

A STUDY OF SPATIAL DATA MANAGEMENT
AND
ANALYSIS SYSTEMS

Clyde Christopher
Jackson State University
Jackson, MS

Richard Galle
Stennis Space Center
NSTL, MS

ABSTRACT

The Earth Resources Laboratory of NASA's Stennis Space Center is a center of space related technology for earth observations. It has assumed the task, in a joint effort with Jackson State University, to reach out to the science community and acquire information pertaining to characteristics of spatially oriented data processing.

INTRODUCTION

During the past 15 years much computer software has been developed for handling spatial data. A number of centers have been set up for processing spatial data, all with different configurations. Some studies have been conducted to collect and disseminate data on software in this area. Scientists at the Science and Technology Laboratory were interested in knowing the software and hardware characteristics in such centers. Additionally, they wanted to know what centers those who are actively engaged in spatial data processing are communicating with. This information is to be shared with scientists in those centers so that all will be fully aware of the state-of-the-art software and hardware for spatial data processing.

METHODOLOGY

The desired information was collected by means of a survey. The survey instrument was designed by Christopher and Galle with approval from proper authorities. Part I asked for background information including names of administrators, primary mission, list of principal application areas, and a list of representative projects currently supported at that facility.

Part II requested data processing methodology. First, it asked for a listing of hardware devices devoted to spatial data processing including the host computer, display devices, digitizers, disk drives, tape drives, plotters, etc. Next, it asked for a list of software packages currently supported by the facility, including compilers, GIS systems, statistical packages, data base management systems, etc. Finally, it asked with

what other facility they were linked and what communication package is being used.

Part III is a request for data characteristics. The survey asks what types of data are used (Landsat, TM, AVHRR, Soils Maps, etc.), what are the data sources, how is it input for processing, what is its internal format (Raster, Vector, etc.), and what CAD related capabilities are available for editing the data sets.

Part IV asked for data analysis characteristics. It specifically asked what types of data processing (statistical, analytical, expert system decision making, polygon declaration, model structuring) supports the applications at that installation. Also, is there merging or overlaying of data and what steps are taken in the processing life cycle.

Part V, the concluding section, asked for representative research which has been undertaken at the facility in the area of spatial data processing and, also, the research objectives which are needed.

The survey was first mailed under a cover letter dated July 10, 1987. A second copy was mailed to those facilities that had not responded by September 8, 1987. The survey was mailed to 349 different installations. Completed forms were received from 115 installations. A survey instrument worded so that all answers could be given by checking the correct item probably would have yielded a higher number of responses. However, that type of instrument would not yield as much detailed information as the one used.

The data will be used to build a data base with appropriate query language so that local users will have access to it. Hopefully, problems in communication will be solved providing access to this data on a nationwide basis.

RESULTS

Background Information

The survey was sent to basically three types of organizations: 1) government agencies, 2) private industry, and 3) educational institutions. In addition to supplying names of key contact personnel at the installation this section of the survey provided us with a statement of the primary mission of the installation and a list of ongoing projects. The government and industrial organizations said they were involved in research in remote sensing and in analyzing terrain and giving technical advice on use of land and water resources. The educational institutions are involved in training students in the use of remote sensing as well as research in that area.

No two installations listed the same project. Typical projects were: "Land use of 13 county region of Tennessee", "Training in forestry and wetlands remote sensing techniques", and "produce 7½ foot wetland maps of the United States". Some indicated that project titles are not available to the public.

The principal application area is an important response in this section. These ranged from "land use classification" to "Geological engineering". However, the most frequent responses were "remote sensing", "image processing", "GIS", or some combination of these.

Data Processing Methodology

Computer systems used to process spatial data range from microcomputers to mainframes. The most frequently listed microcomputers are the IBM PC-AT and the APPLE II PC. In the minicomputer class the VAX 11/780 is most used. In the largest class an Amdahl V7/V8 was listed by one installation. The prime 9650 was listed by several installations.

Table 1. Host Computer(s)

<u>Model</u>	<u>Number of Installations</u>
Microvax II	15
Apple IIc	2
Amdahl (5860/V7/V8)	3
Harris (1000,500,800)	3
VAX 98200,11/730,11/750, 11/780,11/785)	36
Hewlett-Packard 9000	2
IBM PC (ATXT)	12
SUN	3
PDP 11/70	3
Concurrent (Perking Elmer)	5
Prime (400, 550, 750, 465, 9755, 9650, 2655, 2275, 6350, 250, 9955)	
Data General	4
Gould 32/27	5
IBM (3013, 3081, 3084, 4365, 4381)	11
Zenith 248	3
MASSCOMP 5600	5
AT&T (3B5, 3B2)	3

Table 2. Operating Systems

<u>Software</u>	<u>Number of Installations</u>
RIPS	1
VMS	39
VOS 6.1	2
UNIX	11
PRIMOS	30
MVS	7
RT-11	3
MS-DOS	6
MPX	4

RSX-11M	2
OS-32	5
VM/CMS	3
VORTEX	2
AOS	2
ULTRIX	1

Table 3. Digitizers

<u>Model</u>	<u>Number of Installations</u>
GTCO	15
Hitachi	4
CALCOMP 9100	32
Altex	23
Geographics	5
Summagraphics	6
Perkin Elmer (concurrent)	1
Talos	5
NUMONICS	3
Integraph	2
Tektronix	7

Table 4. Display Devices

<u>Model</u>	<u>Number of Installations</u>
Tektronix	58
Comtal	11
De Anza	3
ADAGE	2
Visual 500	8
CALCOMP	2
Lexidata	6
Hewlett Packard	3
Inter Graph	2
ERDAS	3
Hitachi	3

The type of compilers available depends upon the model of computer and the memory size. Some installations list Fortran, only. Pascal and C are also frequently listed. Many installations listed all three of the above languages, but Fortran was listed most frequently. Other languages used included Lisp, Prolog, BASIC, SCAN, COBOL, Modula2, Ada, and PL/1.

Among the image processing and GIS systems listed, ELAS and ARC/INFO were most frequent. Others in the order of occurrence were RIPS, ORSER, GRASS, ERDAS, MOSS, SYNERCOM, and ODYSSEY.

Table 5. Image Processing Software

<u>Model</u>	<u>Number of Installations</u>
ELAS	30
RIPS	6
COS	3
ARC/INFO	45
GRASS	6
ERDAS	22
MOSS	22

SYNERCOM	2
ORSO	2
ATLAS	2
Inter Graph	2
EROS	2

Some installations indicated that their processing involves statistical analyses and they named the statistical package that is available on their system. SAS, SPSS, GLIM, BMDP, MINITAB, OSIRIS, NAG, VECTOR, MATH 77, SSPLIB, IMSL and Microstat are among the packages named.

Database management systems are as varied as statistical packages. However, some who responded did not name a DBMS. One installation stated that it has an inhouse information storage and retrieval system. The usual response was one or more of the familiar systems such as Datatrieve, INFO, PC INF, RBASE, Dbase III, ORACLE, BASELINE, CDOS, SMARTSTAR, SPIRES, or INGRES.

Table 6. Data Base Management Systems

<u>Software Package</u>	<u>Number of Installations</u>
DBASE II or III	20
Datatrieve	5
INFO	36
SMARTSTAR	1
ORACLE	7
BASELINE	1
RBASE	5
Prime-INFORMATION	4
SPIRES	1
OPS-83	1
CDOS	1
INGRES	4

Some large installations at government sponsored organizations indicated that they are connected with state or regional offices through some type of communication link. In general, the educational institutions have local networks. This was the case with several other installations as well, but there were those who indicated that they have no networking.

Data Characteristics

Landsat, TM and MSS data is the most used data in the installations that responded. However, a great variety of data types are used. AVHRR, Soils Maps, SPOT, and GOES were frequently named in the survey. Most installations named several types. Some other types listed were land use maps, USGS topographic maps, ACZCS, census, transportation networks, streams and rivers, watershed and aerial photos. The primary sources of data are EROS, NASA, USFS, USGS, SPOT Image Corporation, and virtually all map sources.

Table 7. Data Types

<u>Data Type</u>	<u>Number of Installations</u>
LANDSAT (MSS AND TM)	72
AVHRR	24
Soils Maps	41
Land Use Maps	12
Census	10
SPOT	21
GOES	6
SIR-B	2
USGS Topographic Maps	18

As a rule digitizers and magnetic tapes are used to input data for processing. Floppy disks were frequently listed as input media. Video cameras, optical disks, and keyboard were each listed by at least two installations.

Table 8. Data Input Scheme

<u>Input Media</u>	<u>Number of Installations</u>
Digitizer	84
Magnetic Tape	86
Magnetic Disk	21
Optical Disk	3
Keyboard	9
Scanning	6
Digital Camera	6

The internal format for data was given as raster or vector or both. CAD related capabilities at various installations include scroll, zoom, draw, density slice, classification, cut, paste, paint.

Table 9. Input Format

<u>Characteristic</u>	<u>Number of Installations</u>
Raster	86
Vector	72
Gridded	4

Data Analysis Characteristics

In answer to the question on analysis of data during processing, we found that a majority of installations do a statistical analysis. Several Installations use the data in decision making for expert systems. A few installations are doing polygon declarations. A small number indicated that they are merging data sets and at least one installation is overlaying data. Some installations were cooperative in sharing with us the steps in the processing life cycle, not an easy task.

CONCLUSION

In this section some installations listed some research efforts that are ongoing in special areas. A non-conclusive list of these follows:

- 1) Incorporation of the SOI-5 soil interpretation records into ARC/INFO
- 2) Incorporate spatial data (GIS) into a decision support system
- 3) GRASS G&D functions
- 4) Link between S and GRASS
- 5) Hydrologic modeling, flood damage analysis
- 6) Image texture recognition
- 7) Geometric rectification, image animation
- 8) Map display techniques
- 9) Digital image processing-both spectral and spatial analysis

(Notes. The researchers acknowledge assistance in this project from the following: Pauline Frances, System Analyst; Jeanette Lewis, Laboratory Coordinator; Andrew Ward, Kimberly Wallace, Rickey Myers, Jody B. Hasten, and Shelton James, students; Jackson State University.)