Reprinted from

Symposium on

Machine Processing of

Remotely Sensed Data

and

Soil Information Systems

and

Remote Sensing and Soil Survey

June 3-6, 1980

Proceedings

The Laboratory for Applications of Remote Sensing

Purdue University
West Lafayette
Indiana 47907 USA

IEEE Catalog No.
80CH1533-9 MPRSD

AN EXAMINATION OF REQUIREMENTS FOR A SOILS RESOURCE INFORMATION SYSTEM

DAVID L. ANDERSON, KIM L. STEVENS,
ROBERT D. HEIL

Colorado State University

I. ABSTRACT

An investigation was conducted into the necessary components, structure and feasibility of a comprehensive, user-oriented, Soils Resource Information System (SRIS). The approach to development involves four stages: (1) analysis, (2) pilot development, (3) prototype development, and (4) implementation. In the analysis phase, soil information users were interviewed to determine their information needs. A pilot SRIS has been developed utilizing five components: a soil map unit detabase, soil interpretation database, Pedon Characterization database, climatology database, and a range database. The pilot is controlled by SYSTEM 2000, a database management system. This system features interactive query and response, and a structured English query language. The ability to answer ad hoc questions posed by a user audience has stimulated interest and provided an awareness of the potential of SRIS.

Results of interviews as they were used in the development and testing of a pilot system and an evaluation of the opportunities/limitations for developing a comprehensive SRIS are presented in the paper.

II. INTRODUCTION

The soils of the United States are an essential but limited resource. Planning for the effective use and conservation of this resource requires access, integration and analysis of a wide variety of soils, climatic, and soil-related data. In an effort to satisfy the requirements of many users of soils information, numerous data collections have been established throughout the U.S. These sources are maintained by scattered agencies, industries and individuals at various national, regional, and local levels. In the past few years, we have seen a draftic increase in interest in soils information systems. Numerous authors have addressed the subject.[2,4]

The major source of soils information in the United States is the National Cooperative Soil Survey.[3] This source includes:

1. Published Soil Survey Reports - This data provides an identification of individual soil types, their areal extent, chemical and physical characteristics, and interpretation of soil behavior under various uses.

2. Soil Interpretation Record, (SCS-Soils-5)[3] - These records provide information on estimated engineering properties and soil interpretation concerning limitations of soils for various uses.

3. National Pedon Data System of the Soil Data System[1] (Pedon) - This is designed to store and retrieve characterization studies of Pedon sites. This site specific horizon-by-horizon description includes chemical, physical, mineralogical and manipulative laboratory tests on selected soils.

Various other sources of information exist which are used by planners and decision makers. Among these sources are soils data collected and stored at various universities, and climatic information collected by weather stations. Because of the fragmented nature of soils information, problems exist which involve the availability of and access to existing data. One of the concepts involved in the developing of a soils information system is the linking of these various sources of information so that questions concerning different data sources can be answered more efficiently.

In an effort to address the above described soils information management problems, the U.S. Department of Agriculture, Soil Conservation Service (SCS) sponsored a cooperative project with the Colorado State University, College of Agricultural Sciences, Department of Agronomy, and Laboratory for Information Science in Agriculture (LISA).

III. PROJECT GOALS

The goals of this project are to determine the necessary components, structure and feasibility of a comprehensive, user-oriented Soils Resource Information System (SRIS). To achieve these goals, a project team was established consisting of a Soil Scientist (working closely with other agency and university soil scientists), an Information System Analyst, and a Computer Programmer.

IV. SRIS DEVELOPMENT APPROACH

The general approach to the development of a Soils Resource Information System involves four stages: (1) analysis, (2) pilot development, (3) prototype development, and (4) implementation.

A. ANALYSIS

The analysis stage is directed at discovering user needs. To accomplish this, the user community was segmented into groups according to their general application of data.

A representative(s) from each group was interviewed by the project team. The objectives of these interviews were to: (1) determine current and/or potential data requirements and applications of data (2) identify data sources and (3) identify desired features, capabilities and components of a Soils Resource Information System. During the first phase of the project, 42 users were interviewed.

A summary of the backgrounds or fields of expertise of the interviewees is shown in Table 1. As indicated, interviews were folicited on a broad discipline base. The soil scientist discipline was further broken down into subfields: (1) consulting soil scientists working primarily in the private sector, (2) research soil scientists, and (3) "Agency" soil scientists, primarily Soil Conservation Service people involved in the collection of soil survey information.

Table 1. Summary of Interviewee Discipline

Interviewee Discipline	# Interviewed
Agronomist	3
Automated Data Processing	5
Climatologist	1
Conservationist	1
County Extension Agent	1
Land Use Planner	1
Natural Resource Specialist	2
Range Scientist	2
Reclamation Specialist	1
Sociologist	1
Soil Scientist	
Consulting	3
"Agency"	13
Research	7
Watershed Specialist	1
TOTAL	42

This list does not represent all disciplines requiring soil information or all those interested in using SRIS. Interviews are continuing in the prototype phase. This list represents those groups interviewed at the time this paper was written.

A summary of the interviews indicates the following data sources and the percentage of requests by interviewees: Fifty-three percent of those questioned needed soils interpretive data obtained from SCS-Soils-5; 50% needed information available from the National Pedon Data Soil Subsystem; 22% needed information on climatology and 20% needed information from other sources. Fifty percent of those questioned needed information from two or more of the sources previously mentioned and various other information such as plant community information, satellite imagery and census of agriculture. Although these are valuable sources of information, it would not be practical to link all these sources with SRIS at the present time because of time and money constraints.

The interviewees indicated that the following features would be desirable for a user-oriented information system: (1) information should be accessible from many remote locations where large computer facilities are not available, (2) the system should be simple enough to be used by persons with little or no computer background, (3) costs should not be prohibitive, (4) ability to access information from different data sources, (5) the capability to examine data and easily update erroneous or out-dated information, (6) a central repository of data gathered by various suppliers, (7) a centralized source of information on available soils data, and (8) graphics capability for displaying soil's boundaries and other resource information on the computer.

B. PILOT

The major objectives in the development of the pilot were to: (1) demonstrate to soils information users possible features and capabilities of a full-scale SRIS, and (2) experiment and evaluate preliminary system design alternatives. The product of this phase of the project was the development of a computer-based pilot Soils Resource Information System. This pilot is a small-scale model of the intended system and contains a limited amount of data for two Colorado counties.

Components of the Pilot. Using information obtained from user input and knowledge of existing data stores, a pilot system was developed at Colorado State University using a CDC 172 computer. The following components (or databases) were developed (Figure 1).

1. Soil Map Unit Database (Appendix 1)- consists of specific information on individual soil types. The source of this data is the Soil Survey Reports.

2. Soil Interpretation Database- The source of this data is the SCS-Soils-t.

At the present time only the Estimated Properties section of the SCS-Soils -5 is included in the pilot. As time and funding become available, interpretations for various uses will be added.

3. Pedon Characterization Database - Includes the chemical and physical data (SL-Record) of the National Pedon Data Subsystem.

4. Climatology Database - Consists of monthly summaries for national and state weather reporting stations.

5. Range Database - Consists of native (climax potential) vegetation and percentage as well as other information relating to range.

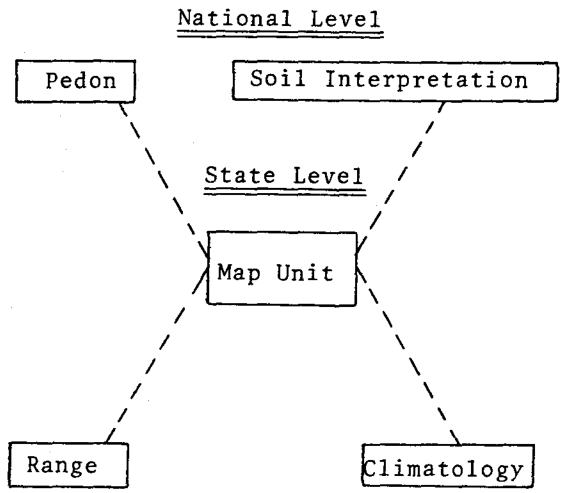


Figure 1. SRIS Component Relationships

Summaries of user interviews indicate that an information system containing these requests directed by this user group.

Relationship of the Databases. Two primary levels of responsibility exist with respect to Soils Information (Figure 1). The National Pedon Data Subsystem and the Soil Interpretation Record, SCS-Soils-5 are nationally centralized data stores. The map unit, and climatic and range data have primary responsibility at the state level. Responsibility for maintaining the database should most likely be maintained at these levels. The bulk of the state level data is site-specific and must meet the needs of a diverse community of users including farm and ranch operators, land planners, researchers and consultants. The size of the databases required to effectively meet these needs inhibits the organization and centralization at a national level.

Data Input. SRIS is designed to obtain maximum benefit from existing data stores. The Pedon, SCS-Soils-5 and climatic information presently exist in machine-readable mode. This information must be translated from its existing format and loaded into the Database Management System (DBMS) by a rather sophisticated program. The map unit and portions of the range components have to be entered into the DBMS manually. We have estimated that the manual loading of map unit (Appendix 1) components for completed soil survey areas in Colorado will take approximately 7.5 person months. This includes filling out the coding form, data entry, verification and editing of 2,600 soil map units. Time estimates have not been completed on the other components.

Structure of Pilot. The capabilities of Database Management (DBMS) processing satisfy most of the requirements identified during the analysis phase. A DBMS is a generalized data storage, retrieval, and management system with data centralization and interactive access capabilities. The DBMS SYSTEM 2000 marketed by MRI Systems, Inc. of Austin, Texas was selected as the DBMS to use for pilot development.

The DBMS features a structured English query language. Users can formulate a majority of their queries in an easy-to-use and simple to understand English sentence structure. With very little training, a user with no computer background can use SRIS without the aid of highly trained computer specialists.

To briefly illustrate the system's capabilities, the following example is given. Suppose we wished to know the map unit names which have sandy loam surface layers in Adams County. To quiry the system, we ask the questions as follows:

```
? ;
---
? list map unit name where surface texture eq sl
---
? and county eq adm;
```

The system output is:

```
  MAP UNIT NAME
***
* ASCALON SANDY LOAM, 1 TO 3 % SLOPES
* ASCALON SANDY LOAM, 3 TO 5 % SLOPES
* ASCALON SANDY LOAM, 5 TO 9 % SLOPES
* ASCALON-PLATNER ASSOCIATION
* ASCALON-VONA SANDY LOAMS, 1 TO 5 % SLOPES
* TERRY-VONA-TASSEL COMPLEX, 3 TO 20 % SLOPES
* TRUCKTON SANDY LOAM, 1 TO 3 % SLOPES
* TRUCKTON SANDY LOAM, 3 TO 5 % SLOPES
* TRUCKTON SANDY LOAM, 3 TO 9 % SLOPES
* VONA SANDY LOAM, 0 TO 3 % SLOPES
* VONA SANDY LOAM, 1 TO 3 % SLOPES
* VONA SANDY LOAM, 3 TO 5 % SLOPES
---
```

To effectively address the needs of a diverse user community, information needs to be stored at a detailed level. The soil map unit, with its unique combination of soil taxa, landscape and natural vegetation, is the basic access component in SRIS.

A three-digit abbreviation, identifying the soil survey area and the soil map unit symbol, constitute a unique identifier of map units in SRIS. Once these two symbols are known, access to the system can be achieved.

The map unit database is designed using two major levels (Appendix 1). The first level contains information which pertains to the map unit. An example of this would be acreage or capability class. These are both reported for the map unit, not the eaxonomic unit. The second level contains information which pertains to the taxonomic unit, e.g., soil series. This level repeats for each taxon within a map unit. An example of information stored at this level would be surface texture and slope. Both are properties of the taxonomic unit.

Contained in the map unit component are links to other databases which contain information pertaining to the map unit or taxon. For example, to find out the permeability of the Ascalon soil, you first ask for the number that identifies the Soils-5 record which contains the information needed. This is done through an interactive query. Then the question is asked as follows:

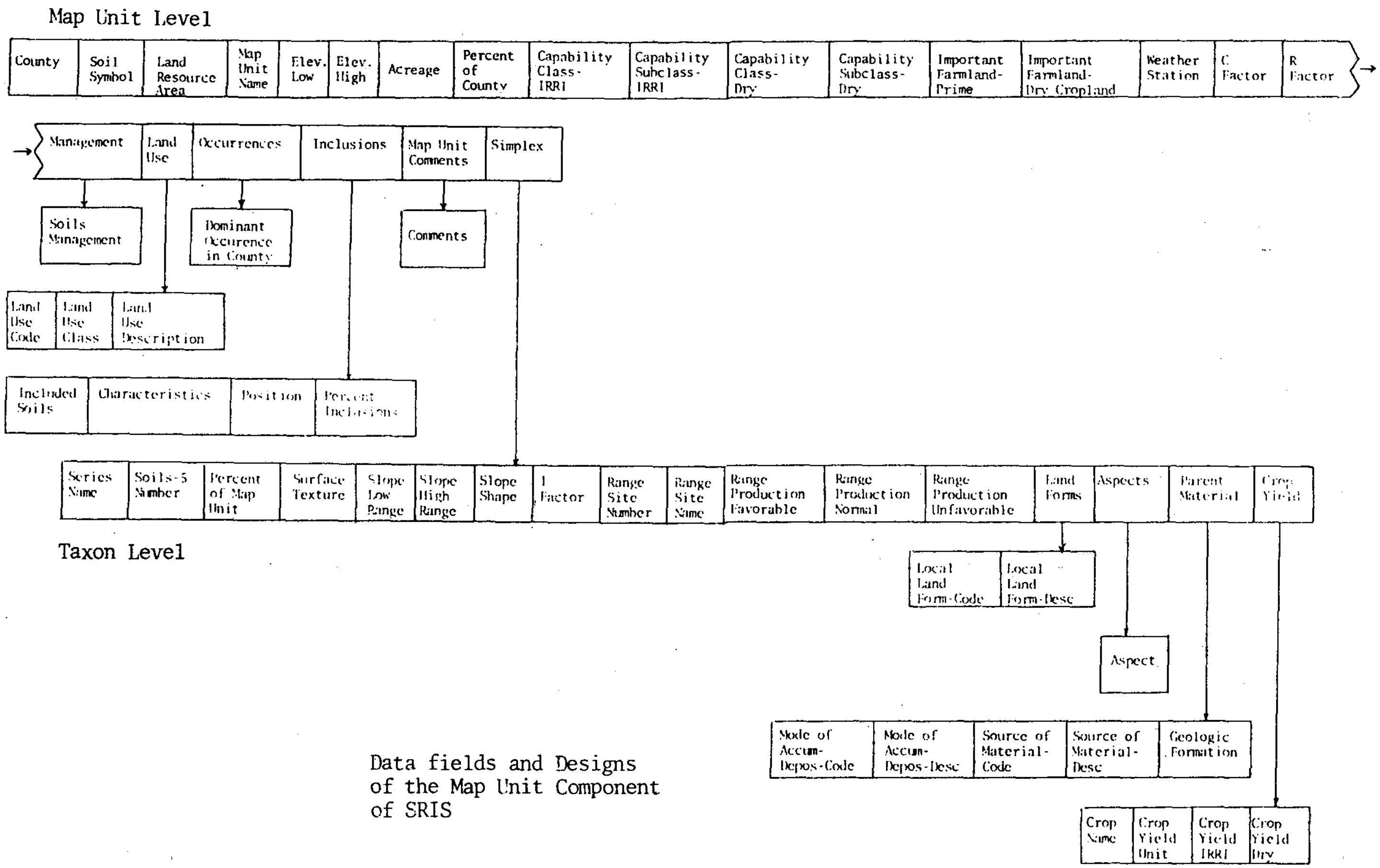
```
---
? list depth low, depth high, permeability low,
---
? permeability high where soils-5 number eq co0003;
```

The system output is:

```
    DEPTH LOW   DEPTH HIGH   PERMEABILITY LOW   PERMEABILITY HIGH
***
*       0            7              6.00                0.20
*       0            7              2.00                6.00
*       0            7              0.60                2.00
*       7           18              0.60                2.00
*      18           25              0.60                6.00
*      25           60              2.00                6.00
---
?
```

This link may seem awkward, however, it is an essential step. This is illustrated by the following example. In Colorado there are sixty-four map units which have Vona soil as a named taxonomic component. To store information at the map unit level, using conventional storage techniques, we would have to store information pertaining to Vona soil sixty-four times. By utilizing separate but linked databases, we store specific information for each taxon once and access the information each time the taxon is used in a map unit. It is easy to see the advantage to this system in terms of data input, storage and updating.

Several problems exist which must be addressed during further analysis and design of SRIS. One of the problems involves the linking of databases. Another



Data fields and Designs
of the Map Unit Component
of SRIS

Appendix 1

problem which remains to be solved is retrieving specific map unit information from the Soils-5 record. One Soils-5 record may have several surface textures and slope phases. To access this information for a specific map unit would involve an additional step which the Map Unit Record SCS-SS-6 form now performs. (This form is used to access information from the SCS-Soils-5.)

The weather station number if the link between the map unit and the climatology database. The number of the weather station which most closely represents the climatic conditions of a map unit is stored as a data field within the map unit component. This number is used as a link between the database in the same manner as the Soils-5 record number was used in the previous example.

The soil series name would perform the linking function between the map unit and the National Pedon Data Subsystem database. A code combining the range site number and an abbreviation for the range site name would be the link between the map unit and the Range Database.

C. PROTOTYPE

Prototype creation is the third major stage in SRIS development approach. The information accummulated during the preceding analysis and pilot stages provides a basis for prototype development. The prototype represents a best estimate of the structure, organization, and content of the SRIS.* The goal of the prototype is to provide a working system for the user community to access for a portion of their soils information needs. The most important output of the prototype stage is the system evaluation by users. The evaluation may indicate that major facets of the prototype have to be modified to satisfy user needs. In this case, the prototype would be redesigned and the prototype stage repeated for another evaluation period. Often, however, the user evaluation indicated a general satisfaction with the prototype model along with minor suggestions for improvement. Minor alterations can be made through "tuning"; the prototype then becomes the production system with intensive data loading taking place during the implementation stage.

*The prototype is a full-scale model, but it contains a small subset of the intended data for the full operational system.

D. IMPLEMENTATION

The implementation approach is done in phases and uses a participatory element. The phases allow a "bootstrap" development whereby the required expertise , knowledge, and evaluation can be accumulated prior to incurring the large-scale costs and committment necessary for full-scale implementation and operation. The participation aspect insures that the system is user-oriented, and will be used when it becomes operational.

V. REFERENCES

[1]Anonymous, 1979. Pedon Coding System for the National Cooperative Soil Survey. USDA Soil Conservation Service, July 1979.

[2]More, A., and S. Bic, (Eds.) 1976. Uses of Soil Information Systems. Proceedings of the Australian meeting of the ISSS working group on Soil Information Systems, Canberra, Australia, March 2-4, 1977.

[3]Resource and Management Information Task Force Report, 1978. A progress report prepared by the Resource and Management Information System Task Force.

[4]Sadouski, A., and S. Bic (Eds.) 1978. Developments in Soil Information Systems. Proceedings of the second meeting of the ISSS working group on Soil Information Systems. Varna/Sofic, Bulgaria, May 30-June 4, 1977.

[5]Stevens, K., E. Jukkola and D. Anderson, 1980. Soils Resource Information System, Technical Report 1.1, USDA. Prepared for USDA Soil Conservation Service by the Laboratory for Information Science in Agriculture, Colorado State University, Fort Collins, Colorado. 99 pages.

[6]Stevens, K., D. Watt and D. Anderson, 1980. Soils Resource Information System, Project Report 1. Prepared for USDA Soil Conservation Service by the Laboratory for Information Science in Agriculture, Colorado State University, Fort Collins, Colorado. 106 pages.

David L. Anderson
B.S. in Agronomy, Oklahoma State University, 1971. Soil Scientist-survey party member with USDA/SCS for 4 years. Party Leader for Kiowa County Soil Survey 2 years. Presently Soil Scientist at C.S.U. in Soils Resource Information Systems. Outstanding Performance Award 1973, Certificate of Merit 1976. Publications include: "Soil Survey of Kiowa County, Colorado" 1977 and "Soil Survey of Logan County, Colorado" 1977.

Robert D. Heil
B.S., M.S., and Ph.D. in Soil Science from South Dakota State University. Soil Scientist for SCS in North Dakota 4 years. Extension Testing Specialist at C.S.U. 5 years. Research Associate at Remote Sensing Institute, South Dakota. Over 40 publications in soil fertility and management, soil genesis, survey and classification, and land use planning. Presently Center Council Chairman for Environmental Resources Center.

Kim L. Stevens
B.S. in Computer Science from C.S.U. in 1973. B.S. in Statistics from C.S.U. in 1976. Working toward M.S. in Management Information Systems at C.S.U. Currently a Systems Analyst for LISA applying computer technology and management science to information systems problems in the agricultural sciences. Previous experience as a scientific programmer building simulation models and information systems for grassland ecosystems.