The Use of Variability Diagrams to Improve the Interpretation of Digital Soil Maps in a GIS

Abstract
A critical layer in geographic information systems (GIS), particularly when utilized in land management decisions, is soil survey information. The spatial component of this information is generally input from digitized survey maps, usually in the form of polygons representing different soil map units. It is a fact that the homogeneity of soils within soil map units varies. Conveying this variability to users is essential to ensure proper use of soil survey information.

Using the transect method, four forested soil map units were examined to assess their homogeneity with respect to the variability of field determined soil taxonomy and physiography and the interpretive variability of selected soil properties for forest land management. The degree of interpretive variability was determined using Shannon's measure of entropy. Variability diagrams and interpretive maps were generated within a GIS. These diagrams and maps, coupled with the digitized soil maps, inform users of the degree of soil map unit variability and the variability of limiting soil properties.

Introduction
Mapping the distribution of soils is an important activity of the United States Department of Agriculture Soil Conservation Service (USDA SCS), Forest Service, and other agencies involved in the National Cooperative Soil Survey Program. A long standing goal of the soil survey has been to provide information to users about soils to facilitate wise land use and management. Information is required on the behavior, limitation, and potential of soils for planning and implementing most land-use activities.

Soil survey maps convey this information by showing soil map unit delineations on an aerial photo base. These delineations are drawn and map units are designed to show like kinds of soils within a specific land area. Soil map units identify soils and their properties. The soil's taxonomic class and site attributes such as slope and erosion potential are also recorded for soils within the soil map units. Soil map units can be either uniform, wherein the major proportion of the soil properties and site attributes of a land area are homogeneous and thus perform uniformly, or complex, wherein the land area has variable soil and site patterns which are generalized. The lack of detailed mapping in these complex areas may not be adequate for land-use and management decisions. Over the past several decades, a number of studies have been undertaken to aid in the design of soil maps as well as to define variability within soil map units.

The majority of soil map unit variability studies have concentrated on the taxonomic variability of soils within soil map units and have been spurred by the adoption of soil taxonomy (Soil Survey Staff, 1975). Little work, however, has centered on the interpretive variability of soil attributes of a given land area or soil properties within soil map units. The need for this type of information has been documented by Miller (1978), but obtaining the information and then conveying to users the variability of the information within soil map units and the possible land-use and management affects has been wanting in many cases.

Studies concerning the variability of soil properties and taxa are extensive. Beckett and Webster (1971) summarized most of the pre-1970 variability investigations and found that as much as half of the homogeneity within a map unit was present within a few square metres and map units were approximately 50 percent pure by taxa. Variability was qualified by stressing that the impurities are often similar and would not require different management. Wilding and Drees (1983) summarized the magnitude of variability observed in various investigations from 1970 to 1980 and found that variability increased as map scales became smaller. They also noted that most soil properties in the map units had a coefficient of variation of 25 to 40 percent. Therefore, an unrealistic number of soil samples are needed to estimate most soil properties at the 95 percent confidence level with 10 percent allowable error.

As a consequence of these studies, it has been suggested that, in addition to taxonomic purity, attention should be given to the variability of interpretations and their influence on use and management (Bouma, 1985; deGrujter, 1982; Miller 1978). In comparing the observed suitabilities of soils for management within map units to the suitabilities of the named soil(s) of the map unit, it was noted that observed soil suitabilities within mapping units can be quite variable, and can differ significantly from the suitabilities implied by the named soil for the mapping unit. The end product of a soil map is not only the map, but also the interpretation. This has been the basis and purpose of the soil survey in the United States. Acceptance of the soil survey as a tool for making land-use and management decisions is contingent upon users having interpretations that are consistent and accurate within the limits of the product. This requires that users have a knowledge of the confidence limits of the se-
lected properties and know that the selected properties vary within the delineated landscape on a soil map (Miller, 1978).

The increasing use of geographic information systems (GIS) further compounds the variability problem, and the use of soil maps as a thematic layer in a GIS is well documented (Best and Westin, 1984; Burrough, 1986; Hendrix and Price, 1986; Hidelbaugh, 1982; Niemann et al., 1987; Smith and Sousa, 1986; Walsh et al., 1987; Vold et al., 1985). Utilizing a GIS, users can retrieve site-, soil-, or use-specific information but seldom the entire soil survey report. The supporting information in the soil survey report provides an indication of the inherent variability in soils to informed readers. However, thematic soil maps can convey an unrealistic degree of homogeneity which tends to obscure natural variation behind smoothly drawn lines and patches of color (Burrough, 1986). Thus, supporting information must be provided to users concerning survey techniques, including the sampling and statistical methods employed, qualitative and quantitative descriptions of soil variability, and the implications for intended use.

Several studies utilizing GIS have been completed which illustrate the impact that the variability within soil map units can have on management decisions. Williams (1985) investigated the feasibility of implementing the USDA SCS's Land Evaluation and Site Assessment with a GIS. Soil information used in the analysis included crop productivity rating and suitability for on-site waste disposal. He found the use of a GIS feasible for applications that were not site specific. Niemann et al. (1987) used soil maps and the Universal Soil Loss Equation model identifying soil erosion potentials for use in monitoring landowner compliance with conservation measures in Dane County, Wisconsin. Vold et al. (1985) conducted a similar investigation in British Columbia and noted the need for field checking individual sites.

Interesting enough is the fact that the GIS technology, which could further the misconception of soil map homogeneity, can easily solve the problem by providing users with variability information. Chrisman (1984) suggested the use of reliability overlays for GIS thematic layers which indicate to users locations where the data can be applied with the most confidence.

In order to present meaningful soil map unit variability information to the user, two research objectives were developed. The first objective was to determine interpretive variability within soil map units for four soil taxa and to quantitatively measure the soil map unit homogeneity. The second objective was to spatially define the variability and incorporate the information into a GIS using the formats of a variability diagram and multi-level interpretive maps.

Methods

Study Site

The study site was located in Houghton County in Michigan's Upper Peninsula (Figure 1). The county is situated on the southern end of the Canadian Shield. The soils are predominantly underlain by Precambrian igneous and metamorphic rocks. Relief ranges from 183 to 457 metres. A detailed description of the bedrock geology and glacial history of the county can be found in D'Avello (1988).

Because 86 percent of the county is forested, and timber and pulp production is a major land use, soil woodland interpretations developed by the USDA SCS were incorporated in this research as a test of soil map unit homogeneity (Supplement 2, Soil Survey Staff, 1980-86). An example of the interpretations is provided in Table 1. Soil woodland interpretations evaluated included

- Windthrow hazard,
lection of transect sites. Fifty-nine transects were allocated proportionately to the four map units. The number allocated was based on the estimated acreage per map unit divided by the total predicted acreage of the four map units in the county.

Placement of individual transects on individual soil maps was based on map unit size and access, with delineations of representative size, shape, and accessibility receiving priority. Transects were located perpendicular to topographic features, such as ridges and valleys, and generally across the widest areas of the map unit. More than one transect was located in large soil delineations.

A constant observation interval of 65 metres, rather than a constant number of observations per transect, was used. This was done to reduce bias of varying transect lengths when using a constant number of observations per transect.

Field observations recorded on SCS Soil Transect Data Sheets at each interval included transect location and soil horizon information such as depth, color, texture, mottling, pH, depth to root limiting layers, and slope. Soils at each transect point were described and classified according to National Cooperative Soil Survey Standards at the series level.

**Statistical Analysis**
For the statistical analysis, individual transects were treated as samples while observations within transects were treated as subsamples. Characteristics of selected soil properties defined placement of a soil into a given interpretive rating. The proportion of interpretive ratings observed for each selected property were calculated for each transect. Descriptive statistics were generated for ratings and taxonomic class from transect totals for each soil mapping unit.

Interpretive variabilities of soil map units were determined using Shannon's measure of entropy (Shannon and Weaver, 1949):

\[ H = - \sum n_i \ln n_i \]

where \( H \) = entropy (variability for this study) and \( n_i \) = proportion of an interpretive rating. Shannon's concept applies not to individual entities, but to the situation as a whole, making it an ideal measure of map unit variability.

Soviet soil scientists Fridland (1976) and Razumov (1986) discuss the use of Shannon's measure of entropy for soil cover and soil horizon complexity studies, respectively.

The maximum entropy value which conveys the greatest variability for a three class system is expressed as:

\[ H = - [(1/3) \ln(1/3) + 1/3 \ln(1/3) + (1/3) \ln(1/3)] = 1.0986123 \]

and the minimum value is expressed as

\[ H = - [(0) \ln(0) + (1.0) + (0) \ln(0)] = 0 \]

Therefore, the amount of information conveyed is inversely related to the value of \( H \), while variability is directly related to the value of \( H \).

Values were standardized to range between 0 and 1 by dividing all values by the maximum \( H \). Three classes were chosen to reflect the degree of variability and information content of the data. The classes and corresponding descriptions were:

- \( H < 0.33 \): Nonvariable
- \( 0.34 \leq H < 0.67 \): Moderately variable
- \( H \geq 0.68 \): Highly variable

**Class Descriptions:**
- Nonvariable: soils occurring in the map unit require similar management.
- Moderately variable: soils occur in the map unit that require either more, less, or a combination of restrictive use and management. Inclusions are common, or the unit is a complex or association, which recognizes two or three soils within a delineation. Making predictions for particular areas within the soil map unit is difficult due to variability.
- Highly variable: both soils that limit and do not limit use and management commonly occur. The map unit is probably a complex or an association, which recognizes two or three soils within a delineation. Making predictions on use for a particular area is difficult due to the degree of variability.

Variability within the soil map units of selected soil properties which influence woodland use and management were also evaluated using Shannon's measure of entropy.

**Table 1. Guide to Equipment Limitations for Mechanical Site Preparation and Planting (Adapted from Supplement M2, Soil Survey Staff, 1986).**

<table>
<thead>
<tr>
<th>Property</th>
<th>Limits</th>
<th>Restrictive Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of water table above 15 inches (months)</td>
<td>Slight</td>
<td>Moderate</td>
</tr>
<tr>
<td>Flooding</td>
<td>None, Rare, Occasional</td>
<td>1-3 Frequent, Long or Less</td>
</tr>
<tr>
<td>Rock Outcrop (%)</td>
<td>&lt;10</td>
<td>10-25</td>
</tr>
<tr>
<td>Depth to hard bedrock (inches)</td>
<td>&lt;0.1</td>
<td>0-3</td>
</tr>
<tr>
<td>Boulders Class</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Stones Class</td>
<td>&lt;3</td>
<td>3-15</td>
</tr>
<tr>
<td>Stones Class</td>
<td>0.1,2</td>
<td>3</td>
</tr>
<tr>
<td>Fraction 3-10 inches in diameter</td>
<td>&lt;25</td>
<td>25-50</td>
</tr>
<tr>
<td>USDA Texture Organic Material</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Clayey Textures</td>
<td>C,SIG,SC</td>
<td>COS,FS,S,</td>
</tr>
<tr>
<td>Sandy Textures</td>
<td>VF S</td>
<td>-</td>
</tr>
<tr>
<td>Slope</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Potential Frost Action</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
These properties are listed in the Soil-Woodland interpretive guidelines (Supplement M2, Soil Survey Staff, 1980-1986), and included: depth to water table; percent of the surface covered with boulders, stones, or cobbles; depth to bedrock; texture of the upper 10 inches; potential frost action; average percentage of cobbles by weight in the 0 to 40 inch depth; AASHTO group for the upper 10 inches; and AASHTO group for the thickest layer between depths of 0 to 40 inches. See Table 1 for examples.

In keeping with the interpretive guidelines shown in Table 1, three levels of detail were used for presenting the interpretive information and are portrayed as a Level 1 Interpretive Map. However, the Level 1 Interpretive Map does not convey enough information to the user where the soil map units are multi-taxa. Consequently, two additional levels of interpretive maps were developed: Level 2 Interpretive Maps (distinct ratings are separated) and Level 3 Interpretive Maps (distinct ratings and distinct limiting soil properties are separated as classes).

Level 1 Interpretive Maps portray a worst case scenario when map units with multiple ratings are present. A rating of slight and moderate for a map unit would be rated moderate.

Level 2 Interpretive Maps maintain distinct classes, which are important for complexes and associations. A rating of slight and moderate would be separate and distinct from a moderate and slight rating.

Level 3 Interpretive Maps are the most detailed with limiting soil properties added as a variable to Level 2 Interpretive Maps. Level 3 categorizes soils of similar ratings and limiting properties.

To illustrate the generation and application of the variability diagrams, the soil map for the study site was manually digitized and entered into a raster GIS with a pixel size of 10 metres square. This approximated the spatial resolution of the 1:20,000-scale soil maps. The digitized map was geographically rectified to the UTM coordinate system. The transect information for the study site was entered into a relational database and, utilizing a program written by Thomasma (1988), the descriptors for the interpretive maps were generated.

Results and Discussion

Based on the statistical analysis of the transects, a soil map unit variability diagram was developed (Figure 2). The map indicates to users that the “purity” of the map units within the study site are nonvariable or moderately variable. The descriptor of moderately variable is portrayed in two shades of green to inform users of the range of ratings found in the field samples. The 10B map units have an H value (entropy value) of 0.41 with field conditions ranging from and including slight, moderate, and severe. The 30B map units have an H value of 0.62 with field samples evenly split between moderate and severe ratings.

A limitation to this general type of variability diagram is that, while the user knows to expect variations within the map unit and to what degree, it does not present the location of the variability, details on the level of variability, or the restrictive feature present. Hence, the development of the Level 1, 2, and 3 Interpretive Maps. These interpretive maps were developed for each of the soil woodland interpretations outlined previously. To illustrate the increase in complexity with decreasing generalization, consider the interpretive maps for equipment limitations for mechanical site preparation and planting (Figures 3, 4, and 5). These interpretive levels are proposed to provide users with the option and choice of determining the desired level of detail required to meet their needs. The interpretive levels are also useful in GIS analysis operations, especially with regard to soil map units with multiple ratings such as the Munising-Skanee complex (30B) and the Trimountain-Paavola complex (89B).

The Level 1 Interpretive Map (Figure 3) informs the user that there are moderate and severe limitations present that will restrict equipment use for site preparation and planting. When overlain with the digitized soil map, the user also knows what soil map units are present and their location. However, no information is provided as to the range of variability within each map unit or restrictive features responsible for limitations.

The Level 2 Interpretive Map (Figure 4) provides the user with information regarding the range of variability within the soil map by map unit location. The Level 3 Interpretive Map (Figure 5) contains the same information found on the Level 2 Map and also informs the user of the restrictive feature causing the limitation.

The impact of the level of detail presented in the variability diagrams can be assessed by calculating acreage for the information classes. Level 1 Interpretive Maps are based on a
worst case basis and therefore group as moderate about 8 acres of a map unit rated slight and moderate. It also groups 847 acres of a map unit rated moderate and severe as severe. This acreage comprises 74 percent of the area rated as severe.

Use of the Level 2 Interpretive Map eliminates this generalization. However, it does not provide information about the restrictive feature(s) contributing to the variability. For example, restrictive features found in the moderate rating are comprised of 56 acres of a map unit due to sandy surface textures; 16 acres due to slope and sandy surface textures; 1,154 acres due to wetness and potential frost action; and 60 acres due to slope, wetness, and potential frost action. These areas make up 4, 1, 90, and 5 percent, respectively, of the area rated as moderate. However, each of these restrictions poses a different equipment limitation when considering mechanical site preparation and planting.

The Level 3 Interpretive Map provides the highest level of detail. Both the range in variability and the restrictive features responsible for the limitation are mapped. This interpretive map, when overlaid with the digitized soil map, not only provides the user with soil map unit names but also with restrictive features present and their location. All of which is critical information for many land-use planning activities.

Intended use will dictate the level of generalization required. A GIS allows the flexibility to provide to users the varied degrees of detail necessary for a given use.

Conclusions
Soil map units in this study contained inclusions of soils that rated differently from the dominant member(s). Soils with ratings better and/or worse occurred and when summed sometimes exceeded the standards for dissimilar inclusions in the soil map unit. This is in no way a reflection of the quality of the soil map or the ability of the soil scientist in making soil maps. Rather, it illustrates the limitations of the "traditional" analog map and the supporting interpretive material which accompanies it.

The variability within the soil map units cannot be reduced as it is inherent in the natural soil bodies on the landscape. The use of variability diagrams and multi-level interpretive maps which portray the variation of soil map units and provide the user with detailed information on limiting soil properties which affect land use can easily be implemented with GIS technology. This permits critical interpretive information to be incorporated into the planning process from the beginning of a project. While our study researched variation only with four map units and considered soil woodland interpretations, the methodology is applicable to all land-use activities.

The use of Shannon's measure of entropy is recommended for production of the variability diagrams because it considers the map unit as a single entity, which it is. It is also important to note that the variation can be calculated from SCS data as it is currently being collected in survey areas employing the transect method or other statistical sampling strategy. No modifications in the soil mapping procedure are required for these areas.

It is noted by Edmonds et al. (1986) that soil scientists could better serve users through increased freedom in conveying map unit descriptions, the ranges in soil properties, distribution and composition within map units, and possible implications for use and management. The research presented here illustrates just that.

Acknowledgments
The authors thank the Environmental Division of the Michigan Department of Agriculture for financial assistance in support of this study and C. Schwemmer, K. Wikgren, and T. Bauer for assistance in soil classification and soil map unit design.

References
D'Avello, T. P., 1988. Interpretive Variability of Soil Map Units and
Application for Use with a Geographic Information System.
deGrujter, J. J., 1985. Transect sampling for reliable information on
mapping units. Soil Spatial Variability (D. R. Nielsen and J.
Edmonds, W. J., and M. Lentner, 1990. Statistical evaluation of the
taxonomic composition of three soil map units in Virginia. Soil
Fridland, V. M. 1972. Pattern of Soil Cover. (Transl. from Russian by
Hendrix, W. G., and J. E. Price, 1986. Application of geographic in-
formation systems for assessment of site index and forest constraints.
368-377.
Hidelbaugh, A. R., 1982. The Soil Conservation Service's interest in
graphic information systems. Comput. Environ. Urban Sys-
tems. 7:295-300.
Johnson, W. M., 1961. Transect Methods for Determination of Com-
poition of Soil Mapping Units. Soil Survey Tech. Notes. USDA
SCS.
Miller, F. P., 1978. Soil survey under pressure: the Maryland experi-
Niemann, B. J., J. G. Sullivan, S. J. Ventura, and N. R. Chrisman,
1987. Results of the Dane county land records project. Photo-
grammetric Engineering & Remote Sensing. 53:1371-1378.
Razumov, V. V., 1986. Comparative analysis of soil variation at dif-
ferent categorical levels of subalpine geosystems. Soviet Soil Sci.
Schwenner, C., 1992. Soil Survey Report for Houghton County,
Michigan. USDA SCS U.S. Gov't Printing Office, Washington,
D.C.
Shannon, C. E., and W. Weaver, 1949. The Mathematical Theory of
Smith, G. S., and D. A. Sousa, 1986. Direct application of the SOI-5
soil interpretations records to a geographic information system.
366-393.
Classification for Making and Interpreting Soil Surveys. USDA
Govt. Printing Office, Wash. D.C.
Thomasma, S., 1988. A program to link ERDAS with a relational
data base system. The Monitor. ERDAS, Inc. Atlanta, Georgia. p.
1.
Vold, T., M. W. Sondheim, and N. K. Nagpal, 1985. Computer as-
65:411-418.
Walsh, S. J., D. R. Lightfoot, and D. R. Butler, 1987. Recognition and
assessment of error in geographic information systems. Photo-
Williams, T. H. L., 1985. Implementing LESA on a geographic in-
formation system - a case study. Photogrammetric Engineering &
(Received 9 August 1991; Revised and Accepted 14 July 1992)

LIST OF "LOST" CERTIFIED PHOTOGRAMMETRISTS

We no longer have valid addresses for the following Certified Photogrammetrists. If you know the
whereabouts of any of the persons on this list, please contact ASPRS headquarters so we can update their
records and keep them informed of all the changes in the Certification Program. Thank you.

Lane Schultz James Steckling Keith Syrett William Thomasset Conrad Toledo Robert Tracy Lawrence Watson Tad Wojenka