	PurPose
A Manufacturing		
1 Automobile	"Local Climatological Data" Surface Weather Observations	Testing new batteries Determining cause of air conditioning failure
2 Weather dependent products: umbrellas, shoes, firelogs, sunglasses, etc	"Local Climatological Data"	Determining impact of weather on sales
3 Software manufacturer	Surface Weather Observations	Database development
4. EPA/consultants	Mixing Height Studies	Air pollution control
B Energy		
1 Electric power companies	Heating & Cooling Degree Data	Determining level of electrical demand through computer models
2 Gas utility	"Monthly Station Normals of Temp, Precip, Heating and Cooling Degree Days"	Determine rate adjustments and expected demand
3 Energy consultant	Wind Energy Resource Information System	Determining possibility of using wind mills
4 Battelle Pacific Lab	DATSAV 2 data (Hourly Surface Obs)	Study for wind energy usage in Third World Countries
5 Oil companies	Navy Marine Summaries International Station CD-ROM	Planning offshore oil drilling platforms
C Agriculture		
1 NWS Agricultural Weather Service	COOP Data	Advisory and research
2 Horticultural firm	International Station CD-ROM	Design of greenhouses and determining areas for crops
3 Herbicide manufacturer	Surface Weather Observations	Determining effects of temperature
4 Horticultural firm	"Frost/Freeze" Publication	Planning for crops
5 Entomologist	COOP Weather Reports	Determining life cycle of insects
6 Commodities exchanges	"Local Climatological Data"	Crop storage planning Effect of climate variations on crop yields

TABLE 2-4: Applications of Historical Climate Data (Cont'd.)		
Customer type	**Data used**	**Purpose**
D Consultants and Engineers		
1 Meteorological consultant	All types of observations Map Analyses Climatic Averages	Expert testimony Climatic Studies
2 Marine consultant	Marine Wind/wave summaries	Port design
3. Engineering firm	FAA Wind Rose	Airport design
4 Engineering firm	Weather Bureau Technical Paper No, 40 NOAA Tech, Memo Hydro-35	Building, highway, dam design/ flood control
5 American Society of Civil Engineers	DATSAV 2	Development of ice Ioadlng Guidelines
6 Architects	COOP/Climatlc Summaries	Planning construction projects
7 University research	North Atlantic Hurricane Tracking Data	Storm risk analyses
8 Unversity research	Summary of the Day Weather Data	Land Use impact study
9 Engineeting firm	Soil Freeze Depth Maps	Construction design and planning
E. Entertainment		
1 Festivals, concerts, sports events, conventions	COOP Weather Reports Climatic Averages	Establish normals and patterns for planning purposes
2 Golf course development	Climatic Averages	Planning useage
3 Resort development Conference	Climatic Averages	Development
4 Insurance industry	Climatic Averages Hourly Precip Data	Planning insurance for entertainment events/verification
5 Authors	'(Local Climatological Data" Climatic Averages	Books: verification of data
F. Communications		
1 Institute of Telecommunications	International Station CD-ROM	Communication system planning
2. Television advertising	Climatic Averages	Establish ideal filming times
3 Ad agencies	Climatic Averages	Determine markets
4 IV/Radio programs Magazines	"Local Climatological Data" Surface Weather Observations	Fact verification

TABLE 2-4: Applications of Historical Climate Data (Cont'd.)

Customer type	Data used	Purpose
5 Cable television firm	COOP Data	Microwave attenuation study for tower planning
G Services		
1 Physicians/medical research centers	Surface Weather Observations Climatic Averages	DiSeaSe/Climate-weather correlations
2 Highway departments	Climatic Averages	Planning snow removal
3 Insurance companies	Climatlc Averages Surface Weather Observations	Planning and verification
4 Insurance companies	Surface Weather Observations COOP Observations	Settle weather related disaster claims
5 Attorneys	Surface Weather Observations COOP Observations	Settle legal disputes
H Housing/real estate		
1 Contractors	Surface Weather Observations "Local Climatologlcal Data" COOP Observations	Determine construction deadline penalties or extensions
2 Real estate developers	Climatic Averages	Site selection for resort and retirement developments
I Transportation		
1 Trucking companies	Climatic Averages	Expedite transport of perishable goods
2 Airline companies	Climatic Averages Map Analyses	Determine favorable air routes
3 Marine shipplng companies	Climatic Averages Storm Tracks Marine CD-ROM	Determine optmal shipplng routes
4 Railroad companies	Climatic Averages	Determine favorable seasons and routes for transport of goods and commodities

As the repository of weather data from the United States and the world, the National Climatic Data Center is charged by law with providing accurate, historical data about US weather and climate. Historcal data about the climate serve a variety of useful applications in agriculture, construction, engineering, law, and transportation This table lists a sampling of uses of climate data in the U S Economy

SOURCE National Climatic Data Center, 1994.

BOX 2-4: Rapidly Changing Technology and Data Storage

Because data must be stored on media available at the time of collection, electronic databases not only benefit from but require periodic storage upgrades. Upgrading data storage provides a number of benefits, including better performance, reduced storage costs, and greater storage density.

Data managers must transfer electronically stored data to new media for two reasons: First, magnetic media deteriorate over time, resulting in some data loss. Second, media often become obsolete as new technologies are developed and the readers for older media lose market share. The rapid shift from long-play records to compact disk technology for commercial music provides an instructive example of such changes in the consumer marketplace.

In the smaller marketplace for scientific data, major improvements in technology can cause special problems in maintaining data. For example, the EROS Data Center has encountered difficulties transferring data acquired from Landsat 1 to new media. Some of these data are of particular significance because they represent the oldest multispectral satellite data available about Earth's surface, and are therefore of considerable importance to research on local and regional land changes. NASA originally recorded the data using proprietary recording equipment that has since failed as a result of age. The development and production team for the original Landsat 1 tape system has dispersed over the years, making it impossible to replicate the system without extraordinary investment.

The EROS Data Center has encountered a different problem with data from Landsat 3. The tapes had been stored at the NASA Goddard Space Flight Center Landsat facility, which was operating in the late 1970s before Landsat operations were transferred to NOAA. Many of these high density tapes now suffer from "sticky tape syndrome, " a condition they developed when subjected to excessive humidity while in storage at Goddard Space Flight Center. The magnetic tape coating absorbed moisture making them unusable without special treatment. The EROS Data Center and the National Media Laboratory have discovered that by heating the tapes in an oven, they can literally cook the moisture out of the tapes, which restores their readability for several weeks. ' After undergoing such processing, they can then be read by the center's tape readers and transferred to more stable modern media, allowing the EROS Data Center to make the data available to customers.

1 "Solving Sticky Tape Syndrome at EROS Data Center," *NML Bits* (Newsletter of the National Media Lab), VOI 3, September 1993

SOURCE: EROS Data Center, Office of Technology Assessment, 1994

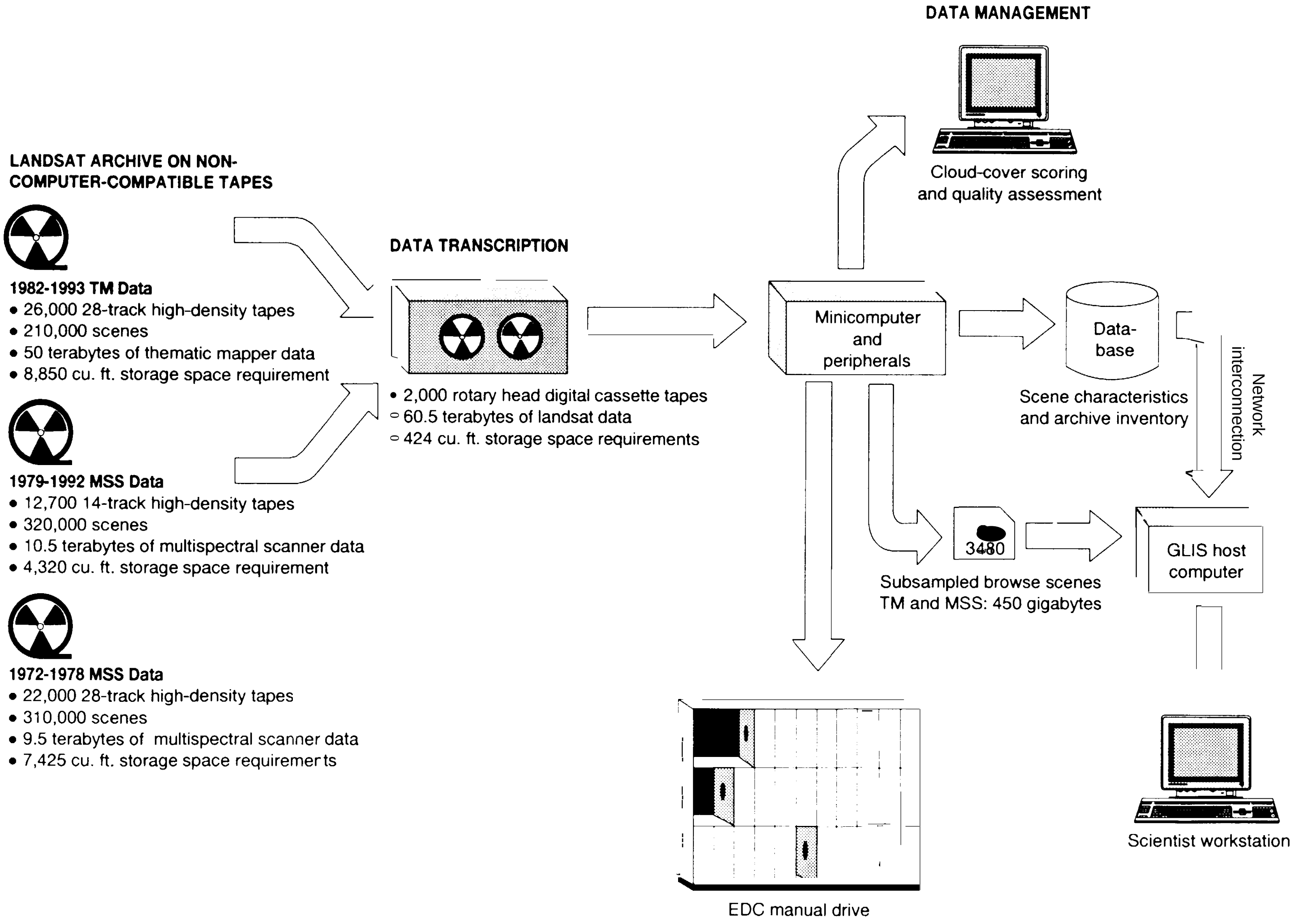
technology becomes obsolete (box **2-4;** figure **2-6).** Fortunately, the existing large market for information technologies gives archive managers some comfort that a storage medium in wide use will not become obsolete quickly.

NAVIGATING THE ARCHIVES

Many archival data are increasingly stored as digital images in specialized **databases.**[41] Yet, data sets are useless if the appropriate information cannot be accessed and applied to solve problems. The development of powerful database software has resulted in systems that can sort through a wide variety of data quickly and provide users with information in several formats. The number and size of data and information systems continues to increase. In addition to the growing collection of data sets archived by the federal **govern-**

[41] Gershon and Dozier, op. cit.

FIGURE 2-6: Thematic Mapper/Multispectral Scanner Conversion System Functional Capabilities



SOURCE: EROS Data Center, 1994.

ment (table 1-2), many states are attempting to develop comprehensive information systems to store and access data specific to their needs. Many of these data are geospatial data.[42] Data and information systems dedicated to managing specific resources (land, timber, oil, minerals, agriculture) are also being developed by companies with a financial stake in such resources.

As archival holdings grow, it will be essential to have the ability to search databases quickly and obtain useful information. Each data center maintains its own database. However, data users, both within the federal government and outside, often have difficulty in determining precisely where data they may need reside, or if the data exist at all. Some data exist in hard copy form such as maps, photographs, lists, and charts. Others are stored in electronic form. Several advisory groups have suggested that the use of available data would be made much more effcient the creation of a metadatabase that listed all data that falls within certain categories in one readily accessible place. Metadata provide summary information about data holdings and guides to accessing them—what data are available, where they are held, and how to access them. Metadata therefore function like a library's electronic card catalog. Having such information is not only helpful in navigating the many available archives, but may make it possible for agencies to avoid creating data sets that are already available elsewhere.

The Federal Geographic Data Committee (FGDC),[43] which is composed of representatives from Federal agencies that generate and use geospatial data, was established to "lead the development of the national spatial data infrastructure (NSDI) and to coordinated its implementation."~ For geospatial data generally, the President has directed:

- FGDC to establish a National Geospatial Data Clearinghouse that would gather geospatial metadata from the agencies,
- each agency to:
  - a) document all new geospatial data with the FGDC metadata standard and make that documentation electronically accessible,
  - b) adopt a schedule for documenting (where feasible) existing data,
  - c) adopt a plan to establish procedures for public access to geospatial data, and
  - d) adopt internal procedures to use the clearinghouse prior to expenditure of federal funds for data collection.

The Executive Order further states that FGDC is to develop standards for implementing the NSDI, and to submit a plan to the Office of Management and Budget (OMB) for:

> completing the initial implementation of a national digital geospatial data framework... and for establishing a process for ongoing data maintenance.

It further instructs FGDC to develop strategies for:

> maximizing cooperative participatory efforts with State, local, and tribal governments, the private sector, and other nonfederal organizations to share costs and improve efficiencies of acquiring geospatial data.[45]

Geospatial remotely sensed data are also subject to these directives. However, as noted above,

42 In other words, they have attributes such as soil type, resources, or other characteristics that can be placed geographically.

43 The FGDC is authorized by the Office of Management and Budget, Circular A-16 to coordinate federal agency involvement in the National Spatial Data Infrastructure.

44 Office of the Vice President, Department Of the Interior, "Recommendation DO1O3," *Accompanying Report of the National Performance Review*, September 1993. The Federal Government generates and uses many kinds of geospatial data. As conceived by the FGDC, NSDI is composed of the large collection of digital, geospatial data that constitutes a major part of the overall Federal information infrastructure. By establishing minimum standards for data acquisition and distribution, the FGDC hopes to "enable analysts and decisionmakers to integrate diverse geographic information" quickly and easily.

45 Executive Order 12906 of Apr. 11, 1994, "Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure."

satellites collect nonspatial data about Earth systems. For the purposes of global change research, environmental monitoring, and other applications, users will need ready access to information about the location and availability y of these data as well.

To avoid duplication in data acquisition and archiving, Congress may wish to consider instructing Federal agencies to coordinate on developing an online database that would hold metadata about all civilian remotely sensed data. Such a database would be able to field queries from government and private customers interested in remotely sensed data. This metadata set, which would be small enough to be located in one site, or distributed among several sites in a network, would not be the source for the data themselves but provide a guide to the holdings thoughout the federal government. It could also store information about data held in local and state government offices, and in private archives. Development of a central metadata set will require the cooperation of the many government agencies, including the Department of Agriculture, Department of Defense (DoD), Department of Interior, Environmental Protection Agency, NOAA, and NASA, as well as agencies that hold other kinds of geospatial data, such as census and land use data. A centrally-organized metadata set would ensure maximum exploitation of data sets that have already been acquired by the government and other users.[46] It would constitute an important component of the larger geospatial metadata set recommended by the FDGC.

GEOGRAPHIC INFORMATION SYSTEMS

Over the past decades, users of spatial data have harnessed the spectacular gains in the power and speed of computer hardware to develop systems capable of meeting their spec ial needs for processing and manipulating spatial data. These systems have become increasingly flexible, allowing users to analyze and manipulate digital images, add new information, and create layers of data that focus on different forms of information. Recent systems also allow users easily to convert various kinds of data to color-coded images that enable researchers to isolate patterns in the data.[47] This recent development is in part a result of new software programs that take advantage of increased processing speeds available in newer workstations.[48]

Data systems capable of assembling, storing, manipulating, and displaying geographically referenced data[49] are known as geographic information systems (GIS). Driven by simple commands, GIS can be used to display and analyze spatial data {box 2-5) in many different ways. GIS users can select from among many categories of information to display on a single digital image, depending on their needs. For example, in figure 2-7. the GIS layers display land use, transportation routes, potential hazardous waste sites, and hydrography. These scenes can be displayed with categories of information (grid lines, roads, zip codes) withheld, and added at a later time, or they can all be displayed at once. The variables are stored in relational data files, and the flexible program relates each variable to the proper image.

[46]See National Research Council, *Toward a National Spatial Data Infrastructure* (Washington, DC: National Academy Press, 1993), the report of the National Academy of Science's Federal Mapping Committee, for a detailed discussion of this point.

[47]For example the High Resolution Infrared Radiation Sounder aboard NOAA's POES satellites takes temperature and humidity soundings in a "column" through the atmosphere. By color-coding the data and plotting observations taken at different geographic locations, scientists can visual I/c changes in temperature and humidity at a given level of the atmosphere as the satellite sweeps across Earth 's surface.

W)\ computer's microprocesso design determines the number of calculations that can be made simultaneously. The clock speed is a measure of how fast the microprocessor operates. The microprocessor must synchron I/e system operations, read instructions from the main memory cache, manipulate data stored there. A microprocessor in a system operating at I MHz would operate 1 million times per second. Many new systems contain 486-based processors that operate at a clock speed of either 33 MHz, 50 MHz. or 66 MHz.

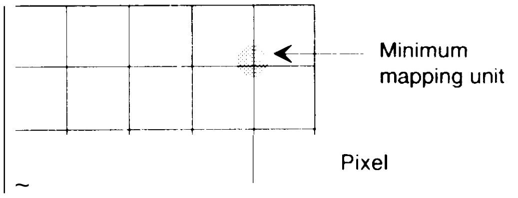
[49]Geographically referenced data are identified according to their location in some space—hence the term spatial data.

BOX 2-5: What are Spatial Data Se

Spatial data sets provide information about conditions at various locations in 2 or 3 dimensions There two broad categories of spatial data raster data and vector data

Raster data sets are obtained by imaging (photographically or by electronic sensor) an area of interest In obtaining a raster data set, the imaging device develops a value of intensity for the physical quantity of Interest (e g land cover features, elevation in meters, land use class) for every cell, or picture element (pixel). The individual pixel represents what is essentially an average Intensity within that area Pixels are defined by their columns and rows; pixels may be square or rectangular

The minimum mapping unit or resolution unit , is the smallest element that can be distinguished by the Imaging system. The pixel size should be smaller than the minimum resolution element, so as not to miss Important features. For example, in the schematic below, if the gray shaded area represents the minimum mapping unit, and is thus detectable by the Imager, it would not show up in the pixel, since it would not cover the majority of any one cell. ' If the pixels were smaller, the feature would show up in the image.



Vector data sets are collectiions of Information based on elemental points whose locations (made up of a point and a direction from an origin) are known to arbitrary precision such as "point A is 25 kilometers north northeast from the map's origin " The placement of a point on a vector-generated map is not limited by the size of a pixel Vector representation allows the map maker to represent map elements much more simply than raster representation. For example, in a vector data structure a circle can be represented in the computer by a point, a radius, and the thickness of the line By contrast, to represent the same circle in raster format would require storing all the pixels that make up a circle in the right sequence to represent the circle as an image Conventional maps developed in a computer have been based largely on vector data sets because vector representation is much more economical of computer memory Because cconventional maps are collections of standard symbols arrayed in a pattern that represents an abstract image Vector data sets such as maps focus primariy on major features (e g., roads, rivers mountain ranges, cities) and present information in a relative way.

Both raster and vector data are Important to consumers of remotely sensed data Increasingly, both data types are being combined to produce products that hold valuable information regarding details of land features, land cover, elevation, *and* relational aspects such as distance between objects For example, the U S. Geological Survey develops combination products known as digital ortho-photo quads, or quadrangle maps that use dlgitzed photographs to correct standard vector maps (produced by USGS for decades) Vendors that produce software for GIS have found a significant market in software that can convert between raster and vector data sets

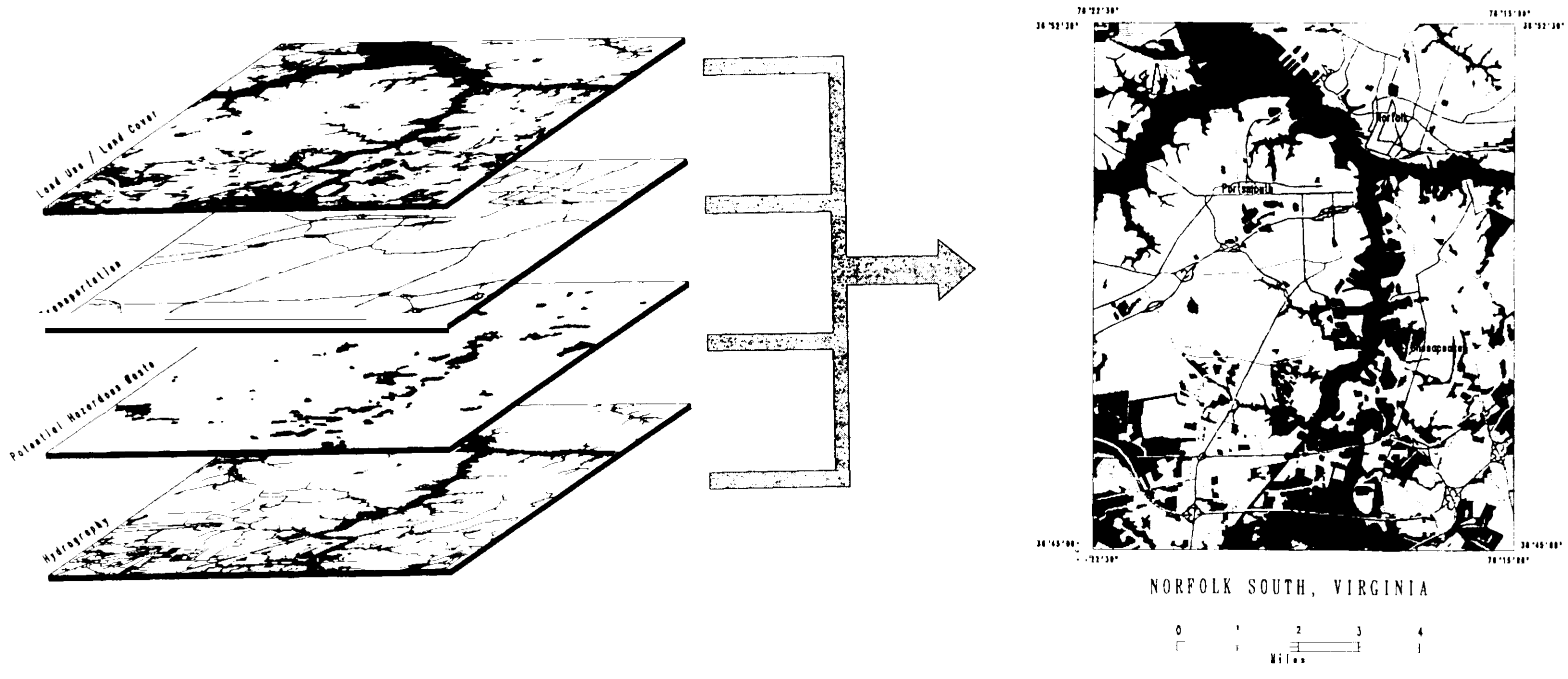
Increasingly, electronic maps are used to replace or augment paper. Paper maps, while an extremely effective way to display information, cannot match the density of similar information stored on floppy disks or CD-ROM A compact disk can store many digital maps and even images of places of particular interest.[2] Further, digital maps can be designed to contain Information of interest to specific classes of users Digital map readers can also be combined with Global Positioning System receivers, providing real-time position Information to the user

[1] If, however, the feature is extremely bright compared to its surroundings, it could be detected as present, though not resolved For example, the Landsat thematic mapper image of the Chernobyl reactor just after it suffered a failure and was burning registered a "hot spot" in the thermal band, which has a surface resolution of only 120 meters The burning reactor therefore could not be Imaged, but the fact that it was burning was never the less observed

[2] For example De Lorne Mapping markets a CD-ROM, Street Atlas, USA, which enables users to locate most U S streets and create and print maps of varying scales The data are stored in vector format for quick call-up and presentation

SOURCE: Off Ice of Technology Assessment, 1994

FIGURE 2-7: Environmental and Urban Analysis Using the Use of Geographic Information System Technology



The National Mapping Division of the U.S. Geological Survey, the Water Resources Division of Virginia, and the Environmental Protection Agency have cooperated to produce a database for a nine-quad area in the Elizabeth River Watershed of the Chesapeake Bay. Data from these and other scientific studies, cartographic and thematic sources, and administrative or regulatory files will be formatted for entry into the GIS. This area is to serve as a pilot project location where GIS activities associated with the Chesapeake Bay Program can be studied and demonstrated. Each of the four layers of information on the left can be printed and studied independently, or can be combined in various ways to produce a composite product.

SOURCE: U.S. Geological Survey, 1993.

The development of GIS has been a major force in the enhanced use of remotely sensed data during the late 1980s and early 1990s. A rapidly growing collection of users now has the ability to use GIS to analyze remotely sensed data and to incorporate other data with them. For example, a remotely sensed image could be used as a foundation for adding ownership boundries, sensitive environmental areas such as wetlands, zoning, historic sites, population densities, and transportation routes. All or part of the additional information could be displayed or printed as needed. GIS users are becoming more familiar with satellite data, and are using them with greater frequency. For example, civil engineers increasingly rely on satellite images of large geographic areas to analyze and explain construction projects, such as the construction of new highways.[50] Planners of routes for pipelines or high-capacity electrical transmission lines find remotely sensed data, coupled with GIS software, extremely useful in locating suitable routes quickly and accurately.sl Field positions derived from the military's global positioning system (GPS) make the accurate determination of geographical locations much simpler than ever before (box 2-6). Appendix B presents other examples of the combined use of these powerful technologies:

> The past few decades have witnessed tremendous development and change in spatial data handling....What some have characterized as the "digital revolution' '...has had a profound effect on almost all activities utilizing geographic information. These effects have been felt in the research and academic sector...in the private sector, which is populated with a large number of digital spatial software system vendors and a larger number of users of such systems; and in the federal, state and local levels of government, which are some of the largest producers and users of digital spatial data.[52]

The rapid increase in computing capabilities and geographic information systems has resulted in many different methods of storing and using data. Because data formats are seldom standardized,[53] data analyzed by one GIS software package may not be readable by another.[54] Lack of standardized data sets can cause serious logistical problems for data users intending to integrate data from multiple sources. This problem is particularly acute in the federal government, where purchases of remotely sensed data are difficult to coordinate across agencies, and operating systems are purchased independently. The development of GIS standards could improve the usefulness of the technology and enhance the market for remotely sensed data.

Data standardization efforts are underway. For example, as noted earlier, federal agenc ies formed the Federal Geographic Data Committee (FGDC) to serve as a forum for the discussion of spatial data issues within the federal govemment.[55] Among other things, the committee has taken the lead in insuring that various software applications

[50] Harold Hough, "Satellite Imagery Charts Course for Civil Engineers in Jacksonville, Florida," *Earth Observation Magazine*, June 1993, pp. 25-28.

[51] To accomodate such users, SPOT Image Corp. now sells data of arbitrary shape by the square mile. It will for example, sell data that stretches along linear features such as pipline rights of way. See "The New Benefits of Space Age *Technology, ''A-E-C Automation Newsletter,* November 1993, pp. 2-7.

[52] Harold Moellering, '* Opportunities for Use of the Spatial Data Transfer Standard at the State and Local Level," *Cartography* and Geo*graphic Information Systems*, Journal of American Congress on Surveying and Mapping, vol. 19, December 1992.

[53] See Victor Callaghan, "The image File Format Mess: What's Your TIFF?" *Advanced Imaging,* March 1994, pp. 44-47,85, for a discussion of the different file formats and the difficulties of working with so many.

[54] See also U.S. Congress, Office of Technol ogy Assessment, *Data Format Standards for Civilian Remote Sensing Satellites*, (Washington, DC: OTA, International Security and Space Program, May 1993) for a discussion of data format issues. This background paper highlights incompatible data formats and storage media, concluding there is no central organization positioned to develop and implement data standards.

[55] Originally chaired by a representative of the USGS, in fiscal year 1994, Bruce Babbit, Secretary of the Interior took over chairmanship of this committee. His chairmanship demonstrates how important the Department of Interior considers the issue of data standards.

BOX 2-6: Remote Sensing and the Global Positioning System

The ability to determine position quickly, simply, and with high accuracy from the Global Positioning System (GPS) constellation of satellites has markedly enhanced the use of remotely sensed data in geographic information systems. Developed by the Air Force for Department of Defense uses, GPS has found extensive use among a wide variety of civilian users.

GPS consists of three major segments: 1) a constellation of 24 satellites orbiting at approximately 11,000 nautical miles above Earth, 2) a mission control segment; and 3) a user segment, consisting of individual fixed and portable GPS ground receivers. GPS satellites have a 12-hour repeat cycle and an approximate 60-degree equatorial spacing. Between six and 11 satellites are always visible 5 degrees or more above the horizon. Ground receivers determine positions by measuring the signal travel time from four or more satellites, Civilian users can achieve position accuracies within 100 meters of the true geographical position. Better positioning (a few meters) is possible by using a correction signal from a ground-based differential position receiver/transmitter. The differential position recewer/transmitter, located at a known position, receives GPS signals from the satellites and calculates a correction, which it then broadcasts to GPS receivers within its range. Users with properly equipped GPS receivers can then automatically correct their calculated positions.

GPS can be used to geocode SPOT or Landsat imagery, and also to find sites within a study area. The merger of satellite imagery and GPS has proved invaluable in creating and updating GIS databases quickly and accurately. Doing so has saved time and money compared to using traditional methods such as field surveying and aerial photography.

SOURCE Off Ice of Technology Assessment, David A Turner and Marcia S Smith, *GPS Satellite Navigation* and *Positioning and the DoD's Navstar Global Positioning System,* 94-171 SPR (Washington, DC Congressional Research Service, Feb 15, 1994)

for GIS within the federal government are compatible. By coordinating with the user community and software developers, this committee has enabled the federal government to adopt a transfer standard.[56] As a consequence of this committee's work, as of February 15, 1994, all federal agencies must purchase and use computer hardware and software that allow the transfer of information among dissimilar systems. Vendors of spatial data processing software must comply with what has become known as Federal Information Processing Standard (FIPS), and can turn to the FGDC for advice. The vendors are now developing hardware and software to comply with the FIPS. USGS has designed a spatial data transfer processor to support transfer of data between various formats.[57]

GIS software has also suffered the drawback of requiring specialized training to use it effectively. Although GIS software is becoming more *'user friendly," most users still require intensive training.

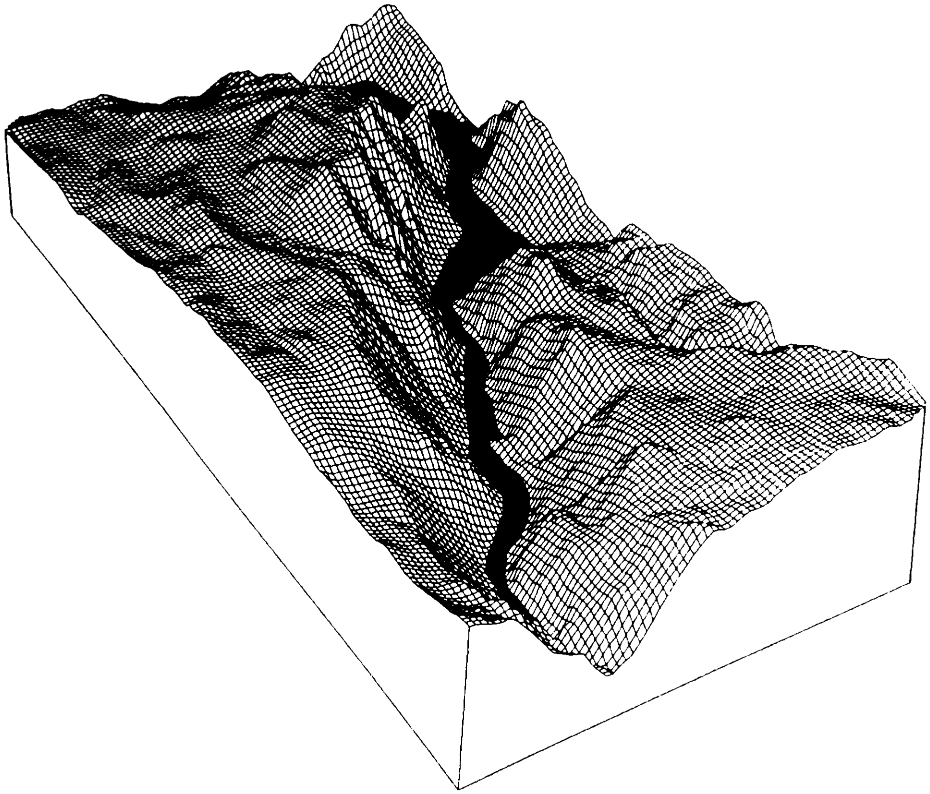
NEW WAYS OF VISUALIZING DATA

People are often more adept at extracting information from properly prepared images than from text, and sometimes capable of culling information from a series of images or video. Increasingly, therefore, researchers use spatial data in video or simulated video. In addition, remotely sensed data sets can be used to simulate (in three-dimensions) landscapes and geologic formations. These too, can be used in video representations of spatial

[56]The Secretary of Commerce approved the Spatial Data Transfer Standard, known as a Federal Information Processing Standard 173 in July 1992.

[57]More information regarding Federal Information Processing Standard 173 and its implementation is available from the Federal Geographic Data Committee, within the USGS.

FIGURE 2-8: Terrain Analysis Using Geographic Geographic Information System Technology



The U S Geological Survey prepared this dlgltal elevation model of the Hetch Hetchy Canyon in Yosemlfe National Park, California using aerial photographs

SOURCE U S Geological Survey, 1993

data. Data collected by remote sensors, when joined with digital terrain models, have been used to simulate airborne approaches to airports, geologic formations, and to produce topographic models of cities and countries around the world (figure 2-8). Converting spatial data to three-dimensional scenes or video presentations is useful because:

- three-dimensional and video representation allow researchers to visualize more information than with two-dimensional still images;
- three-dimensional and video representation provide a sense that the viewer is within the landscape and moving through it;
- three-dimensional and video data sets increase the potential application of remotely sensed data—more stimulating presentations of spatial data make the prospect of commercial use of data more likely.

Using information in three-dimensional format requires increased processing capacity and large amounts of data storage. Many current users of three-dimensional imagery rely on advanced computer work stations (computers that can perform hundreds of million instructions per second, or MIPS). Methods of employing spatial data that were once regarded as frivolous have gained broad acceptance because of the value of visualizing data. For example, the interest in "virtual reality," or "computer simulation of reality," has become more widespread than it was several years ago, in part because image intensive applications have demonstrated the value of high resolution data display.[58]

[58] See "The Third Branch of Science Debuts," Science, vol. 256, Apr. 3,1992, pp. 44-47.