Technical Report No. 89

THE DESIGN OF A SPATIAL DATA FRAMEWORK CENTRAL BASIN WATERSHED, PAWNEE SITE

R. E. Oliver and L. D. Miller

Watershed Sciences Department

Colorado State University

Fort Collins, Colorado

GRASSLAND BIOME

U. S. International Biological Program

April 1971

TABLE OF CONTENTS

ABSTRACT

This report describes the attributes of a framework and analysis model procedure for spatially distributed data at the Pawnee Site. The needs and possible uses of the procedure are discussed along with a description of how some of the data is collected and utilized.

INTRODUCTION

Treatment of ecosystems as spatially distributed systems will require methods for data compilation, summarization, or display on a spatial basis. An adequate framework will serve as a base into which information collected and reduced by various investigations is stored or overlaid to be subsequently used by various spatial models (e.g., hydrology model, productivity model(s), etc.).

DATA CONTAINED IN THE SPATIAL DATA SYSTEM

The closed basin at the Pawnee Site is approximately 1000 acres in extent. The total area is subdivided into 10 ft × 10 ft areas (Fig. 1) or "cells" for the purpose of relating the essential features of the landscape. This size was arrived at as a result of a consideration of present and potential uses and data storage requirements and results in approximately 400,000 cells in the basin. The prime consideration in the selection of the cell size was that if it were any smaller it would result in an unmanageable total number of cells. The size selected is small enough to highly resolve the terrain, soil distribution, etc. from the biological viewpoint. It would have been preferable to use a cell size in the metric system, but this initial plan proved impossible as the topographic maps are in a scale of inches equated to feet and the available coordinatograph for reducing these maps reads only in decimal fractions and multiples of inches. The parameters which will be related to these cells initially include the following:

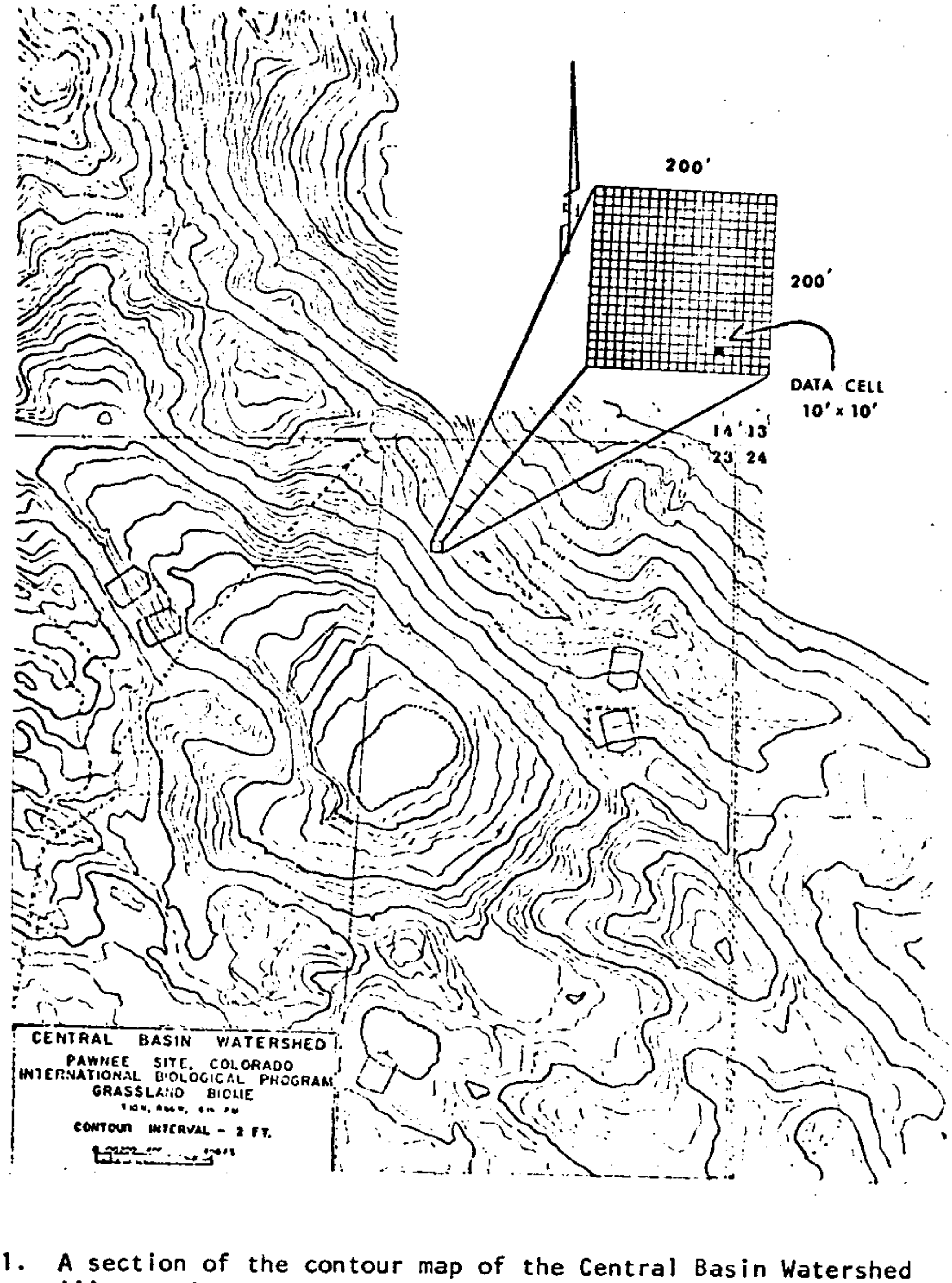


Fig. 1. A section of the contour map of the Central Basin Watershed illustrating the data cell size.

i. Topographic Characteristics (from 2 ft contour intervals maps)
   a. elevation
   b. slope
   c. aspect

ii. Soil type (from the soil type map)

iii. Soil moisture status at a particular time and depth (from interpolation of the readings from neutron access tubes)

iv. Vegetative type distribution (from vegetation map prepared from ektachrome infrared aerial photographs)

v. Percent cover distribution (initially from interpolation of a ground survey and subsequently from remote sensing measurements)

Additional data items can be added as they become available and as a need arises for their inclusion in the model. Overlays specifying dynamic aspects of the landscape, such as the aforementioned soil moisture distribution, will be added to the cells as they become available from remote sensing or ground survey sources. Up to 40 different features can be related to the 400,000 cells and stored on a single disk pack.

DATA

Topographic Data

Black and white aerial mapping photography of the approximately 1000 acres of the Pawnee Grasslands was taken in March 1970, and a contour map was compiled (scale: 1 inch = 100 ft, contour interval = 2 ft). Using these maps dimensionally, stable milar copies were made and are the source of the topographic data for the landscape model.

The forming of the data framework on a digital computer requires the conversion of the data on the available maps into a form which may be operated upon by the computer. This is being accomplished with the use of a coordinatograph, coupled to x and y analog to digital convertors and card punch located in the CSU Engineering Research Center Photogrammetry Laboratory. To accommodate the span of the coordinatograph, the stable base prints were made in sections, each 30 inches × 50 inches or less. Eleven maps of this approximate size are being digitized; each contains a 4 inch overlap and common control points for accurate correlation of the topographic data of adjacent maps.

The ground control points included in the photography and the contour maps were converted by hand to the coordinate system of the ultimate data base (Fig. 2) and scaled for digitizing. Two or more of these points of known coordinates are used for initializing the coordinatograph for the digitizing of each map. The control board for the unit is wired to record the x and y position of the points on the contour lines which lie on any of the family of parallel lines, $x = .1(I)$, where I is any integer. A data point is recorded upon the x translation of + .1 inch representing 10 ft on the ground (Fig. 3). Each of the contours is traced, and the x and y coordinates of the data points are recorded on computer cards. A source map identifier, a contour line identification number, and the contour elevation is recorded on each of the data cards (Fig. 4). The collected data forms a family of terrain profiles of the landscape in the y direction, each profile separated by an x interval of .1 inch on the map or 10 ft on the ground.

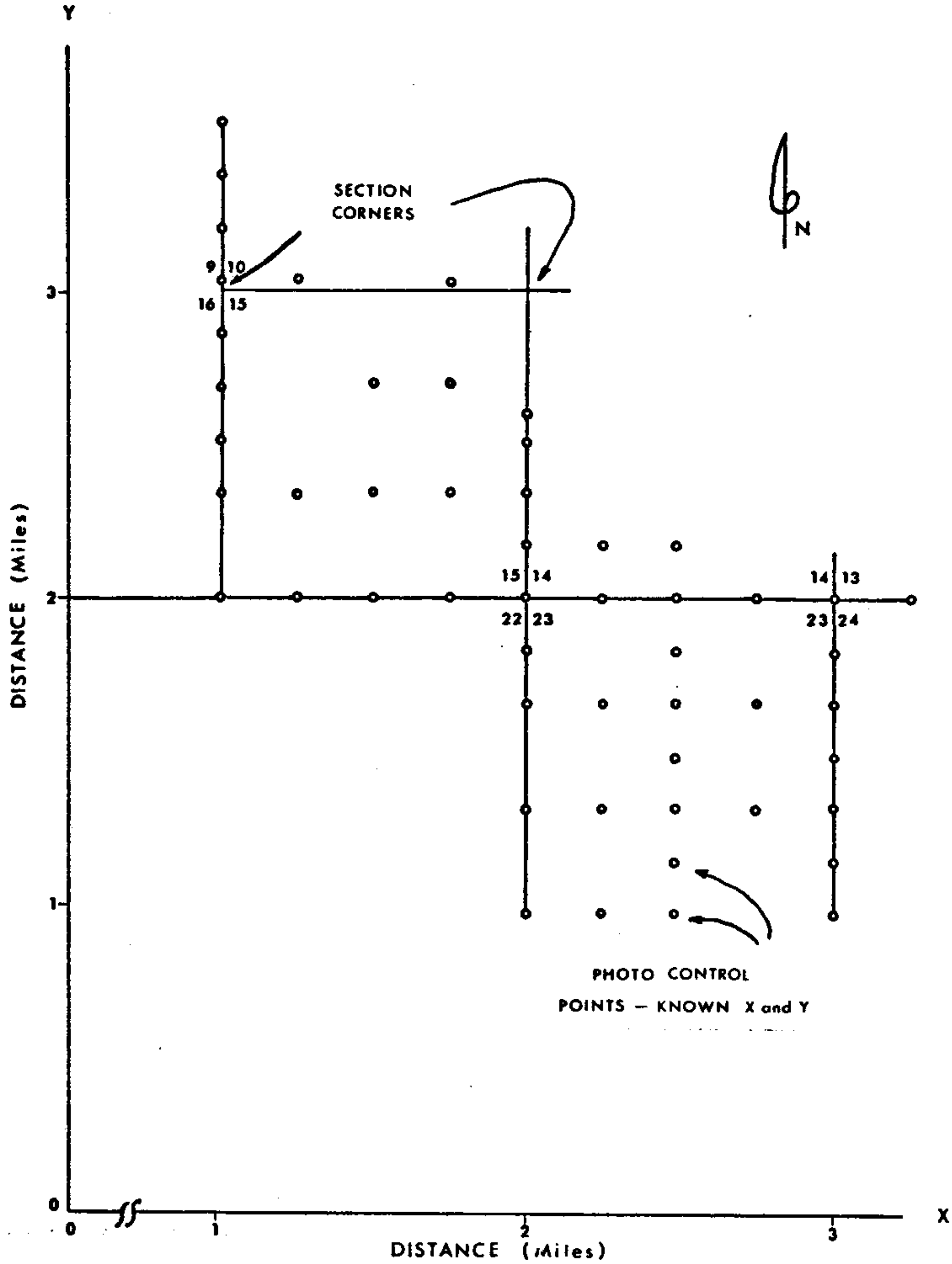


Fig. 2. The coordinate system of the data base at the Pawnee Site. The ground control points are shown.

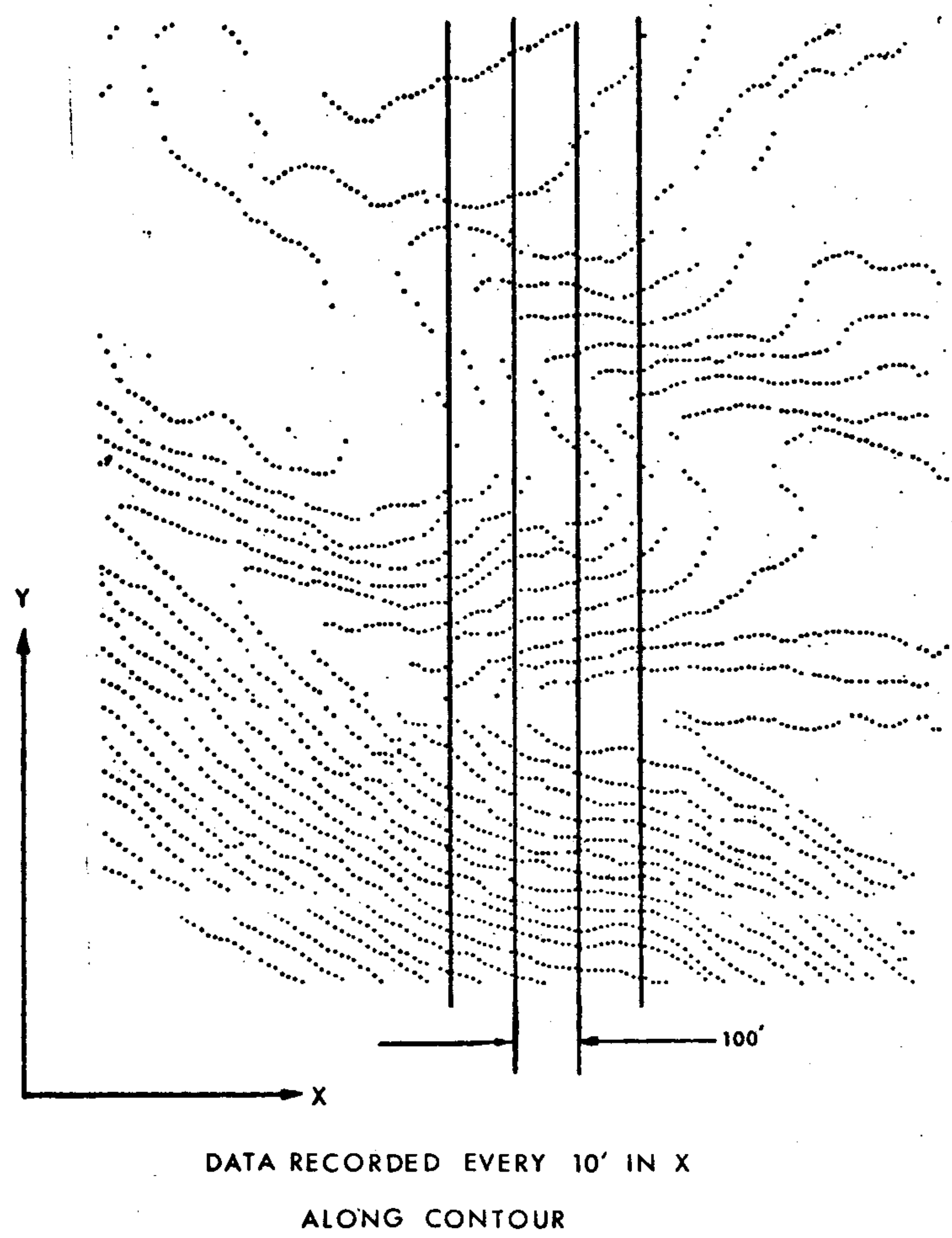
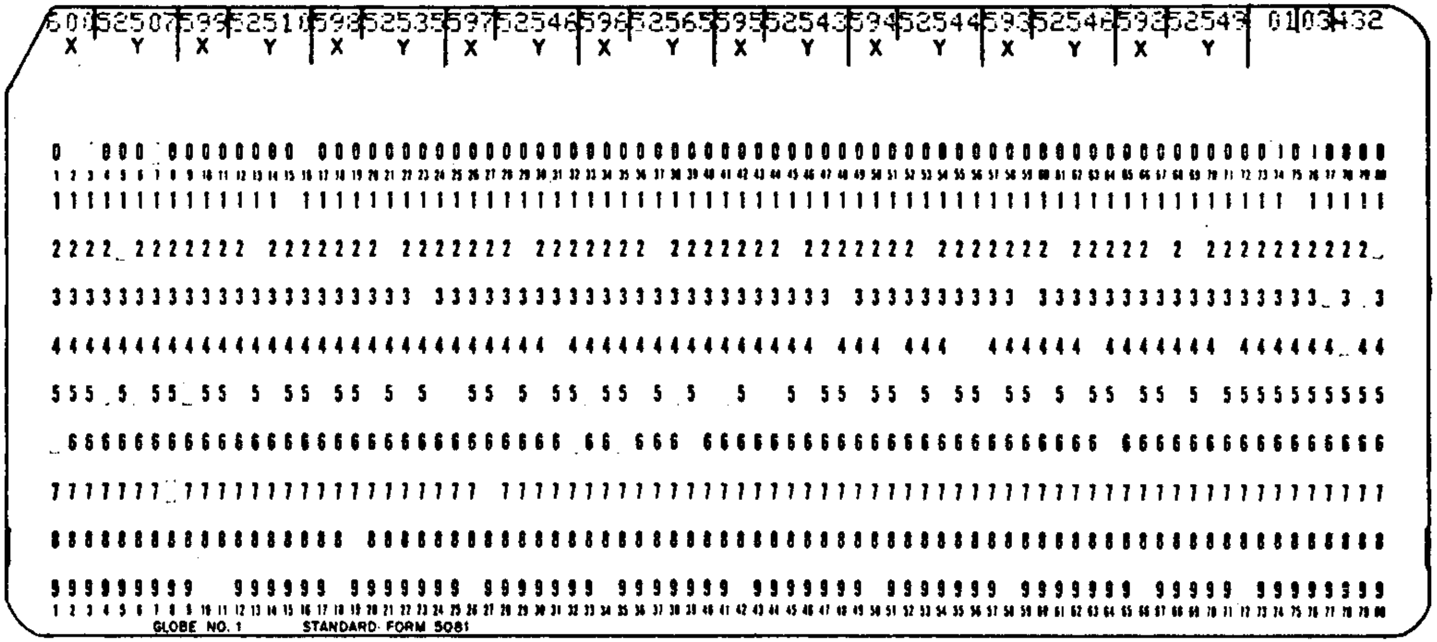


Fig. 3. Topographic data plot for points taken every 10 ft translation in x.



Column(s)	Item
1-3	X (0⟶ 99.9, decimal not recorded)
4-8	Y (0⟶ 99.999, decimal not recorded)
•	
•	(X and Y repeated through column 72)
•	
73	Not used
74-75	Map Number (1,2,...,11)
76-77	Contour Line Identification (1,2,...,99)
78-80	Contour Level (Elevation measured from mean sea level, modulo 5000)

Fig. 4. Format of the data cards used for recording the coordinatograph output when digitizing the contour maps.

After each map is digitized, the recorded data is plotted on microfilm (Fig. 3) to determine whether or not data points were missed during digitizing. This is accomplished by scaling the x and y coordinates of the data to the coordinate system of the microfilm plotter. One must be able to clearly distinguish individual points on the printed microfilm. Hence, a maximum of 146 points may be plotted in each of the horizontal and vertical directions, although there are actually 1024 addressable raster units on each axis on the microfilm CRT. This practical resolving limitation prevents the existence of a one-to-one relationship between the number of original maps and the number of frames of microfilm plotted. One frame of microfilm, therefore, represents an area of landscape 1460 ft × 1460 ft. As an aid in assembling the prints of the several frames of plotted microfilm, data at each edge of the frame is duplicated on the frame containing the adjacent data. This provides about a one-half inch overlap on the scale map presented in Fig. 3. After prints of the microfilm have been made and assembled, the plot (Fig. 3) of the data points is visually compared with the original contour map. The data points should appear at regular intervals in the x direction, and one can readily identify points where data have been missed during the recording phase. The main emphasis for this verification is that large gaps in the data points, or even an entire contour line not present in the data, can readily be determined. Even the absence of a single point can be detected. One can also readily distinguish the presence of duplicate data as these points appear much bolder on the microfilm plot than the other data points. Duplicate data indicates that a portion of the contour line was digitized more than once. When erroneous gaps are verified in the data points, the map is returned to the coordinatograph to remeasure those portions of the map which were missed during the original digitizing.

Next, data points are edited to remove those points which are duplicated. This function, which is performed by the computer, eliminates those points from a given contour line which have the same x coordinate and are within some small, selected epsilon (ε) of another point along the same line. This also eliminates those points which may have been duplicated as a result of deliberately digitizing the overlap areas of adjacent contour maps.

Once the topographic data has been collected, verified, and completeness established, it is ready for processing. This involves sorting the data for each x value into a record to collect the data points together for each of the terrain profiles. Next, the record at each x value is sorted by y values to form a terrain profile.

The terrain profiles (Fig. 5) are analyzed quantitatively as two-dimensional curves. A second degree polynomial is fit to successive triples of data points beginning with the first three points; then the second, third, and fourth points; the third, fourth, and fifth points, etc. Each time the first of the previous three points is dropped, a new point is added. Each polynomial is evaluated for the elevation at each 10 ft inverval in y between the first and second points which corresponds to a grid cell intersection (Fig. 6). This process, when performed on all positions on all terrain profiles, gives the mean elevation of every 10 ft × 10 ft cell in the landscape modelled and becomes the basic data for computing the other topographic parameters.

Sharpnack and Akin (1969) have developed a technique for computing slope and aspect for a grid cell from a matrix of surrounding elevation map cells using a specialized formula for regression coefficients. For the cell layout

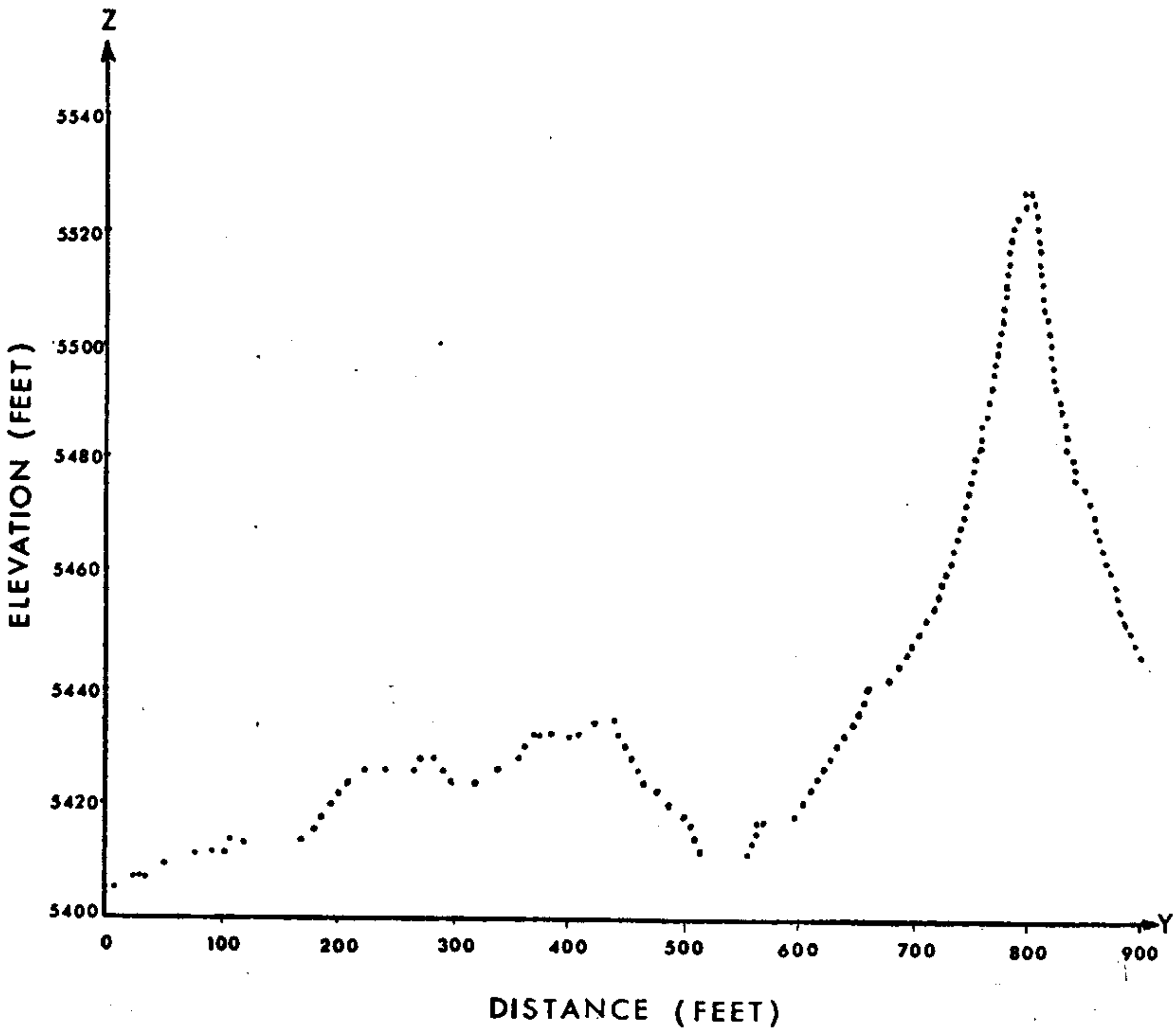


Fig. 5. Topographic profile formed in sorting digital data points.

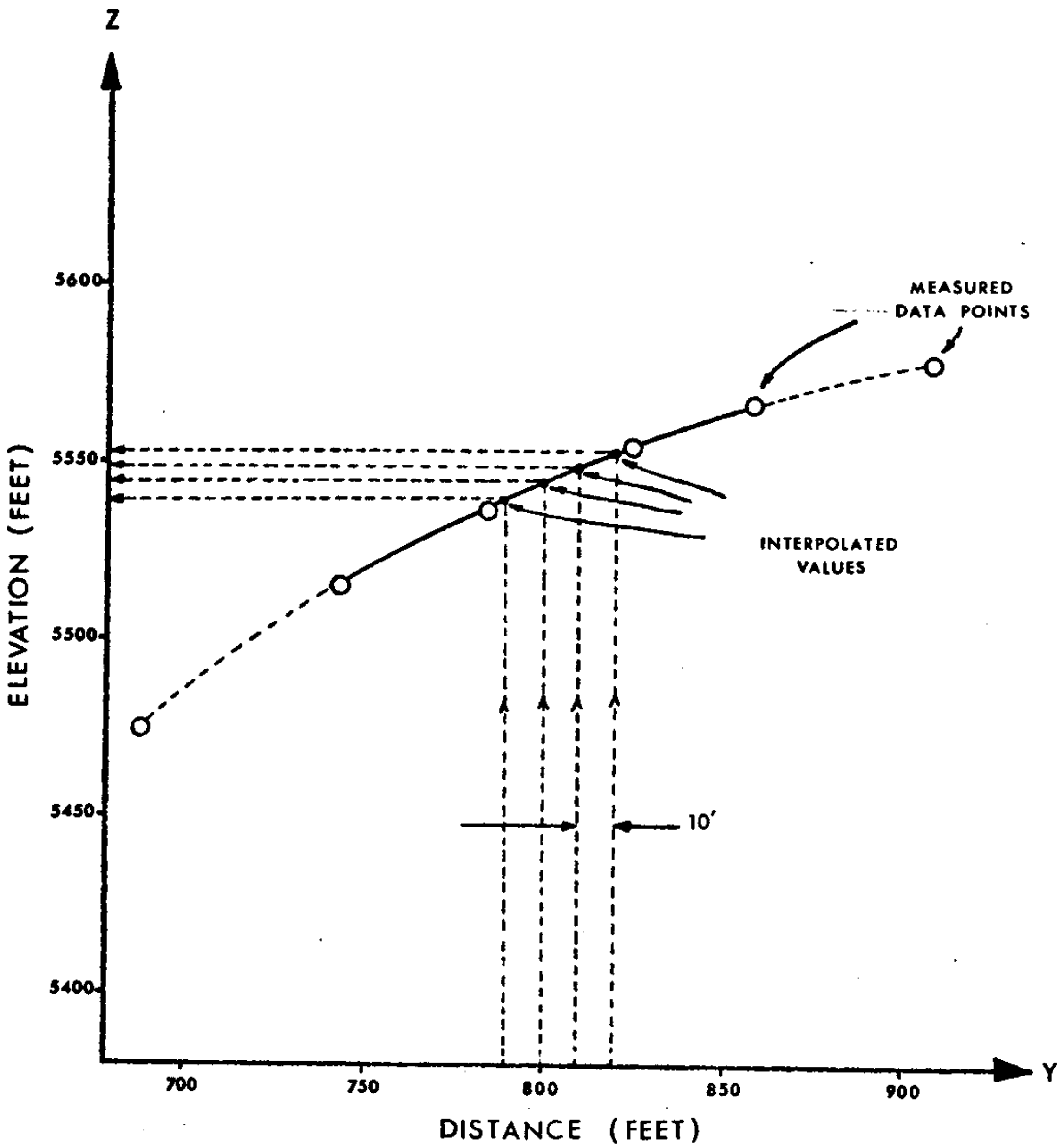


Fig. 6. Interpolation for intermediate cell elevation. Interpolation performed every 10 ft along center section of curve fit to profile.

	$-a$	0	a
$-a$	Z_1	Z_2	Z_3
0	Z_4	Z_5	Z_6
$-a$	Z_7	Z_8	Z_9

↑ N

they have shown that the formula for the regression coefficients is

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} (Z_1 + Z_2 + Z_3 = Z_7 - Z_8 - Z_9)/6a \\ (Z_3 + Z_6 + Z_9 - Z_1 - Z_4 - Z_7)/6a \end{bmatrix}$$

and mean slope can be computed by

$$\text{Percent mean slope} = (B_1^2 + B_2^2)^{\frac{1}{2}}$$

Also, mean aspect from $x = \arc\tan(B_2/B_1)$, $0° < \text{aspect} \leq 90°$. Aspect is converted to azimuth by determining its quadrant from the signs of the slope coefficients.

Soil Type

The next feature to be overlaid into the grid cells is soil type. A map of soil type of the Pawnee Site is available (Fig. 7) and constitutes the basis for overlaying this feature into each cell. The boundaries of each soil type can be digitized using the coordinatograph and the data scaled and converted to the same coordinate system as the topographic data. Each soil type can be assigned a unique classification code and, based on the digitized boundaries, can be projected to the appropriate cells.

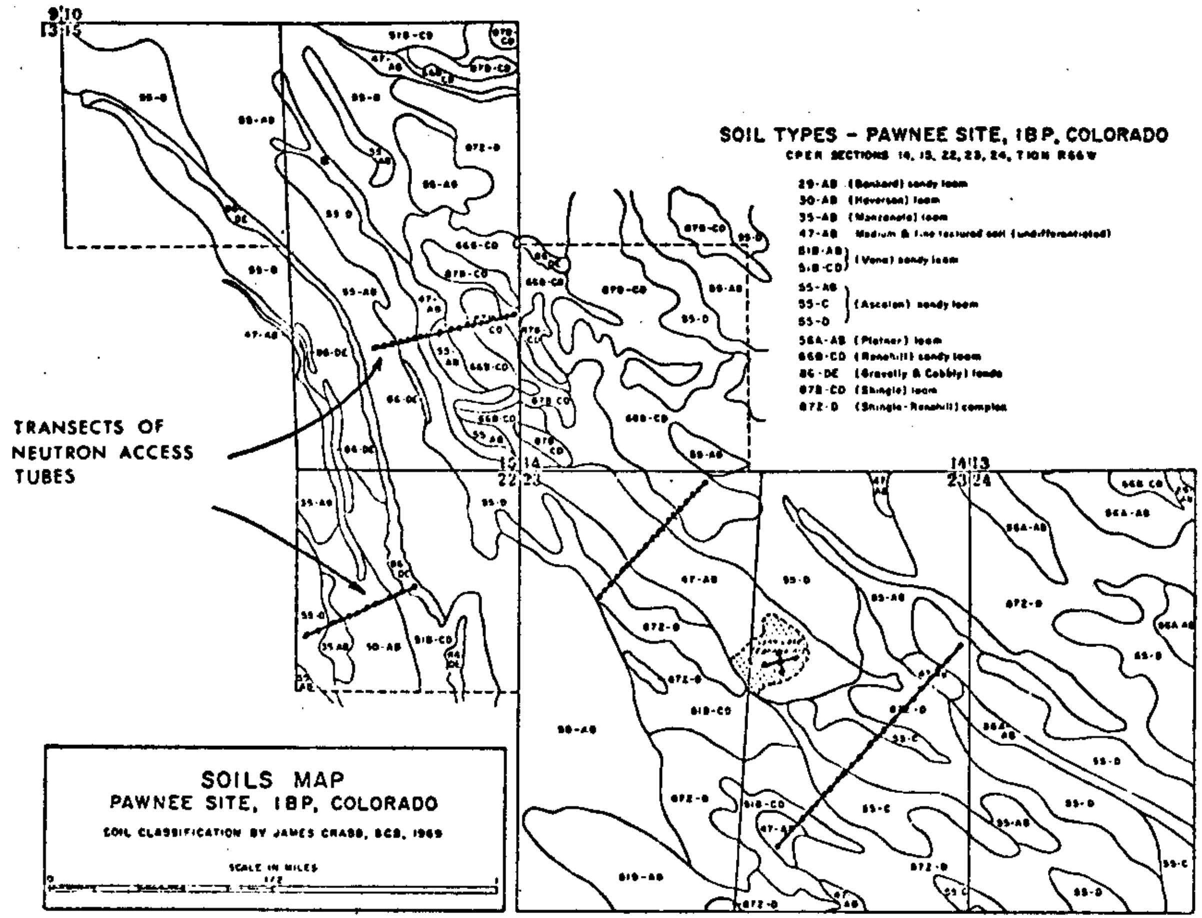


Fig. 7. Soil type overlay. Also shows the location of the neutron access tubes used to measure soil moisture content at selected depths. (After Jameson 1969).

Vegetative Cover Distribution

A vegetation map (Fig. 8) of the basic grassland communities has been completed from 1/12,000 ektachrome infrared and color aerial photos of the site. The boundaries of the various vegetation types can be digitized from this map using the coordinatograph and scaled and converted to the same coordinate system as the topographic data. A classification code associated with each vegetation type can be projected to the appropriate data cells according to the digitized boundaries of the vegetation map units.

Soil Moisture Content

Hydrologic characteristics of the Central Basin are being studied via data obtained from the approximately 150 neutron access tubes located in the basin and the microwatersheds (Fig. 7). From this data regression equations for each soil type can be derived which correlate soil moisture with slope, aspect, elevation, vegetative type or other combinations of characteristics in the model. At any specified time, a reading may be taken of the soil moisture content at a given depth at the locations of each of the access tubes which reside within the boundary of a given soil type. This should provide the dynamic data necessary to predict the soil moisture at the locations of the data cells within that soil type (Fig. 9). The hydrologic studies being conducted may establish that additional related parameters are necessary for determining soil moisture. Such land-use practices as grazing intensity, treatments, etc. may prove necessary for establishing a sufficient soil moisture relationship and can be overlaid into the cells.

Fig. 8. Vegetation type overlay.

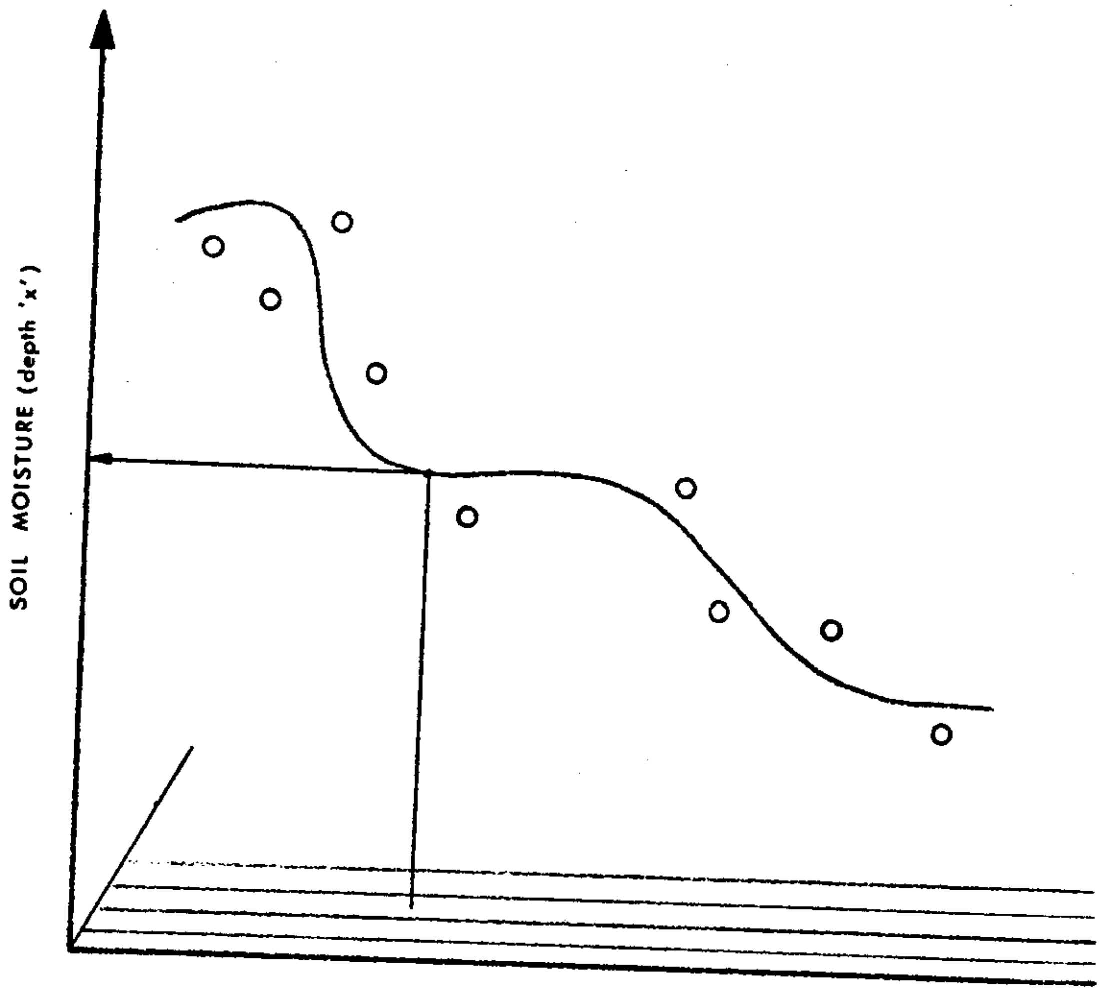


Fig. 9. Possible approach to determining cell soil moisture. Stylized representation of a regression equation determined for a particular soil type and depth on a given day.

How Can the Procedure be Manipulated and Displayed?

The data for each cell can reside on tape or disk located in the CSU computer center and be accessible via the center's CDC 6400 computer system. The data specifying any one or all of the 400,000 cells may be readily accessed by specifying its x and y position (Fig. 10). Any subset of this data can be subjected to statistical analysis. Studies such as the cross correlation of the cell parameters can be made (e.g., how does vegetation type correlate with soil type, slope, aspect, etc.). When remote sensing data is available, correlation studies can be made of relationships between the data in the cells and spectral characteristics obtained from a multiband scanner. The results from these studies may be analyzed using output via microfilm, color display, or high speed printer. This output may be in the form of tables, histograms, gray-scale maps, isoline maps, etc.

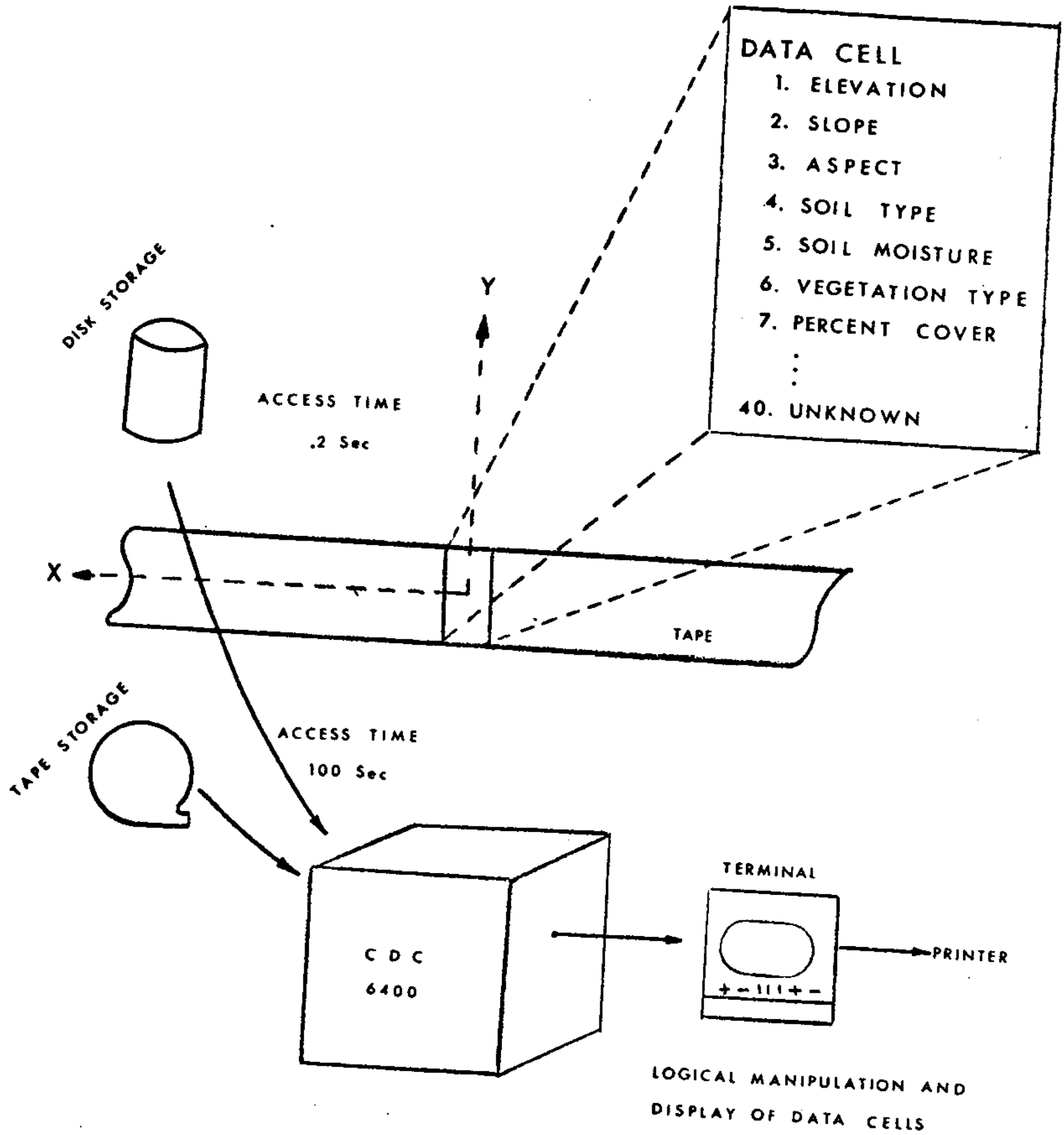


Fig. 10. Manipulating the model of the landscape.

LITERATURE CITED

Jameson, D. A. 1969. General description of the Pawnee Site. U.S. IBP Grassland Biome Tech. Rep. No. 1. Colorado State University, Fort Collins. 32 p.

Sharpnack, D. A. and Akin, G. 1969. An algorithm for computing slope and aspect from elevation. Photogrammetric Eng. 35(3):247-248.