NATIONAL SOILS DATA BASE FROM THE SOIL CONSERVATION SERVICE

C. Gene Johnson
Assistant Head, Cartographic Staff
Soil Conservation Service
U.S. Department of Agriculture
10,000 Aerospace Road
Lanham, Maryland 20801

Increasing demand for data on soils and their geographic location has created a need for a fast, efficient way of converting soil information into a computer-readable form. A computer graphic system called the Advanced Mapping System (AMS) was designed to meet this need for the Soil Conservation Service (SCS).

One of the missions of the Soil Conservation Service is to map the soils of the United States. The soil survey supplies information that is basic to many decisions by people who own or manage land and water resources. Most soil surveys cover one county, but many cover more than one entire county or parts of one or more counties. As of April 1, 1979, 1,161 SCS soil surveys were mapped according to modern standards. Mapping of the entire Nation is scheduled for completion by 1996.

Soil Scientists prepare a comprehensive soil map that can be interpreted for many different uses. Most users, however, are interested in only one or two of the soil characteristics for any specific interpretation.

In 1967, the SCS Cartographic Staff began developing a computer mapping system to provide soil information in digital form. On June 30, 1978, the final acceptance test was completed and the system was accepted. This paper describes the procedure, the system hardware and software, cartographic characteristics, and soil data base characteristics.

Objectives and Application of AMS

The primary objectives of the AMS are to digitize soil data from published soil surveys and to create a national soils data base. This data base contains soil boundary coordinates and the soil symbols. A secondary objective is the production of interpretive maps. Basically, the Advanced Mapping System uses complex soil data to generate maps that are easily understood.

Thematic maps generated by the system, such as those used in SCS river basin studies, watershed projects, geologic investigations, and resource conservation and development areas, are used in planning activities by SCS personnel. The soil data will be used to show the degree of limitations, potential, or productivity of the soils in a town, watershed, or county.

Other information can be digitized for interpretive maps of such features as vegetative cover, range sites, land ownership, geology, topography, and climate.

Procedures for AMS Map Production

In making a soil survey, soil scientists in the field delineate soil boundaries on aerial photographs that have been rectified and scaled or more recently, on orthophotographs. In the SCS state office, the soil boundaries are checked and professionally redrawn on mylar overlays. One overlay shows the cultural and drainage features, a second the soil boundaries, and a third the soil symbols and place names. These mylar overlays are sent to the Cartographic Staff in Lanham, Maryland, for processing. The overlays are used both in printing the soil survey and in producing digitized maps through AMS.

The mylar overlays showing soil boundaries are placed on a digital scanner. The data are processed and, after scanner errors are identified and corrected, soil symbols are placed in each soil area. A drum plotter with a ballpoint pen plots the soil map on paper. An editor checks this sheet, makes the necessary corrections to the data base, and the system produces a corrected map sheet.

Most users want maps in a county format, however, so the individual soil map sheets are joined together and adjusted from the aerial photo base to a planimetrically correct base map. To make this adjustment, the correct base map is placed on the digitizing table and each soil map sheet is displayed on a graphic CRT. Then, vectors are drawn showing the movement of the soil map to fit the base map. The system can adjust 800 points at a time. After the individual soil map sheets are adjusted to fit the new base map, the system joins all the sheets to make one large soil map of the survey area. Border lines on the edges of adjacent sheets are removed and the resulting line segments joined if the segments are within a certain tolerance and if the two joining line segments have the same soil symbol.

An interpretive map, for example, a map of important farmlands, can now be produced from the planimetrically correct county soil map. The symbols for soils in each category of important farmlands are entered into a data file, that is, the soils on prime farmland and farmland of statewide and local importance. Through this step, only those soil lines are retained that form boundaries of the important farmlands. A magnetic tape is generated to plot these lines on an automatic drafting machine.

System Hardware

The system is designed around six minicomputer subsystems: one scanning subsystem, one processing subsystem, three editing subsystems, and an automatic drafting subsystem. Each subsystem can perform independently.

Scanning Subsystem -- This subsystem automatically scans mylar sheets at the rate of ½ to 4 square inches per second. The average size sheet is 10 by 15 inches and requires 6 minutes to scan. The flatbed scanner scans 1-inch-wide vertical strips, moving from left to right across the sheet. The scanner requires black lines approximately 0.01 inch wide on matte-surface mylar. Figure 1 shows the equipment configuration.

Processing Subsystem -- The processing subsystem takes the data from the scanner, converts the data from raster to vector format, joins the 1-inch-wide strips, compresses the data, and transfers them to the system data base. Figure 2 shows the equipment configuration.

Editing Subsystem -- Three identical editing subsystems are used. Each has two graphic CRT stations and one manual digitizing station. Each CRT station is assigned two tasks so that the operator can work on one sheet while the system processes another. Each graphic CRT has a data tablet used in editing the scanner data. Lines are displayed on the graphic CRT, and the operator uses edit commands to correct any errors unresolved by the scanner software. Figure 3 shows the equipment configuration.

Automatic Drafting Subsystem -- From a magnetic tape generated by any of the other subsystems, the automatic drafting subsystem is a flatbed automatic drafting machine with an optical exposure head that draws on photographic film with a beam of light. The optical exposure head has 24 apertures that can draw different line widths. Figure 4 shows the equipment configuration.

System Software

The system software records coordinate data for lines, symbols, and alphanumerics. The system can draw four lettering styles--uppercase and lowercase--and 64 special line types.

Editing -- Editing software can add, delete, or move a feature; correct alphanumeric data; change the end point of a line; connect one line to another; eliminate crossovers; change the line type; join feature parts from two sections of the map; and extract a section of the map by defining as many as 100 coordinate points.

Identification -- All data that are scanned are identified and stored on specific layers within the system. This permits the user to select various layers for producing a map. The operator enters all alphanumeric text into the system, using a keyboard or the menu to indicate the proper location of text units on the CRT.

Paneling -- The average soil survey has about 70 map sheets. To join the sheets together to form a large map, the sheets are placed in position on the graphic CRT and a processing step removes border lines and joins soil boundaries.

Warping -- Scaled aerial photographs with the tip and tilt removed are used to map soil surveys. Because of relief, however, some distortion remains. Before this distortion can be corrected, the soils information must be warped to fit the base map. We have developed the technology to digitize published soil surveys from distorted bases so that the new base is compatible with planimetrically correct base maps produced by other agencies. The operator digitizes points on the soil map from the graphic CRT. The same identifiable point is digitized on the base map on the digitizing table. The system, using as many as 800 warping vectors, then processes information to stretch the map until it fits the new base.

Polygons -- The system works with line segments, but to calculate area it must form and name polygons. After the data base is created, the entire map is checked to ensure that all soil areas are closed. The next command is to name each polygon.

Calculation of Area -- Area calculations can be performed after the map for a county or other survey area is in the digital data base and polygons are created and named. The operator requests the computation of area for the desired map units or interpretive groups.

Scaling -- The system can scale a map to any desired scale.

Color Separations -- Data files are created for each color separation. The file contains a list of colors and related soil symbols. Interpretive maps are prepared by combining similar soil types. The processing step looks at each line segment to determine whether the soil symbols on the left and right are in the data file for a particular interpretive group. If one of the soil symbols is in the data file, the line segment is stored for output because it forms part of the boundary between that interpretive group and another. If both soil symbols are in the data file for the same interpretive group, that line segment is ignored -- because it doesn't form an interpretive group boundary -- and processing continues.

Common Line Removal -- After all line segments are entered and identified, this processing step is required to remove from the data base all common lines between soil areas.

Cartographic Characteristics

Coordinate System -- All coordinate values of the soil lines are in table coordinates referenced to latitude and longitude, state coordinates, or Universal Transverse Mercator coordinates. This will allow the user to convert the data to his required coordinate system.

Mapping Scale -- Generally, detailed maps in the published surveys are at one of these four scales: 1:15,840; 1:20,000; 1:24,000; or 1:31,680. The scale of a county map produced through AMS should be no larger than the scale of the original soil survey. Some users enlarge a soil survey to 1:4800; this practice is not recommended, however, because SCS does not guarantee the accuracy of the soil lines at that scale.

We are using the 1:24,000 U.S. Geological Survey topographic map series as a base for our digitized soils information.

Maximum Ground Resolution -- Soil lines drawn in the field by a soil scientist do not have an absolute coordinate value. However, in a final prepared map -- whether in a published survey or an AMS product -- the soil line should not have been displaced more than 0.010 inch from where it was delineated by the soil scientist on the aerial photograph. At 1:20,000 scale, this 0.10 inch translates to 16 feet on the ground. In some soils it is very difficult to locate a boundary between two soils at less than this distance.

Cartographic Data Base Quality -- The accuracy of soil line coordinate values depends on the quality of the field mapping. Matching water areas to another base is a different kind of problem, however. Boundaries of water bodies are delineated on the aerial photograph as of the date the aerial photograph is taken. When the map is matched to another base map that was completed from aerial photographs of another year, the water boundaries do not always match. It is up to the user to correct the water boundaries if he wishes.

Data Base Characteristics

The soils data are on a magnetic tape that is to Federal Information Processing Standards; this tape is compatible with most other computers. Since all computer mapping display software is different, we have generalized our record format to: the soil symbol on the left, the soil symbol on the right, the number of coordinate values in the line, and the coordinate values defining the line segment from intersection to intersection.

The file is in ASCII but can be converted to EBCDIC if required. The FORTRAN format statement for the first record is (2A5, I5, 9X). The succeeding records will be (2(F6.3, 6X). The magnetic tape format will be 9 track, 800 bpi (see figure 5). The length of the file is variable, depending on the size of the area digitized. Each record contains 24 characters and the logical record length is 240 characters.

Processing Rate

The task of digitizing and entering soil line information into a data base has just begun, and no schedule has been established. Appproximately 150 soil surveys are published each year, but only 60 or 70 can be digitized each year with our present equipment and staff. A survey's publication and digitization dates will generally be different. At present, publication updates are planned every 25 years.

The processing rate of the system with present equipment is capable of twenty-four 10 by 15 inch soil map sheets per 8-hour day. Processing steps include scanning, editing, entering of symbols, warping, paneling, and preparing color separation overlays. The system has been used for production of interpretive maps and map finishing. The map finishing has taken much longer than anticipated and will not be done on the system in the future. Efforts will be made to digitize soils data for the data base and to prepare interpretive maps that can be overlaid on existing base maps.

Conclusion

More and more users are asking for soils data in digital form. Digital data are becoming a normal part of the soil survey. Map data should be integrated with the soil survey interpretive files for easy preparation of interpretive maps. Published soil surveys provide information that decision makers need for making sound decisions. With digitized maps produced through AMS, this information is easier to display on interpretive maps.

Scanner Subsystem

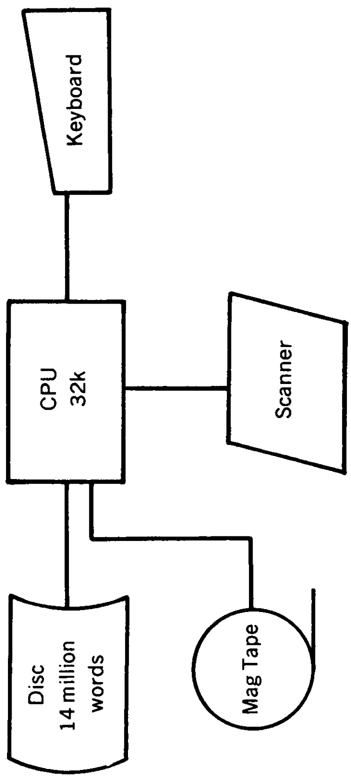

Figure 1

Processing Subsystem

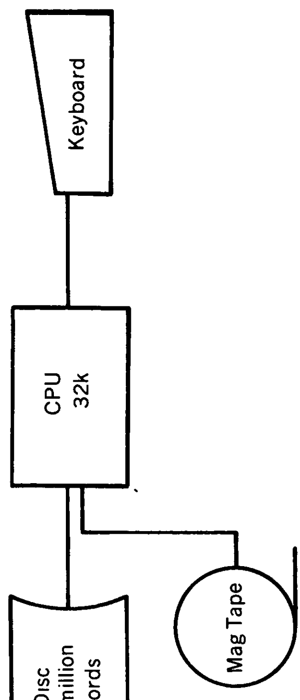

Figure 2

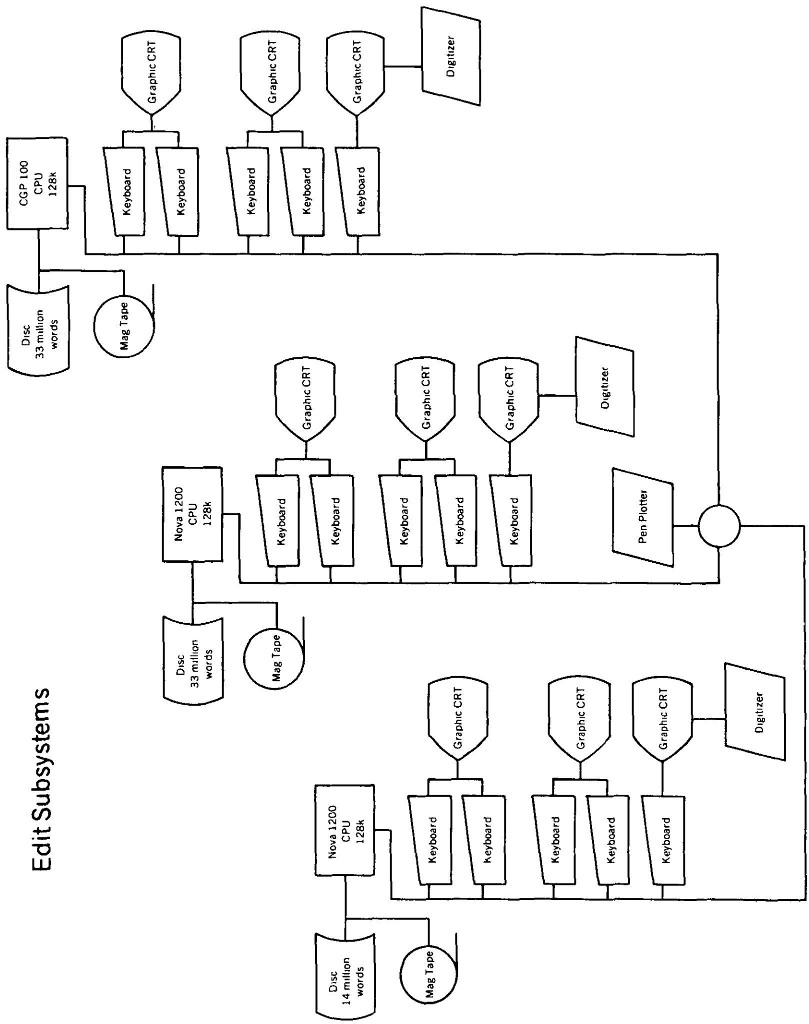
Edit Subsystems
CGP 100
CPU
128k
Disc
33 million
words
Mag Tape
Keyboard
Keyboard
Keyboard
Keyboard
Keyboard
Graphic CRT
Graphic CRT
Graphic CRT
Digitizer
Nova 1200
CPU
128k
Disc
33 million
words
Mag Tape
Keyboard
Keyboard
Keyboard
Keyboard
Keyboard
Graphic CRT
Graphic CRT
Graphic CRT
Digitizer
Pen Plotter
Nova 1200
CPU
128k
Disc
14 million
words
Mag Tape
Keyboard
Keyboard
Keyboard
Keyboard
Keyboard
Graphic CRT
Graphic CRT
Graphic CRT
Digitizer

Figure 3

Drafting Subsystem

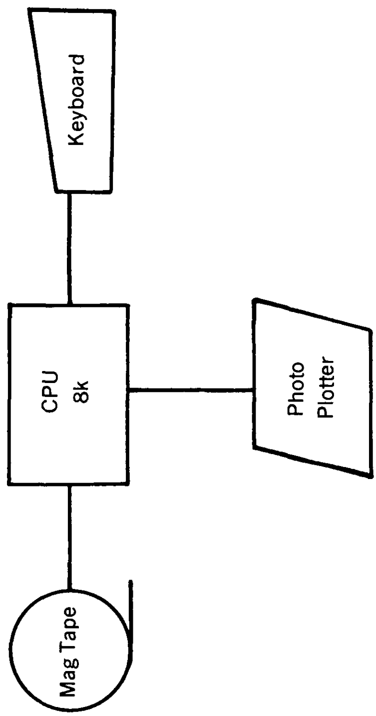

Figure 4

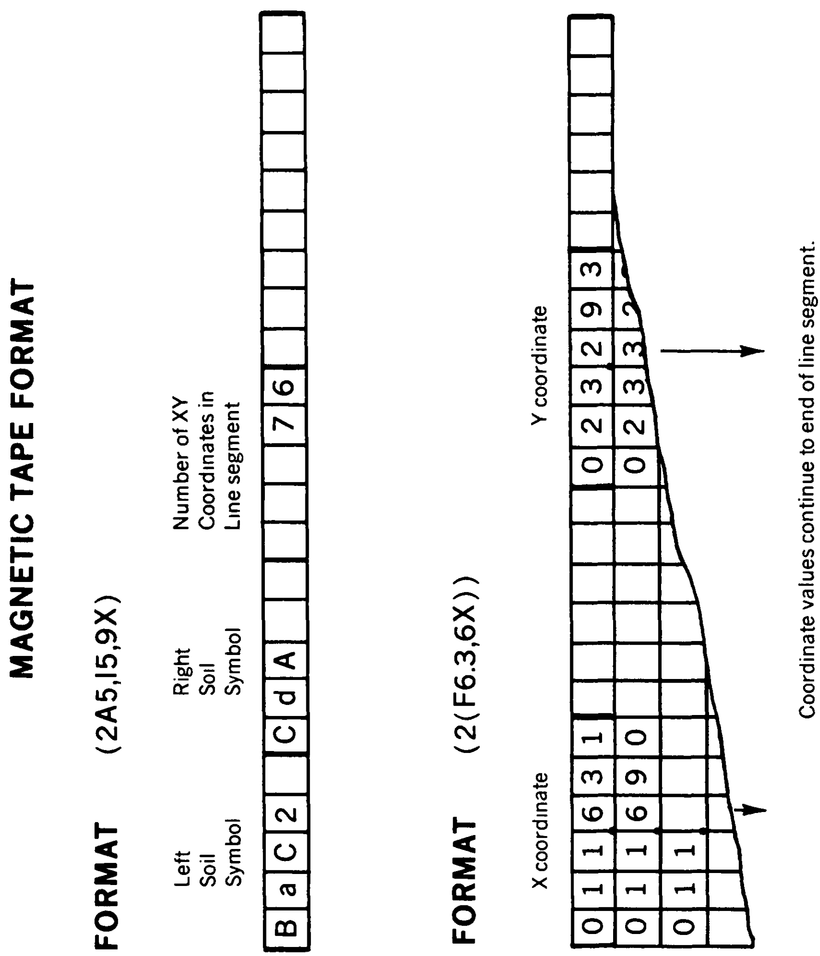
MAGNETIC TAPE FORMAT
FORMAT (2A5,I5,9X)
Left Soil Symbol
Right Soil Symbol
Number of XY Coordinates in Line segment
B a C 2
C d A
7 6
FORMAT (2(F6.3,6X))
X coordinate
Y coordinate
0 1 1 6 3 1
0 2 3 2 9 3
0 1 1 6 9 0
0 2 3 3
0 1 1
Coordinate values continue to end of line segment.

Figure 5