AN ABSTRACT OF THE THESIS OF

Joel K. Poore for the degree of Master of Science in

Soil Science presented on February 21, 1986.

Title: Quantitative Shape Analysis of Soil Map Delineations in Benton County, Oregon

Redacted for Privacy

Abstract approved:

J. H. Huddleston

Soils are 3-dimensional bodies that make up natural landscapes. In addition to the morphological properties used to characterize soils, soil bodies also have the properties of size and shape. Soil maps are made in an effort to provide information on the spatial distribution of different kinds of soils. Soil mappers draw to scale, as accurately as possible, the sizes and shapes of the different kinds of soil bodies they observe in the landscape. Beyond that, however, very little quantitative information relative to size and shape is provided to soil map users.

Quantitative shape characterization presents several opportunities to learn about soil genesis and soil interpretations for land use. Intriguing questions include "Why does a soil body have the particular shape it has?", "Does each map unit possess an intrinsic shape or range of shapes?", "Do existing map unit interpretations apply equally to delineations of different size and shape?", "How can shape data for individual delineations be

aggregated into an overall description of soil patterns in different geographic areas?", "What effects do soil patterns have on land use?".

None of these questions can be answered without first having an appropriate technique for characterizing the shapes of individual delineations. The objective of this research, therefore, was to examine several possible shape indexes and isolate those few which had the greatest utility for characterizing shape. These few were then used to examine shape distributions within a few selected map units and compare shapes between map units.

Data were collected by digitizing 452 delineations sampled from 13 different kinds of soil bodies identified in the soil survey of Benton County. For each delineation, 43 potential indexes were calculated. These included primary measurements, such as area and perimeter, and figure attribute ratios such as Horton's form ratio, Miller's circularity index, Schumm's elongation ratio, and Fridland's coefficient of dissection. A convex hull was circumscribed around each delineation, and the same primary measurements and attribute ratios were calculated for the convex hull. Additional indexes were calculated by comparing values determined for a delineation and corresponding values for it's convex hull.

One additional technique used was to fit each polygon and convex hull with a 22-sided vertex lag polygon. Calculation of distances between vertices of this polygon leads to the derivation of a vertex lag index of shape. Variations on the vertex lag theme provided several additional indexes.

Correlation analysis showed that the set of 43 indexes was highly intercorrelated. In order to reduce this set to a smaller set of minimally correlated indexes, the entire data set was subjected to a factor analysis. The result was a set of three dominant factors, which together accounted for 86.1% of the total variance in the data set. Each factor was interpreted by considering the nature of the shape indexes that loaded heavily on it, and a single index was selected to represent each factor on the basis of maximum interpretability.

The first factor was interpreted as a measure of the complexity, or irregularity, of a delineation. The vertex lag index for the delineation was selected as the best single index to represent this attribute of shape. The second factor included all of the primary measurements. Though not a measure of shape *per se*, primary measurement data were viewed as significant elements in the spatial description of soil map delineations. Polygon area was taken as the best index to represent the effects of primary measurements. The third factor was interpreted as a measure of form. In this case, Schumm's elongation ratio, as measured on the convex hull, was found to be the most interpretable index of form.

These three attributes, size, form, and complexity, provided the best quantitative description of shape. The indexes that represent them were found to be minimally correlated and maximally interpretable.

Each of the 13 kinds of soil bodies sampled was characterized in terms of the three aspects of shape using descriptive statistics and frequency histograms. Comparisons between samples were evaluated using the Mann-Whitney U test. The data suggested that delineations belonging to a single soil mapping unit do have distinctive distributions of size, form, and complexity. Shape differences between mapping units were most evident when comparing soils on different landforms, parent materials, and slope gradients.

Quantitative Shape Analysis of Soil Map Delineations in Benton County, Oregon

by

Joel K. Poore

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed Febuary 21, 1986

Commencement June 1986

APPROVED:

Redacted for Privacy

Professor of Soil Science in charge of major

Redacted for Privacy

Head of department of Soil Science

Redacted for Privacy

Dean of Graduate School

Date thesis is presented February 21, 1986

Typed by Joel Poore for Joel K. Poore

ACKNOWLEDGEMENT

I would like to dedicate this theses to my two children,

Jesse and Tyson

Special thanks to Herb Huddleston for providing direction, organization, and encouragement during this study. Thanks to Jerry Simonson and Jon Kimmerling for teaching me, and for their suggestions regarding the theses. I would also like to thank Randy Brady for his support, and Stan Carpenter for the much needed help with data analysis. Finally, I offer my warmest appreciation to my wife and parents for love that is constantly shown to me.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

QUANTITATIVE SHAPE ANALYSIS OF SOIL MAP DELINEATIONS IN BENTON COUNTY OREGON

INTRODUCTION

Individual areas demarcated on a soil map are called soil map delineations. The shape of each delineation of a consociation approximates the actual shape of a soil body (polypedon) belonging to a single class of soil. Exact representation of pure soil bodies on a map is not feasible. Soil maps would not be legible if exact boundaries were drawn between each polypedon in a survey area. Soil map delineations, therefore, represent spatial occurrences of map units rather than natural soil bodies.

A consociation soil map unit is defined by a dominant class (or phase of a class) and an acceptable percentage of inclusions of different classes of soil. Recognizable inclusions and small boundary irregularities are often omitted or generalized on soil maps due to map scale.

One soil map unit in a soil survey can be used to define a large number of individual delineations. Delineations defined by one map unit are spatially separated by delineations of different map units or non-soil areas.

Soil maps made using this concept of map units, as are those included in recent soil surveys in the U. S., provide an immense amount of spatial soils information. Each mapped soil body, or delineation, has the properties of size and shape, which can be and should be studied quantitatively (Kellogg, 1963).

Subjective descriptions of soil map delineations are included in a recent soil survey of Wasco county, Oregon (Green, 1978). Descriptions include the Wamic loam 1-5% slopes found in "broad, smooth, convex areas" and the Frailey loam 3-30% slopes in "broad, irregularly shaped areas". A universally accepted quantitative shape index could replace such descriptions in order to improve communication and assist scientific study.

The population of delineations may have a characteristic shape or distribution of shapes that is an intrinsic property of a soil mapping unit. With a usable shape index, the distribution of shapes can be described and interpreted. Developing such descriptions and interpretations for soil mapping units has been hindered by a lack of quantitative methods for measuring two-dimensional shape. Advances in computer techniques for processing spatial information from soil surveys allows new questions to be raised concerning soil body size and shape. The set of questions include;

1. How can the shape of soil map delineations be described in a quantitative manner?
2. Do soil map delineations of a specific map unit present a range of shapes similar to the range of other soil properties used to define the map unit?
3. What correlations exist between the shapes of delineations and the soil and landscape properties used to define the map unit?
4. What is the relationship between the shape of a delineation and the genesis of the soil body being approximated?

5. Are use and management interpretations for mapping units independent of the shape of individual delineations?
6. Can a better understanding of soil pattern be achieved by developing an objective understanding of the shapes of delineations as elements of pattern?

It is impractical for a single study to address all of the above questions. Relationships between the shape of soil bodies and their genesis and management cannot be studied without an accepted quantitative descriptor of shape. Analysis of soil pattern cannot proceed without preliminary spatial analysis of individual pattern elements, i.e. soil map delineations. Thus, the objectives of this research are:

1. To develop a meaningful, quantitative system for classifying the shape of individual soil map delineations; and
2. To evaluate the characteristics of, and differences between selected soil mapping units in terms of delineation shape.

Very little work by soil scientists has been devoted to the quantitative study of soil body shape. The limited information from soil science was combined with shape indexes provided by literature from a wide variety of disciplines concerned with two-dimensional shape description. Selected parameters and indexes were then calculated for many soil map delineations representing 13 different mapping units. The results were used to characterize the shapes of the soil map delineations. The distributions of a few key shape parameters for each sample were used to characterize the map units and evaluate differences between them.

LITERATURE REVIEW

Several disciplines, including geology, psychology, cybernetics, geography, biology and soil science have been concerned with two-dimensional shape measurement. Examination of research from these disciplines provides a clearer understanding of what shape is, how to measure shape, and the evaluation of process-form relationships. The shapes of soil map delineations and other similar figures have been measured by a variety of procedures.

Shape Definitions

Shape, with respect to soil map delineation properties, must be defined before a meaningful shape descriptor can be selected. Several authors have defined this property. Shapes of biological objects, such as chromosomes and skulls, have been described simply as "outlines with landmarks" (Bookstein, 1978). Landmarks are points on a boundary which are homologous and have anatomical or geometric significance to a figure's definition. Distance and angles between landmarks are used to define the shape of the figure being studied. This approach to shape analysis has not yet been used to study soil map delineation shape. Possible landmarks for delineations are soil map boundary intersections (nodes), long axis endpoints, and area centroid.

Another definition of shape is "the set of properties possessed by any closed figure of at least two dimensions, which has a planar representation, and which possesses precise boundaries" (Boots and

Lamoureux, 1972). This definition fits the characteristics of soil map delineations reasonably well.

The precision of boundary location depends in part on the mappability of map units, and in part on the mapping skills of the soil scientist. Ideally, boundaries are placed at the point of "maximum lateral rate of change" of soil properties used to define a map unit (Knox, 1965). Abrupt, clear, gradual, and diffuse boundary width classes have been developed to account for gradual changes in the natural soil continuum (Hole, 1978). For shape analysis purposes, the soil map boundary position is one dimensional with no width. Area, complexity, and elongation are the spatial properties of soil bodies which soil scientists have concentrated on in the past (Hole, 1953; Fridland, 1965).

Attneave and Arnoult (1956) studied the perception of shape in an attempt to identify measurements made on a figure's boundary with the most psychological relevance. Stenson (1966) defined shape as the sum total of visual cues expressed by a figure's boundary. The spatial properties which he concentrated on were physical measurements such as length of boundary, i.e. perimeter, indexes which measure boundary complexity, and the general elongation of a figure.

Freeman (1977) defines shape as the "composite effect of curvature" of an open or closed curve. Soil map delineations are closed curves, but the boundary segments between boundary intersections, or nodes, may be considered open curves. These open curves represent the form produced by the interaction of two sets of soil processes, one set acting on each side of the boundary. The shapes of the open curves could be studied using techniques for studying the shape of waveforms discussed by Pavlidas (1978) to learn more about process-form relationships in soils.

Use of Indexes

An ideal shape index should have the following properties (Bunge, 1966; Lee and Sallee, 1970; Stoddart, 1965):

1. provides a range of index values distributed along a continuum, which allows meaningful class intervals, and places similar shapes close together and dissimilar shapes far apart.
2. assigns each shape a unique number.
3. does not include less than shape by concentrating on sub-properties alone, i.e. area or figure elongation.
4. does not include more than shape by including elements of pattern such as orientation and puncturedness.
5. is simple to calculate.
6. is readily restated in shape terms.

Lee and Sallee (1970) and Hudson and Fowler (1966) point out the major problem for developing a single numerical value to describe shape completely. The set of all planar shapes has an infinite number of degrees of freedom. The set of all real numbers used to quantify shape has only one degree of freedom.

In spite of this difficulty, many so-called shape indexes have been developed. These indexes can be useful for understanding process-form relationships. Various techniques of shape analysis measure only those portions of total shape information that are relevant to the definitions of the figures being studied. For application to soil map delineations, the requirements of an ideal shape index should still be satisfied, at least until more experience allows correlation of soil processes with certain sub-properties of shape.

There are three general categories of shape measurements. Figure attribute ratios, sometimes called form ratios, are simple mathematical ratios which compare attributes of empirical shapes to those of standard shapes. Direct shape measurements express the spatial relationships between actual preselected boundary points. Boundary following techniques identify portions of the boundary that significantly influence the totality of perceived shape. Each of the three approaches attempts to summarize all the shape information for a figure into a single number or set of numbers.

Figure Attribute Ratios

Attribute ratios are a coarse method of measuring shape. The ratios have been termed "variations on a theme" (Boyce and Clarke, 1964). The most common themes are boundary complexity and figure elongation. These simple indexes are constructed from three primary measurements: area, perimeter, and long axis (Muehrcke, 1978). These values are used to compute a corresponding set of primary measurements for selected standard shapes. The circle and ellipse are very common standard shapes due to the regularity of their boundaries and mathematical definitions of their primary measurements.

Ratios are calculated by comparing a primary measurement for an empirical shape, such as the area of a soil map delineation, to a corresponding primary measurement for a standard shape. Using a circle or square to represent compaction (high area to perimeter ratio), Taylor (1970) suggested that deviations of a shape from compaction include an increase in elongation, boundary indentations, fragmentation, and puncturedness. The latter two deviations may well indeed be a property of delineations, but are elements of pattern rather than shape.

Figure attribute ratios which measure complexity or elongation have received the most attention from soil scientists. Boundary complexity has been measured by taking the ratio of the length of delineation boundary (perimeter) to the perimeter of a circle having an area equal to the delineation's area. Hole (1953) calls this index the soil body pattern index. Classes of shape complexity using this index include very simple (<1.3), simple (1.3-1.7), moderately simple (1.7-3.5), moderately complex (3.5-5.5), and very complex (>5.5) (Hole, 1978).

Elongation is a recognized property of soil bodies which can be measured, in a general sense, by form ratios. One technique uses the ratio of a delineation's longest axis to it's shortest axis (the shortest distance between boundary points which are endpoints of a line perpindicular to the long axis) (Fridland, 1972; Piech, 1980). Classes of elongation using this index include disks (<2), spots (2-5), and stripes (>5) (Hole,1978). Selection of a short axis seems to be quite arbitrary, however.

Bilton (1983) measured the elongation of several delineations of forest soils in Tillamook County, Oregon, using Schumm's elongation ratio. This index is calculated as the ratio of the diameter of a circle having an area equal to a delineation's area to the delineation long axis (Schumm, 1956). The index did not appear to distinguish between elongated soil bodies on ridges and elongated soil bodies along drainageways, however.

Many other combinations of primary measurements could be used to construct form ratios. Basically, they all give general information about a figure's elongation or boundary complexity. Boots and Lamoureux (1972) have assembled an extensive bibliography which describes most of the form ratios used for applied shape analysis. Form ratios are simple to calculate, and they yield a unitless index value. They are sometimes difficult to interpret because two different shapes can have nearly the same index values.

Form ratios are useful as shape indexes when studying figures which have predictable boundary configurations. Research on the shape of drainage basins (Chorley, Malm and Pogorzelski; 1957) and marine atolls (Stoddart, 1965) has shown figure attribute ratios to be of value. Interpretation of the ratios is possible due to past experience with the spatial attributes of these figures. Experience working with these shapes has allowed the selection of appropriate standard shapes, i.e. a lemniscate standard for drainage basins and the ellipse for marine atolls.

Variability of soil body shapes is not predictable, and the actual shape which corresponds to each index value is not yet known.

It is also questionable whether appropriate standard shapes have been used to construct form ratios for soil body shapes. Perhaps with more experience with the use of attribute ratios to measure the shape of soil bodies, it may be possible to select one or more standard shapes that are particularly suited to the shape information characteristic of soil map delineations. It is also possible that a single attribute ratio is very well suited for expressing the shape of soil bodies.

Direct Shape Measurements

Another type of shape index could be called direct shape measurement. These indexes summarize spatial relationships between actual boundary points or areas of a figure. They allow a clearer picture of actual boundary configuration than form ratios, but they tend to be more tedious to calculate. One example, which is similar to the form ratio concept, is the "direct symetric difference metric" (Lee and Sallee, 1970). This method compares an empirical shape to a standard shape (circle) by direct overlay and measurement of the areal union and intersection. The metric was found to group Sudanese villages according to shape well.

It appears that when shape differences are predictable, the metric has merit. Since soil body shape is unpredictable at this time, there is still a good chance that two quite different shapes could have nearly the same index value. Use of modern area overlay procedures as discussed by Monmonier (1982) would make this method less tedious to calculate.

Another method of direct shape measurement is the Vertex Lag method (Bunge, 1966). The method has been tested and evaluated by Fridland (1972) and Stoddart (1965), and it has good potential for soil body shape analysis. The original shape is first approximated by a polygon with some number (n) of vertices which are actual boundary points. Clearly, the greater the number of vertices, the better the approximation of shape, but the amount of calculation increases as well.

The vertices are numbered 1 to n, and each vertex is equidistant from the preceding vertex. The distance between every other vertex is summed, then every third vertex, every fourth, and so on until (n-2)/2 distance sums are calculated. Another set of (n-2)/2 values of the distances squared is calculated using the same lag procedure. For example, a shape approximated by a vertex lag polygon with eight equal length sides and eight vertices has a set of three sums and three sums of squares as the numerical description of shape. A single shape value is not produced.

Fridland (1972) used the vertex lag method to classify the shapes of elementary soil areals (similar to soil map delineations). He used polygons having only eight sides, but although his shapes were grouped reasonably well, he recommended that more vertices should be used to capture the irregular boundary configurations of many soil bodies. He also suggested further testing of the method by analyzing mapped soil bodies from a range of pedo-geographical areas. More experience with the method is also needed to learn how to reduce the sets of distance and distance squared values into a single index value or classification. Stoddart (1965), using an eight sided polygon, found that the vertex lag method was not sensitive enough to capture the subtle shape differences between marine atolls.

The measurements required by the vertex lag method have been made in the past with the manual use of adjustible length bars on hinges, projection of the shape onto a grid base, and manual distance measurements and summation (Bunge, 1966; Fridland, 1972; Stoddart, 1965). Modern methods of boundary digitization and distance calculations using computers enable further advances with the use of this method of shape measurement.

The radial line shape index developed by Boyce and Clarke (1964) is another direct shape measurement. A preselected number of radii are extended outward from the center of gravity of a shape. The sum of the equal angles between radii equals 360 degrees. The radial distances from the center of gravity to the figure's boundary are compared to the distance from the center to the perimeter of a circle having the same area. Stoddart (1965) and Cerney (1975) point out that besides being awkward to calculate, different boundary configurations can produce identical index values. If the index could preserve the distance values which correspond to the position of each radius, the set of values could represent shape. Reducing a set of shape measurements for a figure to a single value often loses the ability to assign a unique number to each shape.

Bribiesca and Guzman (1979) introduced a shape measurement called "the shape number". To determine the shape number for a figure, the boundary must be digitized and stored as a Freeman chain code, i.e. a string of octal digits from and returning to a point of origin (Freeman, 1974). The index number can be independent of area and unique to each shape. Athough the method appears to meet our requirements, applied research with the method is lacking, interpretation of results is not well established, and very specific data input and computer software are required.

Another method of direct shape measurment begins with the construction of a histogram of standardized distances within a figure. Taylor (1970) constructed this sort of histogram for each of four outlines of countries. Moment measures derived from the distributions were used to describe the apparent differences in shape between these figures. Bunge (1966) discusses the principle of using the distances within figures to describe spatial properties and geographic processes.

Boundary Following Techniques

The most recent methods of shape measurement would be classified as boundary following techniques. Boundary digitization is needed, along with algorithms to make many calculations while tracing a figure's boundary. Pavlidas (1978) reviews several of these modern approaches to shape measurement.

Boundary followers are very precise, but it should be remembered that shape measurement need not be more precise than the figure itself (Bunge, 1966). Techniques such as the "tangent angle function" (Bookstein, 1978), Fourier descriptors (Zahn and Roskies, 1972), and the "medial axis transformation" (Blum, 1967; Bookstein, 1978) are well adapted to measuring fine shape detail on enlargements of reality. Soil map delineations, however, are generalized reductions of reality that do not require the highest amount of precision.

One boundary following technique, however, may have particular relevance to shape measurement of soil map delineations. This technique involves measuring the convex deficiency of a figure by isolating individual concavities or indentations (Duda and Hart, 1973). Sklansky (1974) describes the calculations needed for an area deficiency index, indexes of concavity depth and complexity, and an overall convex deficiency index. He also discusses how to find the minimum perimeter polygon, which is a separate polygon having the shape of a rubber band stretched around the outline of an irregular shape. Another name for this figure is the convex hull. His use of convex deficiency measurement, and the irregular, indented nature of most soil map delineations, suggest that further application of this method would be useful. Several algorithms for calculating the convex hull of a digitized boundary are available (Akl and Tousaint, 1978; Jarvis, 1973; and Graham, 1972).

Index Selection

Selecting the best index from a list of potential shape indexes is common to many of the applied shape analysis studies in various disciplines. Matthews (1981) stated several considerations to be made before making this selection. A precise definition of the figures being studied is necessary so that numerical differences have meaning. The attribute of shape being measured should be known. The limitations of the shape index also should be known.

Shape index selection is usually based on the desirability of having a single index value which is easy to calculate, one which selects a median shape which is also the median shape for other indexes, and can be readily restated in shape terms. The selection of the best index for describing the shapes included in a study would consider the ease of interpretation and correlation data between indexes.

One shape index was selected from a list of six indexes in a study of marine atolls (Stoddart, 1965). The selection criteria were a review of each frequency distribution and a comparison of the interpretability of each index. The regular outline and elliptical nature of atolls was a controlling factor for selecting an ellipticity index. Also, all of the nine median atolls for the ellipticity index were in the median class for one or more of the other five index distributions.

Stenson (1966) used factor analytic techniques to find which physical measurement or group of measurements correlated best with perceived complexity of 20 randomly generated forms. Most of the total system variance was accounted for by a single factor. The factor was defined by four indexes of complexity. The indexes are the number of turns on the boundary, length of boundary (perimeter), perimeter squared to area ratio, and the variance of the internal angles of the form. Grouping potential indexes with factor analysis indicates the subproperties of shape which are being measured, or whether the property of shape itself is being captured by one or more potential shape indexes.

A variable selection technique used in atlas construction seems to be applicable to shape index selection. The "key-variable approach" involves the elimination of variables, i.e. shape indexes, which are highly correlated (Monmonier, 1974). Once each index has been calculated for each shape in a sample, all that is needed is the correlation coeficients between each index. The method helps to sort out shape indexes which provide unique information.

The most modern methods available need to be applied to shape analysis problems (Bunge, 1966). Fortunately, computer processing of spatial information is becoming a practical tool for such analyses. Two excellent sources of information about design, development, and management of geographic data processing systems, similar to those needed for shape analysis, are available (Nagy and Wagle, 1979; Monmonier, 1982). Types of data which can be manipulated by this technology include: a) Alphanumeric information such as census reports and map classification descriptions, and b) pictorial or graphic information such as photographs and maps.

Modern, detailed soil surveys provide the necessary information for managing a geographic information system as described above (Soil Survey Staff, 1951). Spatial boundary information and regional definitions are included in each soil survey. In current surveys, the shapes of delineations for a particular soil map unit are described with somewhat subjective phrases.

In this study, potential shape indexes are tested on a sample of well defined shapes, i.e. soil map delineations. The numerical results and frequency distributions for each index are evaluated.

Correlations between the potential indexes are also analyzed. These procedures are followed in order to find those shape indexes which are minimally inter-correlated and maximally interpretable.

The review of literature did not provide a standard method of comparing groups of shapes based on a single index. The available literature on shape analysis does provide information on each of the steps for a shape analysis. Very few previous studies provide results on the complete procedure needed for this study of soil map delineations.

SAMPLING PROCEDURE

Sampling soil map delineations for this quantitative shape analysis project was a four-step process. First a modern, detailed soil survey was selected. Then soil mapping units were selected to provide a wide range of delineation shapes and pedogenic processes. Next the population of delineations for each map unit was identified and sampled. Finally the boundary of each sampled delineation was recorded as an ordered string of (X,Y) coordinates.

The soil survey of Benton County Area, Oregon (Knezevich, 1975) was selected as the source of soil map delineations. The majority of the survey area was mapped with an intensity of an Order II, detailed soil survey. The scale of mapping is 1:20000. Familiarity with the soils and landforms of the area also influenced the survey selection. Map units and delineations were drawn from a single survey in order to control possible inconsistencies of mapping by different soil survey parties. Map units were selected to facilitate overall comparisons between soil series, between phases of the same series on different slopes, and between different series having the same slope ranges.

Of the five factors of soil formation (time, parent material, vegetation, climate and relief), parent material and relief appear to have the most influence on the shape of natural soil bodies. This affected the selection of mapping units. The survey area contains a wide variety of soils formed on different parent materials and landform positions. Soils located on three of the generalized

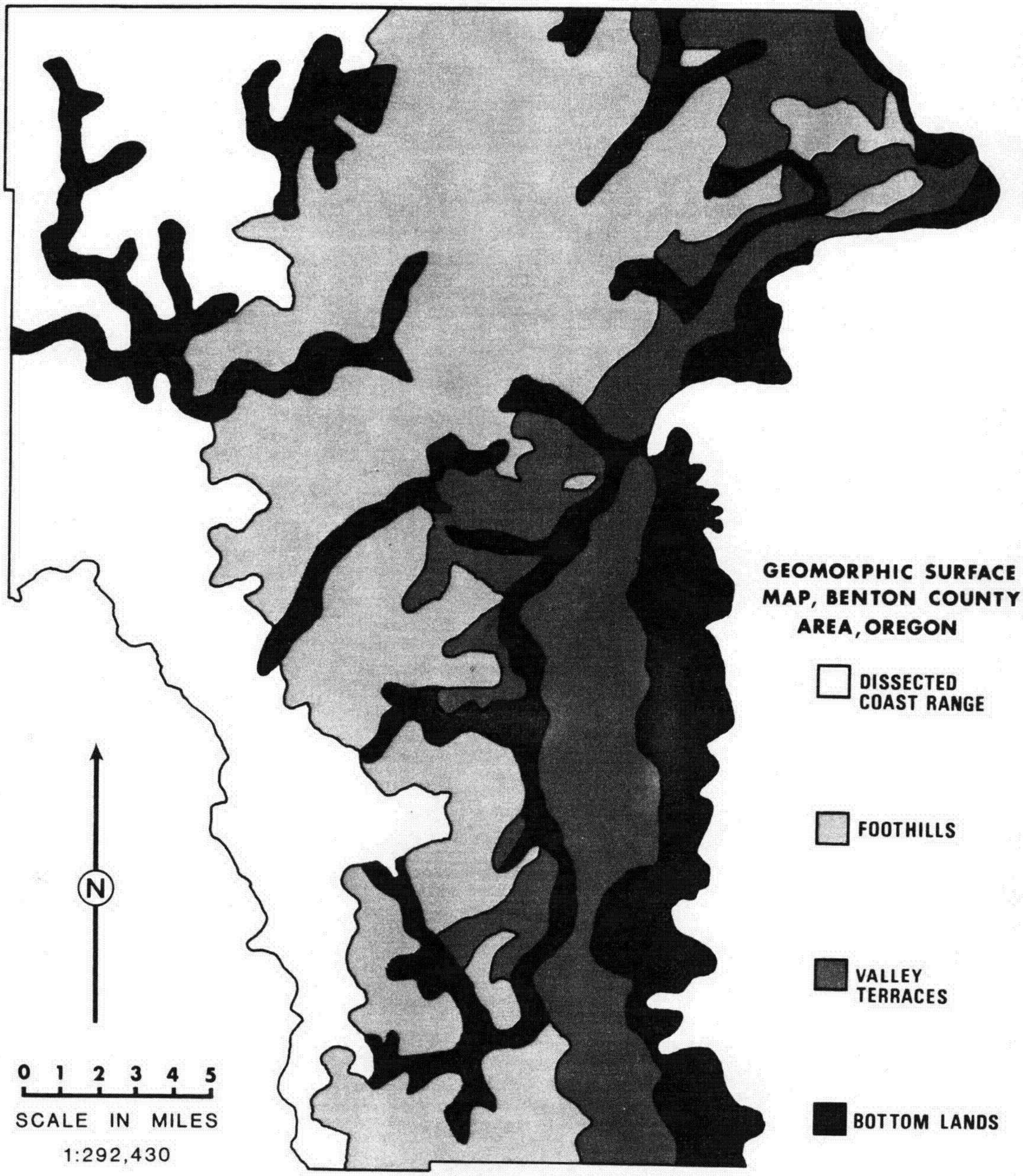


Figure 1. Generalized geomorphic surface map of Benton County Area, Oregon.

geomorphic surfaces shown in Figure 1 were selected for shape analysis: foothills, valley terraces, and alluvial bottom lands. The "dissected coast range" was mapped with a lower intensity than the other three surfaces, thus, no map units from this region were selected.

Five mapping units defined by soil series provided five populations of delineations to sample. The Dixonville and Bellpine series represent soils found on the foothills of the coast range. The Woodburn soil series is found on the younger, valley terraces running parallel to the Willamette River, which defines the east boundary of the survey area. The Chehalis series is located on recent alluvium deposited along the Willamette. The Waldo series represents soils developed on very recent alluvium along tributaries of the Willamette, which drain the coast range and foothills.

Bellpine and Dixonville soils are mapped according to designated slope phases. For both of these series, there are four slope phases, namely 3-12%, 12-20%, 20-30% and 30-50%. Thus, delineations sampled from these eight populations allow comparisons of delineation shape based on slope differences within the same series. In order to sample Bellpine and Dixonville series on a series level, contiguous phases of the respective series were sampled as a single delineation.

The Woodburn series has two slope phases; 0-3 and 3-12%. Only a few delineations of Woodburn 3-12% are found in the the survey area, however. Thus, the population of Woodburn 0-3% delineations was sampled as if that series had only one slope range. Chehalis and Waldo series have only one slope phase (0-3%), so there was no need to combine contiguous slope phases to generate delineations of soil series. Table 1 provides a summary of the properties for the selected soil mapping units. Landscape position, slope, parent material, and degree of weathering are some of the properties of delineations that can be compared on the basis of shape.

The next step was to draw a random sample from each of the thirteen map unit delineation populations. Each delineation of a population was located and identified on the soil survey map sheets. Only whole delineations were sampled, i.e., the boundary of the survey area was not considered part of a delineation's boundary. Each delineation so located was assigned a unique identification number. A random sample was then drawn from this list of numbers for each population of map unit delineations.

The sampling rule was to obtain, if possible, a 50% sample for each mapping unit. The number of delineations in the sample for a map unit could be no less than 25, as was the case for the Willamette 3-12% map unit. The sample could be no greater than 50. These sampling rules were used to ensure large enough samples to cover the range of shapes, but not create so large a sample that it could not be analyzed in the time available. The number of delineations comprising each map unit sample is shown in Table 1.

Each of the sampled delineations was traced from the soil maps and copied onto hard bond paper to increase durability and maintain dimensions of the delineations. The number of delineations on a sheet was arbitrary, and no aspect of orientation or relationship to other delineations was maintained. Each delineation was considered

Table 1. General information for the sampled soil mapping units

MAP UNIT	SYMBOL	SLOPE	SAMPLE SIZE	SOIL CLASSIFICATION	PARENT MATERIAL	LANDFORM
Bellpine	B	3-50%	36	clayey, mixed mesic xeric Haplohumult	sedimentary colluvium	foothills
Dixonville	D	3-50%	42	fine, mixed, mesic pachic ultic Argixeroll	basic igneous colluvium	foothills
Chehalis	C	0-3%	30	fine silty, mixed, mesix cumulic ultic Haploxeroll	recent, mixed alluvium	alluvial bottom land
Woodburn	Wo	0-12%	50	fine silty, mixed, mesic aquultic Argixeroll	silty alluvium	valley terrace
Waldo	Wa	0-3%	49	fine, mixed, mesic fluvaquentic Haplaquall	recent alluvium	along streams and drainage ways

MAP UNIT	SYMBOL	SLOPE	SAMPLE SIZE
Bellpine 3-12%	BC	3-12%	25
12-20%	BD	12-20%	32
20-30%	BE	20-30%	35
30-50%	BF	30-50%	25
Dixonville 3-12%	DC	3-12%	36
12-20%	DD	12-20%	30
20-30%	DE	20-30%	35
30-50%	DF	30-50%	27

an individual.

Each of the 452 delineations was digitized to record the boundaries as ordered strings of (X,Y) coordinates. A 14 inch L-frame, SAC Graf/pen digitizer combined with a tectronix 4051 micro-computer was used. A simple boundary input program was used to digitize the boundaries and save them as vector type polygons. Point mode digitizing was used instead of continuous mode to control the number of boundary points for storage and processing. The origin point for each polygon was arbitrarily chosen. From the origin, points were ordered in a clockwise manner.

METHODS OF SHAPE MEASUREMENT

The sampling procedure produced 452 separate strings of (X,Y) coordinates, each of which represents the shape of a soil map delineation. Using the coordinate strings, 43 potential shape indexes were calculated for each delineation polygon. Some of the indexes have been used previously by other researchers. The others are new indexes derived during the course of this study. All are listed in Table 2.

Primary Measurements and Attribute Ratios

Since primary measurements are the basis for calculating many indexes of shape, their determination was the first step. The three basic primary measurements for a two-dimensional figure are area, perimeter, and long axis. Although these measurements are not actually shape indexes, their values were retained in the list of potential indexes.

Formulas for computing area and perimeter values for an irregularly shaped, digitized polygon are discussed in depth by Monmonier (1982). Delineation area is computed by accumulating a running balance of triangular areas. Areas of successive triangles are computed as a variation of the vector cross product. Perimeter is computed by accumulating the lengths of individual boundary segments using the pythagorean theorem. A small segment of a computer program written in BASIC language is used here to illustrate the computation of area and perimeter values.

Table 2. Symbols and names of potential shape indexes.

PRIMARY MEASUREMENTS	FIGURE ATTRIBUTE RATIOS
AE - Delineation Area	X1 - Hortons Form Ratio
PE - Deineation Perimeter	X2 - Circularity Index (Miller)
AH - Convex Hull Area	X3 - Elongation Ratio (Schumm)
PH - Convex Hull Perimeter	X4 - Ellipticity Index (Stoddart)
LA - Delineation Long Axis	X5 - Minimum perimeter index
AV - Area of Concavities	X6 - Cartographic Complexity
PV - Perimeter of Concavities	X7 - Elliptical Irregularity
WA - Weighted Average Concavity Area	X8 - Coefficient of Dissection
	X9 - Boundary Complexity Index

Z1 thru Z9 are the same as X1 thru X9 except the indexes are calculated for convex-hull rather than delineations.

CONVEX DEFICIENCY	VERTEX-LAG
P2 - Area Deficiency Ratio	VLI - Vertex-Lag Index
P3 - Extra Perimeter Ratio	VLC - Vertex-Lag Classification
P4 - Concavity Perimeter Ratio	CVLI - Convex-Hull VLI
P5 - Concavity Dominance Index	CVLC - Convex-Hull VLC
MI - Concavity Depth Index	P1 - CVLI - VLI
MX - Concavity Complexity Index	SIX - Shape classification

INDEX COMPARISONS

Q1 - Convex comparison (Stoddart's Ellipticity Index)
Q2 - Convex comparison (Schumm's Elongation Ratio)
Q3 - Convex comparison (Elliptical Irregularity Index)
Q4 - Convex Comparison (Fridland's Coef. of Dissection)
Q5 - Convex comparison (Boundary Complexity Index)

For a polygon with K-1 vertices it is necessary to set X(K) equal to X(1) and Y(K) equal to Y(1) to insure boundary closure.

```
10 PE = 0 : AE = 0
20 For I = 1 to K-1
30 D = SQR(((X(I)-X(I+1))^2) + ((Y(I)-Y(I+1))^2))
40 PE = PE+D
50 AE = AE+((X(I)*Y(I+1)) - Y(I)*X(I+1)))
60 NEXT I
70 PE = PE/100 : AE = (AE*0.5)/10000

PE = DELINEATION PERIMETER (CM)
AE = DELINEATION AREA (CM^2)
```

The length of each delineation's long axis was also determined and recorded. The method used in this study was to use the capability of an electronic digitizer to measure the distance between potential long axis endpoints. Trial and error was used until a maximum distance was found. The long axis value was recorded and manually input into the final primary measurement data set. The actual long axis endpoints were not recorded. Later in the study, a short program was written to calculate the long axis distances for comparison with the manually computed distances. The long axis values computed both ways were very similar. A simple program written in BASIC language illustrates the calculation of long axis when the polygon is represented as a coordinate string.

```
10 LA = 0
20 FOR I = 1 TO K-1
30 FOR J = 1 TO K-1
40 D = SQR(((X(I)-X(J))^2) + ((Y(I)-Y(J))^2))
50 D = D/100
60 IF D>LA THEN LA=D
70 NEXT J
80 NEXT I

LA = Long axis (cm)
```

The three primary measurements were used to calculate several figure attribute ratios which describe various components of delineation shape. The standard shapes used in this study were the circle and the ellipse. Delineation primary measurements were used to calculate similar primary measurements for the standard shapes. Selected primary measurements were placed in equations which mathematically compare an irregular delineation shape to corresponding standard shapes. Nine ratios, identified as X1-X9, were used in this study. Their definitions and formulas are given below:

AE - AREA OF DELINEATION POLYGON

PE - PERIMETER OF DELINEATION POLYGON

LA - LONG AXIS OF DELINEATION POLYGON

X1 - HORTON'S FORM RATIO (HORTON, 1932)

Defined as delineation area divided by long axis squared.

X1 = AE/LA^2

Index values range from 0.0 to 1.0

X2 - MILLER'S CIRCULARITY RATIO (MILLER, 1953)

Defined as delineation area divided by the area of a circle with a perimeter equal to delineation perimeter.

X2 = AE/(PI*((PE/(2*PI))^2))**

Index values range from 0.0 to 1.0

**PI = 3.14

X3 - SCHUMM'S ELONGATION RATIO (SCHUMM, 1956)

Defined as the diameter of a circle with an area equal to delineation area, divided by delineation long axis.

X3 = (2(SQR(AE/PI)))/LA

Index values range from 0.0 to 1.0

X4 - STODDART'S ELLIPTICITY INDEX (STODDART, 1965)

Defined as the delineation long axis, divided by the minor axis of an ellipse whose area and long axis are equal to delineation area and long axis.

X4 = LA/(2(AE/(PI*(LA/2))))

Index values range from 1.0 to infinity.

X5 - MINIMUM PERIMETER INDEX (THIS PAPER)

Defined as two times the delineation long axis, divided by the delineation perimeter.

X5 = (2*LA)/PE

Index values range from 0 to 1.0.

X6 - CARTOGRAPHIC COMPLEXITY INDEX (THIS PAPER)

Defined as delineation perimeter, divided by the calculated maximum boundary length (MBL) that fits in a circle whose diameter equals the delineation long axis.

X6 = PE/MBL

MBL = 2*d+((PI*((LA/2)^2))/d) where, d=0.3cm.

Index values range from 0 to 1.0.

The variable d is a factor of map distance for a soil map at a scale of 1:20000. The factor is equal to the diameter of a circular minimum sized map delineation (National Cooperative Soil Survey, 1980).

X7 - ELLIPTICAL IRREGULARITY INDEX (THIS PAPER)

Defined as perimeter of an ellipse with a long axis and area equal to the delineation long axis and area, divided by delineation perimeter.

X7 = ((2*PI)*SQR(((LA/2)^2 + ((4*AE^2)/((PI*LA)^2)))/2))/PE

Index values range from 0 to 1.0.

The equation for the perimeter of an ellipse, given a value for area and long axis is given by Weast, et al. (1964).

X8 - FRIDLAND'S COEF. OF DISSECTION (FRIDLAND, 1972)

Defined as delineation perimeter, divided by the perimeter of a circle whose area is equal to delineation area.

X8 = PE/3.54*SQR(AE)

Index values range from 1.0 to infinity.

X9 - BOUNDARY COMPLEXITY INDEX (PIECH, 1980)

Defined as the delineation perimeter squared, divided by the area of a circle whose perimeter is equal to delineation perimeter.

X9 = PE^2/(4*PI*AE)

Index values range from 1.0 to infinity.

(reciprocal of an index used by Miller (1953))

Construction and Use of the Convex Hull

The next step was to use the delineation polygon coordinate strings to find the coordinate strings which define each delineation convex hull. The convex hull of an irregularly shaped polygon represents the relatively smooth shape formed by conceptually stretching a rubber band around the polygon (Figure 2). The coordinate string which defines the convex hull for a delineation is a subset of the delineation coordinate string.

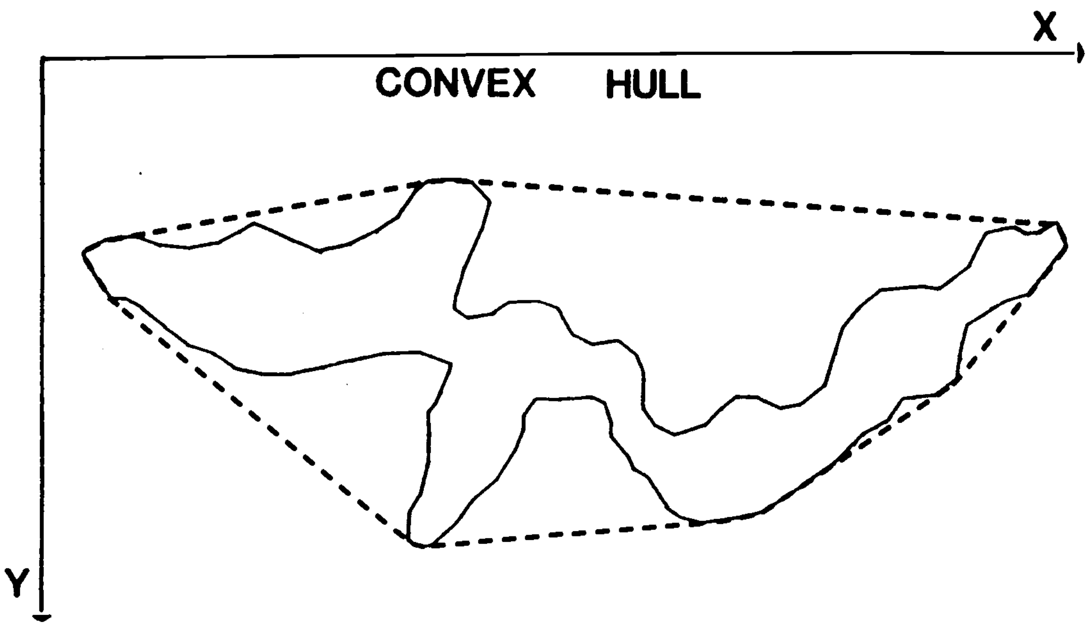


Figure 2. A delineation polygon and convex hull.

The computer algorithm used to determine the set of points for each convex hull is similar to the approach taken by Akl and Tousaint (1978). First, the delineation string is sorted so that the origin, or (X(1),Y(1)), is the point with the minimum X value. For a polygon with (K) number of vertices, start at J = 1 and do (a)-(f) below for every three consecutive points J, J+1, J+2 until K is reached. The first (X, Y) coordinate in the string is equal to the last coordinate to insure boundary closure. The algorithm places all points on the convex hull into a new array of X and Y coordinates.

```
(a) S=((Y(J+1)-Y(J))*(X(J+2)-X(J+1))) +
                        ((X(J)-X(J+1))*(Y(J+2)-Y(J+1)))
(b) IF S<=0 AND J+2=K : STOP
(c) IF S<=0 THEN J=J+1 : GO TO (a)
(d) ADD POINT J+1 to CONVEX HULL POINT ARRAY
(e) IF J>1 THEN J=J-1
(f) GO TO (a)
```

The new convex hull point strings were used to calculate the area (AH) and perimeter (PH) of each convex hull. The calculations were made in the same manner as delineation area and perimeter were calculated. The value for long axis of a convex hull is the same as the corresponding delineation long axis. AH and PH are considered primary measurements, and when mathematically compared to AE and PE, another primary measurement, area of concavities (AV), can be computed. AV is simply the difference between AE and AH. Therefore;

$$AV = (AH-AE)$$

The perimeter of concavities or concave perimeter (PV) can also be determined, but this requires the sum of concavity "mouth" distances. To do this for a delineation, the two coordinate strings representing the delineation and convex hull polygons are sorted so that the origin for each string is the same (X, Y) coordinate. The strings are simultaneously examined, point by point. The point at which the delineation begins to deviate from the convex hull marks an endpoint for a concavity mouth. The point at which the delineation and convex hull converge again represents the other endpoint for the "mouth". The straight line distance between endpoints is called the "mouth" distance. Consecutive points in the delineation string which connect the two endpoints define the spatial characteristics for that single concavity. A delineation can have none or several concavities. Therefore;

PV = (PE-(PH-TM)), where TM = the sum of concavity mouth distances for a delineation.

Once the string of (X, Y) coordinates defining each convex hull was known, the same nine figure attribute ratios were determined for the 452 convex hulls. These are identified as Z1-Z9. The formulas used to calculate them are identical to the formulas for X1-X9, except that AH and PH are substituted for AE and PE respectively (Table 2). Long axis values remain the same. The figure attribute ratios calculated for each convex hull, along with their identification symbol, are listed below:

Z1 - HORTON'S FORM RATIO

Z2 - MILLER'S CIRCULARITY RATIO

Z3 - SCHUMM'S ELONGATION RATIO

Z4 - STODDART'S ELLIPTICITY INDEX

Z5 - MINIMUM PERIMETER INDEX

Z6 - CARTOGRAPHIC COMPLEXITY INDEX

Z7 - ELLIPTICAL IRREGULARITY INDEX

Z8 - FRIDLAND'S COEFFICIENT OF DISSECTION

Z9 - BOUNDARY COMPLEXITY INDEX

The convex hull allows the development of seven additional shape indexes which essentially measure aspects of convex deficiency. These include the ratios P2-P5 and the indexes identified as WA, MI, and MX. The construction and use of these indexes is very similar to work done by Sklansky (1974). Each index is defined below:

P2 - AREA DEFICIENCY RATIO

Defined as the ratio of delineation area to convex hull area expressed as percent.

P2 = (AE/AH)*100

Index values range from 0 to 100%

P3 - EXTRA PERIMETER RATIO

Defined as the difference between delineation perimeter and the convex hull perimeter, expressed as percent of hull perimeter.

P3 = ((PE-PH)/PH)*100

Index values range from 0 to 100%

P4 - CONCAVITY PERIMETER RATIO

Defined as the ratio of perimeter of concavities to total delineation perimeter, expressed as percent.

P4 = (PV/PE)*100

Index values range from 0 to 100%

P5 - CONCAVITY DOMINANCE INDEX

Defined as the ratio of the average sized concavity to the total area of concavities for a delineation, expressed as percent.

P5 = (WA/AV)*100 (WA is described below)

Index values range from 0 to 100%

WA - WEIGHTED AVERAGE CONCAVITY AREA

Defined as the average sized concavity for a delineation. A weighted average is used to minimize the effect of small concavities and emphasize the larger concavities as having more impact on a delineation's shape.

WA = (ΣAi^2)/AV, where Ai equals the area of the ith concavity and AV equals the total area of concavities (AH-AE).

Index values range from 0.0 to infinity

Example: A digitized polygon has 5 concavities.

#	Area	% AV	Area X %AV
1	35.4	0.52	18.32
2	23.7	0.35	8.21
3	6.6	0.10	.64
4	2.3	0.03	.08
5	.4	0.01	.00
Total	68.4	1.00	27.25

Arithmetic Average = 13.68 WA = 27.25

The MI and MX shape indexes were developed to account for the nature of a delineation's concavities. MI is an index of concavity complexity and MX is an index of concavity depth. The subsets of the delineation boundary point string which represent concavities were identified by direct comparison of the convex hull and delineation (X,Y) coordinates. The two indexes are described below:

MX - CONCAVITY DEPTH INDEX

For one delineation, the depth of each concavity is measured from the midpoint of the mouth to the most distant boundary point on the concavity boundary. Each depth is multiplied by the ratio of that concavities area to the total area of concavities for that delineation. The sum of the weighted depths is equal to MX.
Index values range from 0 to infinity.

MI - CONCAVITY COMPLEXITY INDEX

A measure of the relative complexity of the boundaries of the concavities for a delineation. For each concavity, the most distant boundary point from the midpoint of the mouth is identified. The linear distance from the most distant point to each endpoint for the mouth is calculated and summed. The distance represents a conceptually smooth concavity. The distance is divided by the actual boundary distance. This ratio is weighted by multiplying by the ratio of that concavity's area to the delineation's area of concavities. The sum of weighted complexity values equals the final index value. The boundary complexity ratio for the largest concavity will dominate the MI value.

Index values range from 0 to infinity.

Several indexes were developed to learn more about the figure attribute ratios and how they vary when calculated for the original delineations and their respective convex hulls. Q1-Q5 are these direct numerical comparisons and each is briefly explained below:

Q1 - (X4-Z4)/Z4

Defined as the ratio of the ellipticity of a delineation minus the ellipticity of it's convex hull to the convex ellipticity. Ellipticity is measured by Stoddart's Ellipticity ratio.

Index values range from 0.0 to infinity

Q2 - Z3-X3

Defined as the absolute difference between a delineation and it's convex hull with respect to form, or elongation as measured by Schumm's elongation ratio.

Index values range from 0.0 to 1.0

Q3 - Z7-X7

Defined as the absolute difference between a delineation and it's convex hull with respect to the elliptical irregularity index.

Index values range from 0.0 to 1.0

Q4 - (X8-Z8)/Z8

Defined as the ratio of delineation complexity minus the boundary complexity of it's convex hull to the convex hull complexity. Boundary complexity is measured by Fridland's coefficient of dissection.

Index values range from 0.0 to infinity

Q5 - (X9-Z9)/Z9

Defined as the ratio of delineation complexity minus the boundary complexity of it's convex hull to the convex hull complexity. Boundary complexity is measured by the Boundary complexity index.

Index values range from 0.0 to infinity

Vertex Lag Measurements

The vertex lag method of shape measurement provides a quantitative assessment of actual boundary configuration for a delineation. The method is shape preserving, because spatial relationships between boundary points are measured. The accuracy of the method depends on the accuracy of the approximating vertex lag polygon in capturing the majority of a delineation's shape information. In a sense, the method measures boundary complexity, but the nature and relative positions of concavities, i.e. complexity, of a delineation are included also.

Any polygon can be approximated by a vertex lag polygon having some number, n, of equal length sides. In the past, eight sides have been used to approximate shapes (Bunge, 1966; Fridland, 1972; Stoddart, 1965). Increasing the number of vertices allows a more exact representation of the often irregular shapes of soil map delineations. This is practical, however, only by using a computer to generate vertex lag coordinate strings and to perform at high speed all the necessary calculations. Further benefits of computerization include more accurate shape descriptions and less chance of human error.

In this study a 22-sided vertex lag polygon was used to approximate each delineation polygon. Starting with an arbitrarily chosen vertex, each vertex was numbered from 1 to 22.

The lag procedure requires calculation of distances between successive pairs of points,as indicated in Figure 3. Each lag step has 22 pairs of vertices which are used to calculate 22 distances. Because there are (n-2)/2 unique sets of pairs, or "lags", a 22-sided vertex lag polygon yields ten lags, each having a set of 22 distance values (Table 3).

The next step was to calculate the sum of the 22 distance values for each lag. Then the sum of the squares of the 22 distance values was calculated for each lag. This procedure produced

Table 3. Pairs of vertices used for vertex lag distance calculations

	Vertex Number																				
lag 1 -	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2
lag 2 -	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3
lag 3 -	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4
lag 4 -	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4	5
lag 5 -	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4	5	6
lag 6 -	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4	5	6	7
lag 7 -	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4	5	6	7	8
lag 8 -	10	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4	5	6	7	8	9
lag 9 -	11	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4	5	6	7	8	9	10
lag 10 -	12	13	14	15	16	17	18	19	20	21	22	1	2	3	4	5	6	7	8	9	10	11

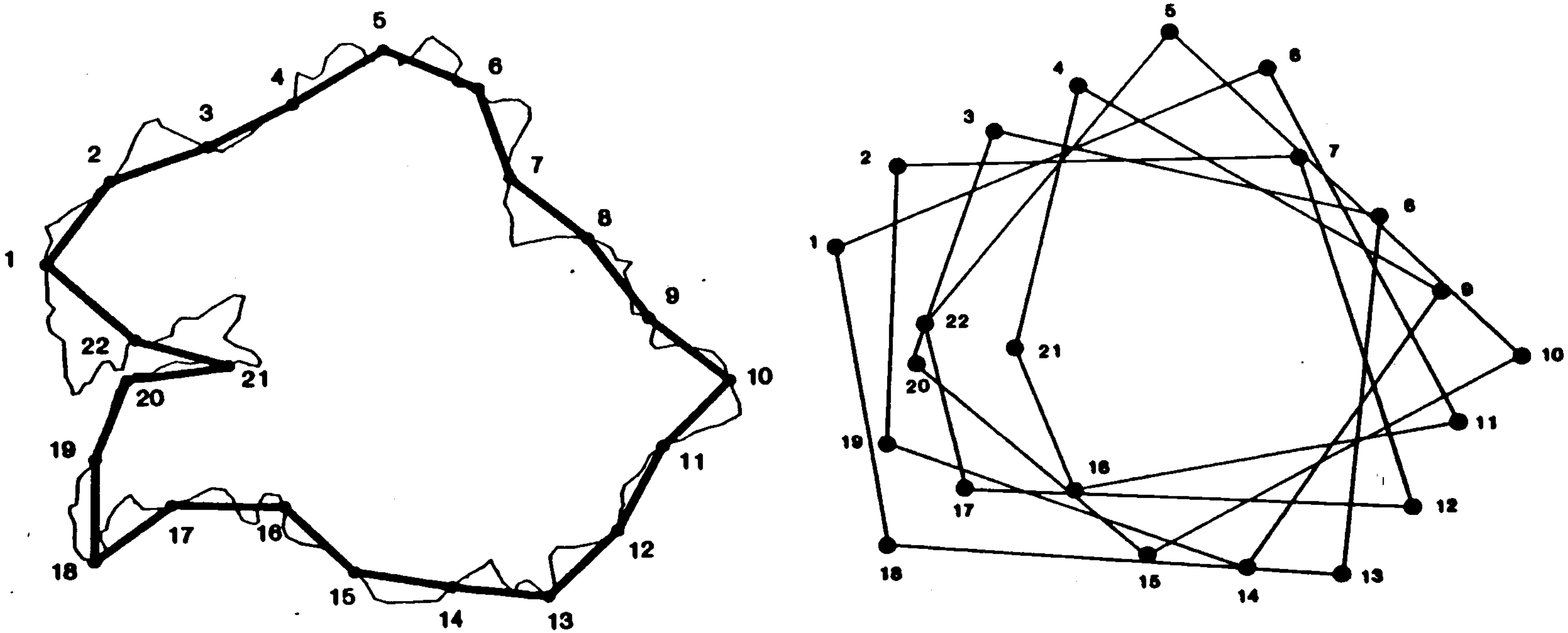


Figure 3. Vertex lag polygon and lag-4 vertex pairs

20 values which are unique to the shape of the approximating polygon. If two delineations have the same set of 20 values, their shapes would be essentially identical, and the shapes of their approximating vertex lag polygons would be exactly identical.

The final step in the vertex lag procedure was to reduce the set of 20 sums and sums of squares to a single index value to represent shape. One method of doing this was to compare the set of values for a delineation vertex lag polygon with the set of values for the vertex lag polygon of a circle. Because each edge of a vertex lag polygon is equal in length, the distance measurements for all polygons were standardized by dividing by their respective edge lengths before computing sums and sums of squares. This permitted comparison of shapes of different size, and was necessary in order to derive a single index value from the set of sums and sums of squares.

The following is a portion of the BASIC program used to calculate sums and sums of squared distance values from the vertex lag polygons. ES is the edge length for a respective vertex lag polygon. Adding 10% of ES to ES was done before standardizing the set of vertex lag values to units of edge length. This was done in order to make sure that a delineation curve fell under the curve for a circle for subsequent index calculations.

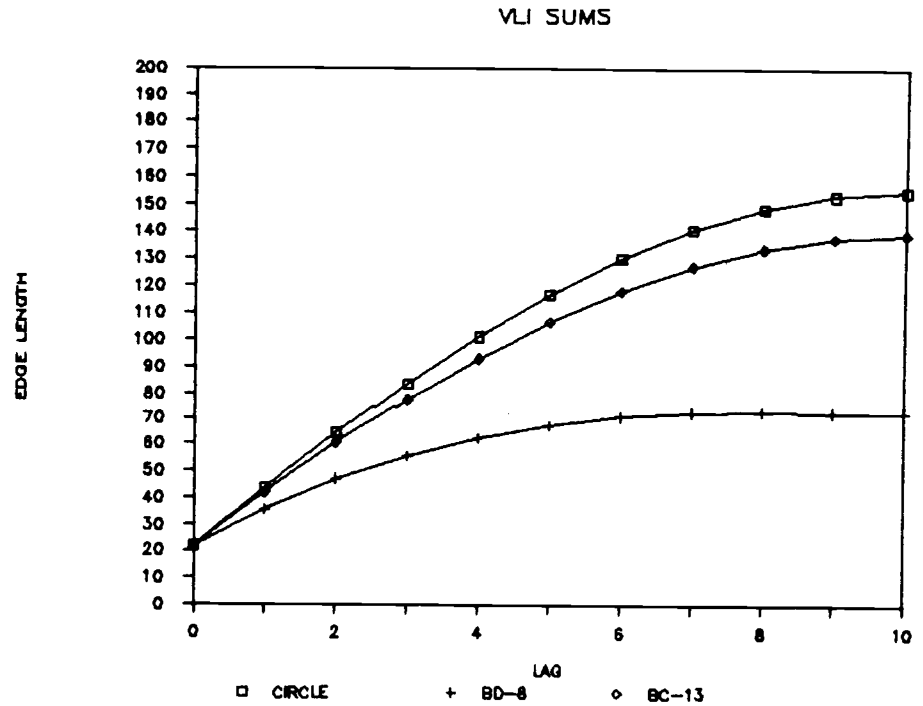
```
125 ET=(ES*.1)
126 ES=ES+ET
130 FOR I=1 TO 10
140 J=0 :K=0
150 FOR R=1 TO 22 :N=R+(I+1) :IF N>22 THEN N=N-22
160 D1=SQR(((BX(R)-BX(N))^2)+((BY(R)-BY(N))^2))
```

```
170 D1=D1/ES :D2=D1^2
180 J=J+D1 :K=K+D2
190 NEXT R
200 DD(I)=K! :D(I)=J!
210 NEXT I
```

Data for a circle with a diameter of 2 cm is shown in Table 4. These data are plotted against the lag number in Figure 4. The results for two delineations are also plotted. Because the circle boundary is perfectly smooth, it's graph always forms the upper bound in each of the two graphs. The more complex and irregular the delineation polygon, the greater the distance between the delineation graph and that of a circle. This is true for both the plot of sums and the plot of sums of squares. The generalized curves in each graph are very similar in shape. Thus, the ratio of the area beneath any given delineation curve to the area beneath the standard curve for a circle is the basis of the Vertex Lag Index (VLI).

TABLE 4. Set of 20 vertex lag values for a circle. Normalized distance is in units of edge length.

	RAW		NORMALIZED	
	SUMS	SUMS OF SQUARES	SUMS	SUMS OF SQUARES
LAG - 1	12.40	6.99	43.5	86.2
LAG - 2	18.28	15.19	64.2	187.5
LAG - 3	23.79	25.72	83.5	317.4
LAG - 4	28.82	37.75	101.2	466.0
LAG - 5	33.25	50.28	116.8	620.5
LAG - 6	37.02	62.30	130.1	769.0
LAG - 7	40.02	72.80	140.6	899.0
LAG - 8	42.22	81.02	148.3	1000.2
LAG - 9	43.55	86.23	153.0	1064.5
LAG - 10	44.00	88.00	154.6	1086.4



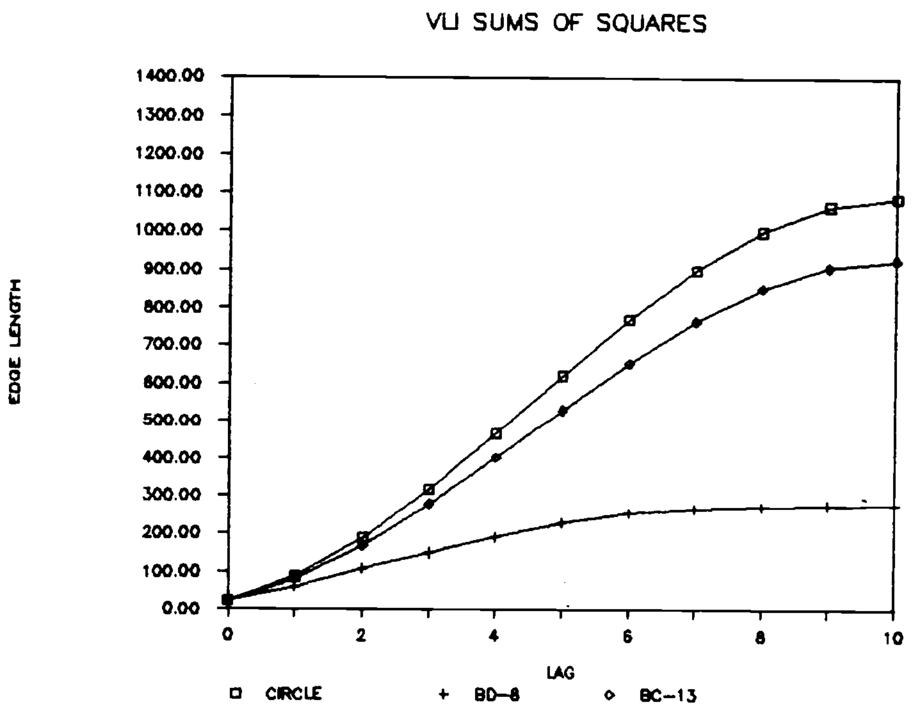


Figure 4. Plots of normalized sums and sums of squares for a circle and two delineation polygons.

For each plot, the areas under the graphs for a circle were set equal to 1000. The area for a delineation's sum curve was expressed as an integer proportion of the circle's sum curve. The same was done for the graph of sums of squares. The resulting two integer proportions were averaged to yield the final VLI index. The potential range of values is 0 to 1000.

Another attempt to reduce the 20 distance values to a single integer is called the vertex lag classification (VLC). The classification assigns each delineation shape a value between 1 and 50. A low value indicates a very complex boundary configuration, and a high value represents a very regular or smooth boundary configuration.

The first step in the classification procedure was to find the minimum and maximum values for each lag sum and sum of squares, considering the ranges of values generated by analysis of all 452 delineations. These values and their respective ranges are shown in Table 5. The minimum values were rounded down to the nearest integer value, and the maximum values were rounded up to the nearest integer. Then each of the 20 ranges was divided into 50 equal class intervals. The widths of the class intervals for each range are also shown in Table 5. Each lag sum for a delineation was assigned to the appropriate class (1-50). This was also done for the sums of squares. The resulting 20 classification values were then averaged for a final VLC index value.

Table 5. Minimum and maximum values used to construct 50 class intervals for each lag step.

	SUMS			SUMS OF SQUARES		
LAG	MIN.	MAX.	CLASS INTERVAL	MIN.	MAX.	CLASS INTERVAL
1	30	46	.32	46	9	.96
2	37	66	.58	69	197	2.56
3	41	84	.90	83	328	4.90
4	45	102	1.14	105	475	7.40
5	49	117	1.36	132	628	9.92
6	51	129	1.56	142	774	12.64
7	53	139	1.72	151	904	15.06
8	49	147	1.96	124	1005	17.62
9	43	151	2.16	98	1069	19.42
10	40	153	2.26	86	1091	20.10

The actual delineation shapes included in this experiment determined the amount of boundary complexity represented by each of the shape classes of VLC. Thus, delineation sampling determined the amount of complexity equating to a VLC value of 1, and the degree of regularity yielding a VLC value of 50.

The vertex lag procedure was also used to calculate an index number and classification value for each of the convex hull coordinate strings. These indexes are identified as CVLI and CVLC. The possible ranges for CVLI and CVLC are 0 to 1000, and 1 to 50, respectively.

Direct comparison of VLI values and CVLI values sheds some light on the sensitivity of the vertex lag method as to how well it differentiates between the indented delineation shape and the convex shape of the hull. The index (P1) for this comparison is the absolute numerical difference between CVLI and VLI.

$$P1 = CVLI - VLI$$

Index values range from 0 to 1000.

SIX is another shape classification index that utilizes vertex lag information. SIX, however, includes two major components of shape: elongation, as measured by Schumm's Elongation ratio on the convex hull (Z3), and complexity, as measured by the Vertex Lag Classification (VLC). The range of possible Z3 values (0.0 to 1.0) was divided into three class intervals. VLC values, 1 to 50, were divided into ten class intervals. A matrix was formed with the three elongation classes on the Y axis and ten VLC classes on the X axis. The two index values were used together to assign each delineation a value of 1-30 within this matrix. Figure 5 illustrates how SIX classifies shape.

Summary of Indexing Procedures

The systematic analysis of each delineation polygon is illustrated in figures 6 and 7. The delineation polygon was first digitized, and the string of X,Y coordinates was used to create a corresponding convex hull. Primary measurements were calculated for both the soil map delineation and the convex hull. Figure attribute ratios were also calculated for each delineation and the hull. Comparison of some of the figure attribute ratios generated additional indexes. Indexes of convex deficiency were also calculated by direct comparison of a polygon with the convex hull. Finally, both the delineation polygon and the convex hull

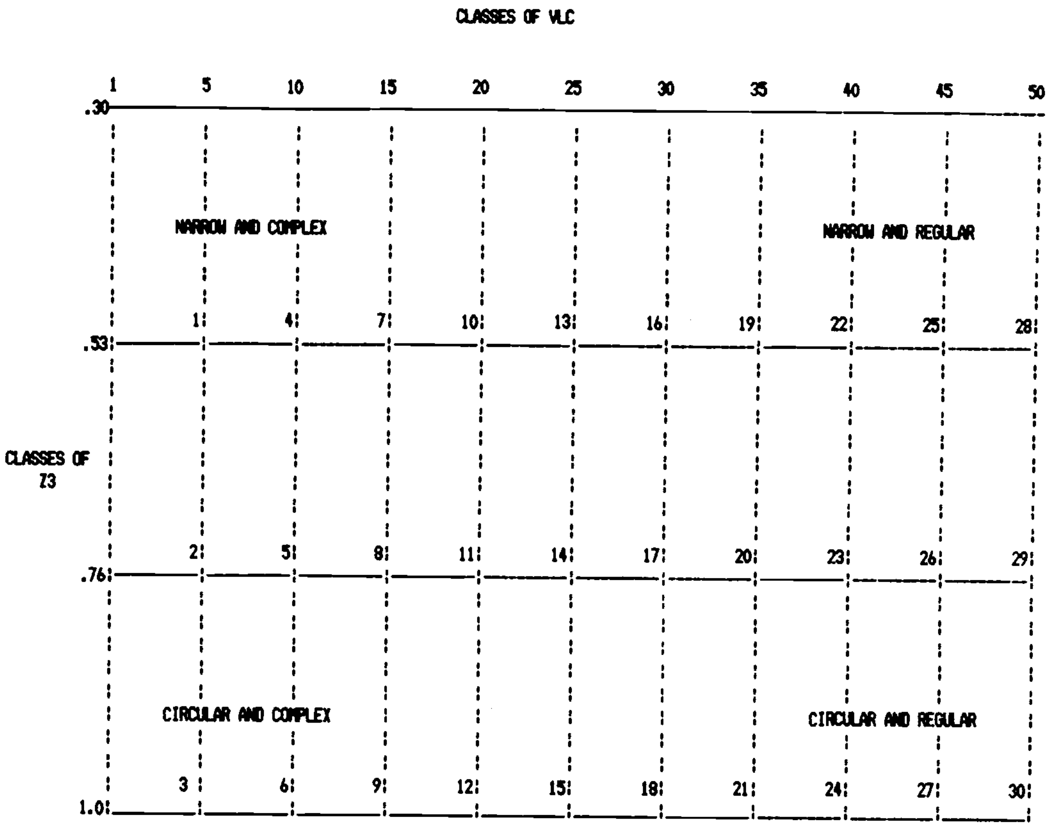


Figure 5. Classification matrix used to determine values for SIX.

were used to create 22-sided vertex lag polygons. These polygons yielded distance measures which were then used to construct additional indexes of delineation boundary complexity. All together, 43 indexes were calculated for each of the 452 soil map delineations in this study.

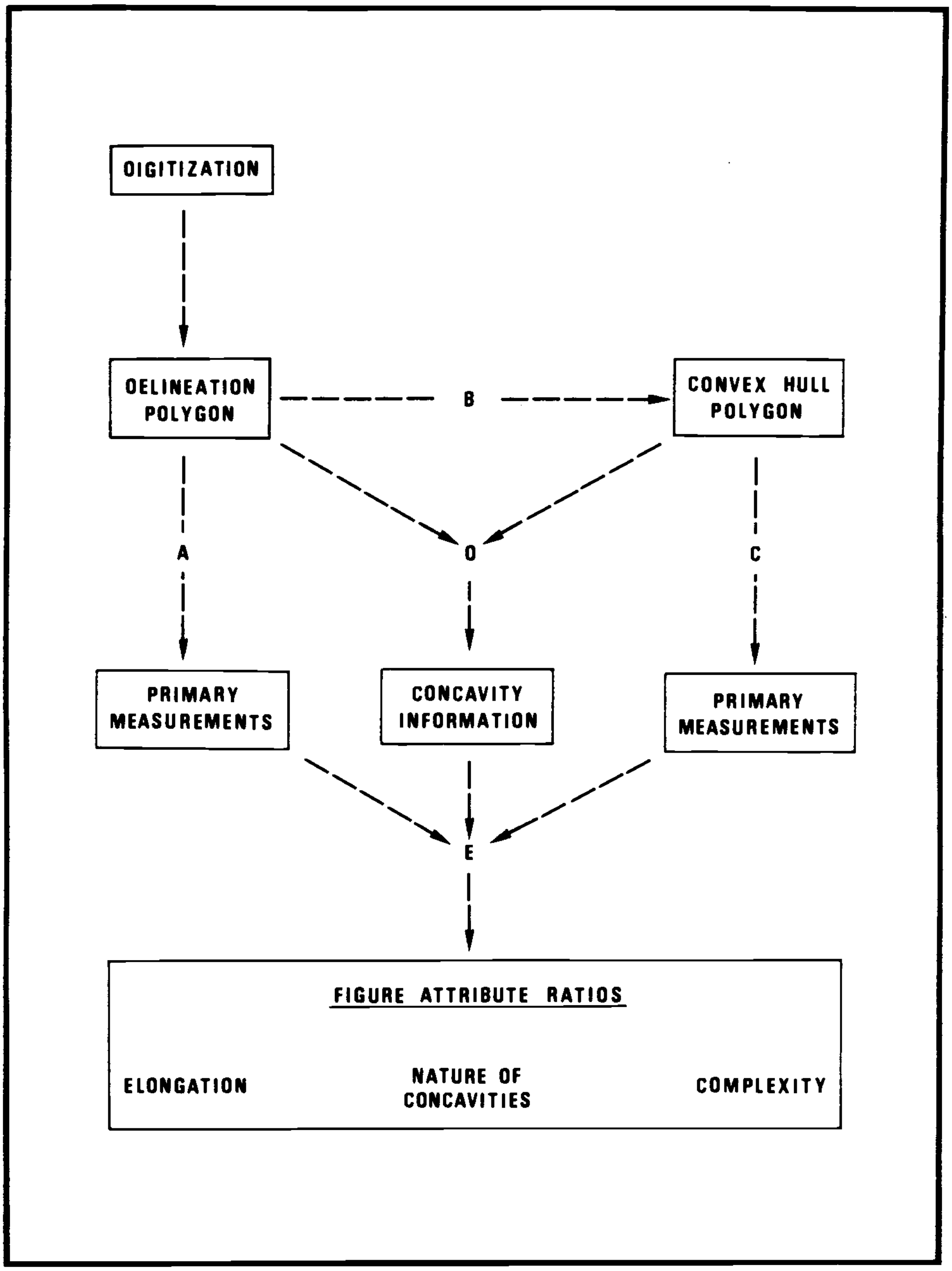


Figure 6. Steps for calculating all figure attribute ratios.

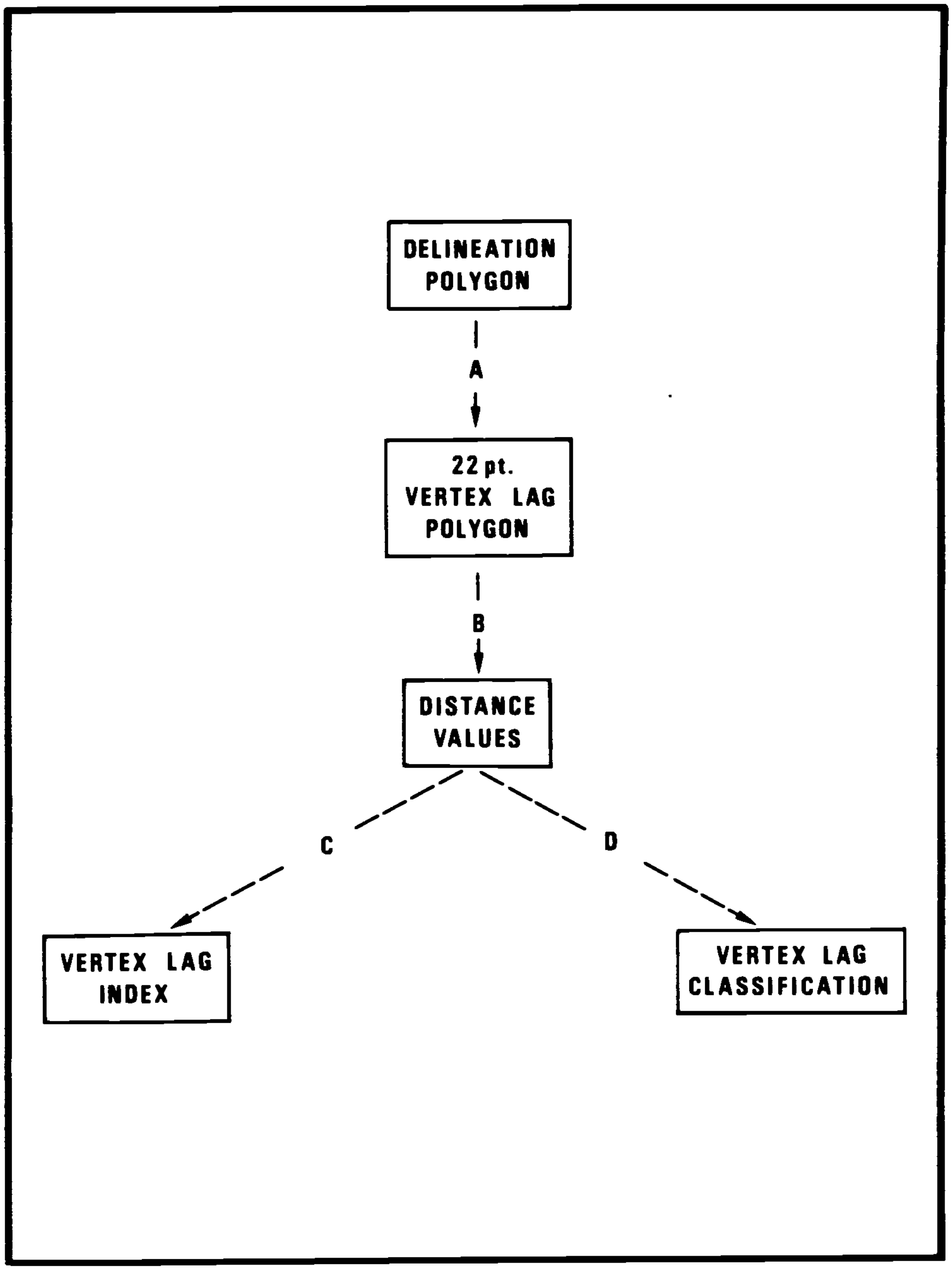


Figure 7. Steps for calculating the vertex lag index and classification.

INDEX EVALUATION

Each of the 43 potential indexes was calculated for each digitized soil map delineation. The data were then used to run a factor analysis to group similar indexes and clarify which spatial properties were measured. Descriptive statistics illustrate the range and distribution of values that each index provides. Spearman's rank order correlation coefficients help to identify similar and dissimilar indexes. These three statistical tools were used to select only indexes that provide interpretable and unique spatial information about soil map delineations. Only the selected indexes were used for in-depth comparisons of delineations within and between the sampled soil mapping units.

The desired result of the evaluation process was to find at least one index that adequately measured shape. The definition of shape used was the sum total of all visual cues on a figure's boundary. Some visual cues that became apparent during tracing and digitization were elongation, amount and depth of concavities, and overall boundary complexity. The list of potential indexes includes shape preserving methods, figure attribute ratios, i.e. shape non-preserving, convex deficiency measurements, and primary meaurements.

Factor Analysis

Factor analysis of the entire data set, based on the varimax criterion (Kendall, 1975), reduced the number of indexes from 43 correlated variables to 6 independent factors. The factor loadings are shown in Table 6. Each factor is defined by indexes with relatively high factor loadings on it. A value of + or - 0.50 was used to distinguish between high and low loadings. Indexes either have a high loading on only one factor, load heavily on more than one factor, or do not load heavily on any of the factors. Most indexes, however, load heavily on only one factor.

The percent of variance accounted for by each factor and the indexes with high loadings on each factor are shown in Table 7. This table indicates that most of the variance in spatial properties of soil map delineations is accounted for by the first three factors (86.1%). Factors 4, 5 and, 6 have a cumulative percent of only 13.9%. Table 7 also shows the groups of indexes that measure similar properties. The next step was to select the most interpretable indexes from these groups.

Factor 1 indexes -- Factor 1 is the dominant factor in the analysis accounting for 50.3% of the variance within the entire data set. Therefore, the type of indexes in factor 1 capture the dominant spatial property that soil map delineations have. There are 14 indexes that have relatively high loadings on this factor. Factor loadings for Q2, Q5, and P2 are higher on factor 5 than factor 1. These indexes are not considered as factor 1 indexes. The descriptive statistics for the remaining 11 indexes are included in Table 8. These statistics help one to interpret each index better, and to begin discarding similar and undesirable indexes.

TABLE 6. Varimax rotated factor matrix. Values in parentheses indicate high loadings.

INDEX	FACTOR 1	2	3	4	5	6
AE	0.18	(0.91)	0.14	-0.13	-0.12	0.11
PE	0.29	(0.89)	0.05	-0.27	0.14	0.18
LA	0.23	(0.78)	-0.15	-0.15	0.26	0.37
AH	0.19	(0.95)	0.09	-0.15	0.03	0.06
PH	0.23	(0.88)	0.03	-0.27	0.16	0.25
AV	0.20	(0.89)	0.01	-0.17	0.26	-0.01
PV	0.30	(0.88)	0.04	-0.27	0.15	0.18
WA	0.27	(0.67)	-0.09	0.00	0.45	0.13
VLI	(-0.87)	-0.23	-0.02	0.34	-0.19	-0.04
VLC	(-0.87)	-0.24	-0.02	0.33	-0.19	-0.05
P1	(0.91)	0.13	0.11	0.18	0.17	0.05
SIX	(-0.83)	-0.23	0.05	0.33	-0.19	-0.06
CVLI	-0.21	-0.23	0.13	(0.88)	-0.11	-0.00
CVLC	-0.20	-0.22	0.13	(0.85)	-0.09	-0.02
P2	(-0.50)	-0.11	0.32	0.17	(-0.71)	-0.18
P3	(0.85)	0.33	0.11	-0.05	0.17	0.16
P4	0.37	0.20	-0.11	-0.10	0.22	(0.60)
P5	0.05	-0.21	-0.24	0.14	0.22	-0.18
MI	(0.64)	0.28	0.07	0.07	0.11	0.15
MX	0.47	-0.03	0.21	-0.01	-0.04	-0.04
X1	-0.17	-0.00	(0.84)	0.26	-0.30	-0.16
X3	-0.12	-0.02	(0.85)	0.27	-0.38	-0.16
X4	-0.09	0.14	(-0.60)	0.25	(0.58)	0.12
X2	(-0.62)	-0.18	0.43	0.37	-0.38	-0.25
X8	0.41	0.37	-0.26	-0.46	(0.61)	0.14
X9	0.29	0.42	-0.19	-0.45	(0.64)	0.09
X5	(-0.68)	-0.29	(-0.61)	0.18	-0.16	-0.09
X6	-0.12	-0.34	0.32	0.04	-0.11	(-0.83)
X7	(-0.78)	-0.30	-0.41	0.26	-0.21	-0.14
Z1	0.16	0.04	(0.93)	0.23	-0.04	-0.12
Z3	0.18	0.02	(0.94)	0.25	-0.07	-0.12
Z4	-0.25	0.03	(-0.83)	-0.29	0.20	0.12
Z2	-0.02	-0.19	(0.59)	(0.72)	-0.18	-0.14
Z8	-0.04	0.29	(-0.50)	(-0.74)	0.23	0.11
Z9	-0.05	0.33	-0.46	(-0.73)	0.23	0.10
Z5	-0.30	-0.22	(-0.86)	0.31	-0.07	0.05
Z6	-0.25	-0.32	0.28	0.04	-0.13	(-0.84)
Z7	-0.34	-0.30	(-0.60)	(0.60)	-0.10	-0.00
Q1	0.29	0.15	-0.24	-0.15	(0.88)	0.10
Q2	(0.61)	0.10	-0.03	-0.10	(0.69)	0.12
Q3	(0.89)	0.20	0.12	0.12	0.24	0.20
Q4	(0.67)	0.27	-0.09	-0.13	(0.64)	0.15
Q5	(0.59)	0.28	-0.06	-0.12	(0.67)	0.11

Table 7. Grouping of variables based on factor loadings greater than + or - 0.50.

Factor #	1	2	3	4	5	6
Eigenvalue	19.1	9.3	4.2	2.7	1.5	0.9
% of Variance	50.3	24.5	11.3	7.2	4.1	2.6
Cummulitive %	50.3	74.8	86.1	93.3	97.4	100.0
1	P1	AH	Z3	CVLI	Q1	Z6
2	Q3	AE	Z1	CVLC	P2	X6
3	VLI	AV	Z5	Z8	Q2	P4
4	VLC	PE	X1	Z9	Q5	
5	P3	PV	X3	Z2	Q4	
6	SIX	PH	Z4	Z7	X9	
7	X7	LA	X5		X4	
8	Q4	WA	X7		X8	
9	X5		X4			
10	MI		Z2			
11	X2		Z8			
12	Q2					
13	Q5					
14	P2					

Table 8. Descriptive statistics for factor 1 indexes based on the combined sample.

	MIN.	MAX.	MEAN	STD. DEV.	SKEW	COEF. OF VARIATION
P1	-177.90	427.72	61.28	117.80	0.68	192.20
Q3	0.01	0.52	0.14	0.10	0.83	72.14
VLI	334.53	999.99	774.96	142.71	-0.76	19.16
VLC	5	48	32	9	-0.71	28.66
P3	-0.08	88.19	15.45	15.07	1.58	97.54
SIX	3	30	20	5.5	-0.61	27.50
X7	0.28	1.24	0.84	0.15	0.88	17.86
Q4	0.01	2.73	0.40	0.35	1.67	87.50
X5	0.24	1.10	0.72	0.14	-0.62	19.44
X2	0.04	0.90	0.41	0.20	0.28	48.78
MI	1.00	2.40	1.17	0.17	2.21	14.53

Correlations between indexes are determined using a rank correlation method because of unfamiliarity with the shape indexes and the wide range of skew values for their distributions. The rank order correlation coefficients for the Dixonville sample of delineations and factor 1 indexes are shown in Table 9. Correlations with SIX were not available, but all other correlations are quite high. One index from this group will adequately represent the factor 1 indexes as long as it is the most interpretable.

Table 9. Rank order correlation coefficients for factor 1 indexes using the Dixonville series sample.

	P1	Q3	VLI	VLC	P3	SIX	X7	Q4	X5	X2
P1	---									
Q3	.93	---								
VLI	-.88	-.89	---							
VLC	-.89	-.89	.99	---						
P3	.90	.98	-.92	-.93	---					
SIX	---	---	---	---	---	---				
X7	-.86	-.92	.89	.89	-.93	---	---			
Q4	.85	.93	-.93	-.93	.95	---	-.87	---		
X5	-.79	-.81	.76	.75	.58	---	.95	-.70	---	
X2	-.61	-.78	.89	.89	-.82	---	.70	-.92	.48	---
MI	.84	.89	-.83	-.83	.90	---	-.86	.85	-.77	-.70

Four indexes derived from the vertex lag method of shape measurement have high loadings on factor 1. The vertex lag index (VLI) and classification (VLC) have identical loadings of 0.87. The numerical difference between VLI and the vertex lag index for a delineation's convex hull (CVLI) is the index labeled P1. P1 has the highest loading on factor 1 of any index. The shape classification index (SIX), which combines VLC with a measure of elongation, has the sixth highest loading on factor 1.

VLI values for the sampled delineations range between 334.5 and 1000.0. In order for a delineation to have a value of 1000, it has to have a perfectly smooth circular shape. Indentations and general boundary irregularity increase as VLI decreases. Theoretically, a set of lag sums and sums of squares from the vertex lag method is unique to only one shape. However, the delineations in Figure 8 have the same VLI values and appear to have different shapes.

Reducing the 20 lag distance values to just one value may mask the true shape descriptive properties of the vertex lag method. On the other hand, B-27 and BD-16 (Figure 8) may have nearly the same shape, and it is the visual perception of shape that is not accurate. According to topological theory, shape is independent of transformations such as stretching or bending (Duda and Hart, 1973).

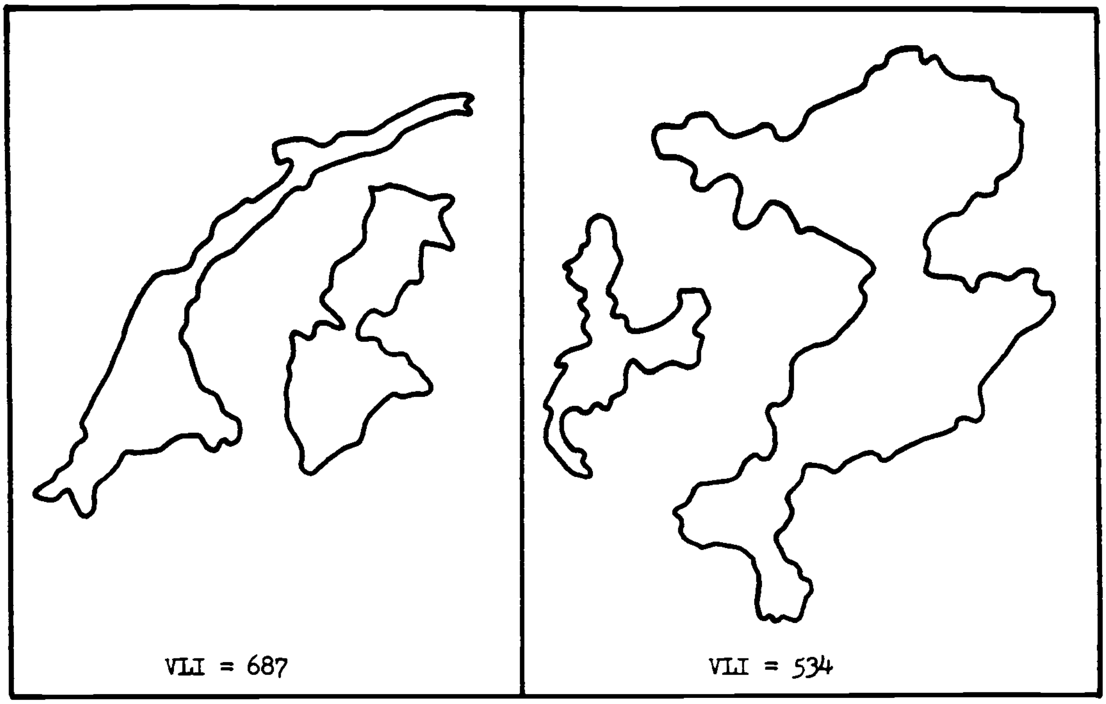


Figure 8. Pairs of delineations with equal VLI index values.

Thus, if VLI represents shape, perhaps the difference in appearance is due to some degree of stretching. This idea was tested by using a zoom transfer scope to compress the larger shape for each pair in Figure 8 along the long axis until both members of a pair had the same long axis. When this was done, the delineations with identical VLI values had strikingly similar perceived shapes.

The potential range of 0 to 1000 for VLI was divided into 10 equal classes. The delineations for the five soil series mapping units were placed in their appropriate classes. Figure 9 includes delineations from each of the ten shape classes. As the VLI values become higher, the amount of boundary complexity decreases in an interpretable manner.

The coefficient of variation for VLI (Table 8) is low at 19.16. This indicates that the soil map delineations cluster close to the mean of 775.0. The distribution for all the delineations has a slightly negative skew.

The variation in VLC values (CV=28.66) is higher than for VLI. This is due, in part, to the fact that the experimental range of the vertex lag distance values is used to produce VLC. The potential range of VLC is 1 to 50. The minimum value for the entire set of delineations is 5 and the maximum is 48. The mean VLC value is 32. VLI and VLC are highly correlated (r=0.99), as shown in Table 9.

Because VLC values are dependent on the range of values in the data set, any interpretations of this index are limited to this particular set of delineations. VLI values do not have this limitation, and any two dimensional shape will fall somewhere within

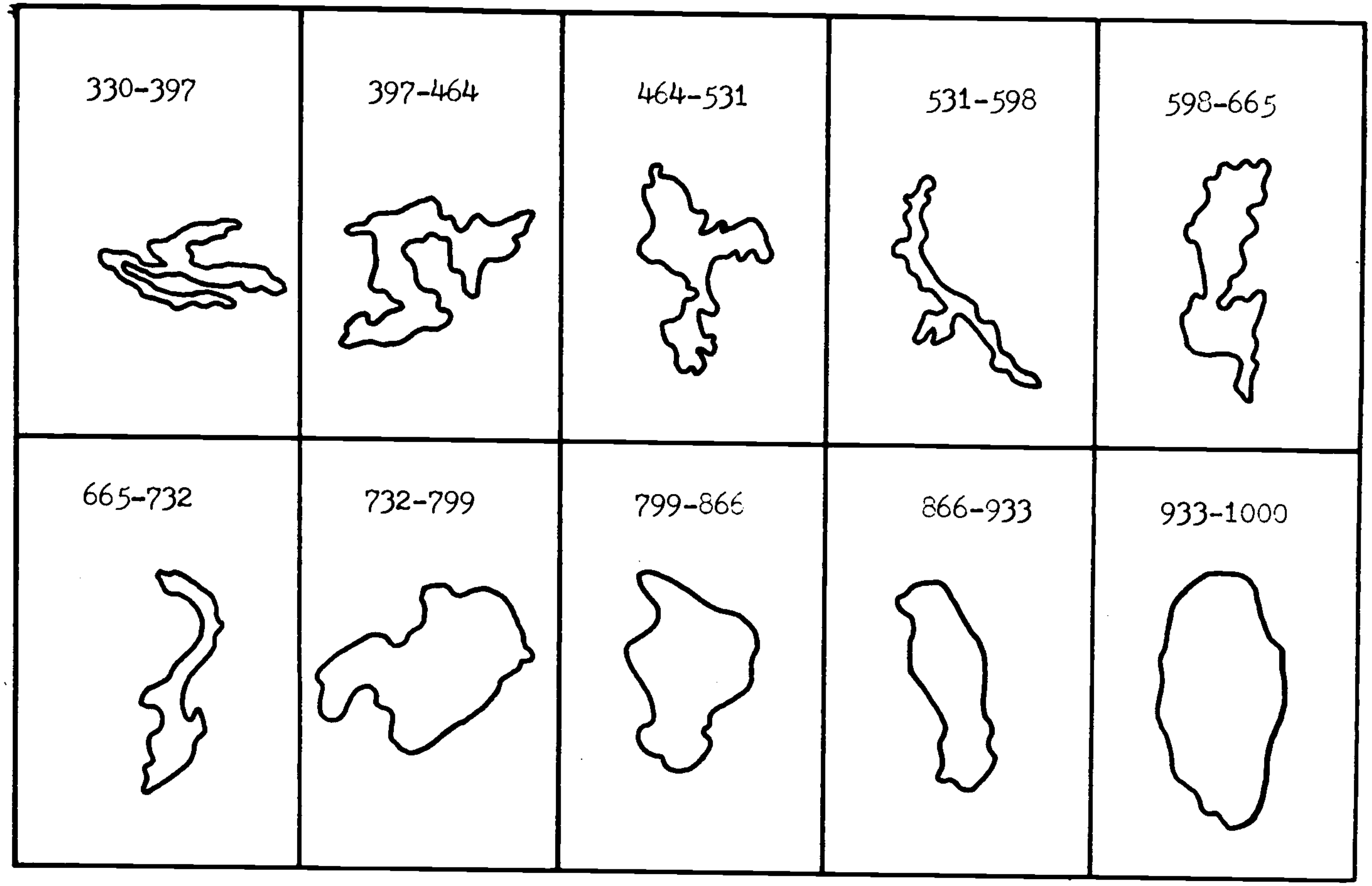


Figure 9. Representative delineations for 10 classes of VLI

the range 0 to 1000. Thus VLI is preferred over VLC as a universal shape index for delineations. VLC could be used for specific data sets in order to separate the data for easier interpretation.

The numerical difference between the vertex lag indexes for a delineation and convex hull is the index P1. P1 had the highest loading on factor 1. The interpretation of P1 data is based on three premises. The first is that VLI is a true measure of shape for any figure. The second is that delineation shape is a function of differing degrees of both boundary complexity and form or elongation. The third premise is that the convex hull of a delineation represents only the form for a delineation. Therefore, subtracting VLI from CVLI produces an index of the amount of boundary complexity that influences delineation shape. Expressed in another way;

Shape = Complexity + Form
Complexity = Shape - Form.

Since the convex hull is more regular than a delineation, CVLI was believed to always be larger than VLI. P1 would then be a positive number between 0 and 1000 as a potential range. The minimum experimental value for P1 in Table 8 is -177.9. The coefficient of variation of 192.2 is the highest of any factor 1 index. The mean value is 61.28 and the distribution had a slight positive skew. If the assumptions on which P1 is based are true, a negative value indicates that the delineation shape is dominated by form with little influence from complexity.

Delineations with negative P1 values were further analyzed. The convex hull polygons were checked and found to fit each delineation properly. The plot of Pl vs. area in Figure 10 shows that most delineations with negative Pl values are quite small. Perhaps the vertex lag programming for the distance values has too much error to account for subtle boundary differences of small delineations.

Tracings of larger delineations with negative Pl values are also included in Figure 10. The majority of these delineations are elongated and narrow with indistinct indentations. This may support the complexity plus elongation equation or may be caused by inaccurate vertex lag approximations. The origin points and edge lengths are not the same for a delineation and convex hull approximation. Narrow extremities may or may not have been accurately approximated by the vertex lag polygon with 22 sides. This problem with polygonal approximation could be solved by using a greater number of edges in the vertex lag method. The theoretical questions that the Pl results raise are intriguing. However, Pl was not selected for further analysis of delineation shape because of the uncertainty involved with the index interpretation.

SIX is another index designed to incorporate elements of form and complexity into a single index of shape. VLC is a measure of complexity. Schumm's elongation ratio for a delineation's convex hull (Z3) is a measure of form. The range of VLC values (1-50) was divided into ten classes and the potential range of Z3 (0.0-1.0) was divided into three classes. This combination forms a matrix of 30 cells. Each delineation was assigned a classification within the range of 1 to 30.

Delineations are distributed well along the entire range of VLC.

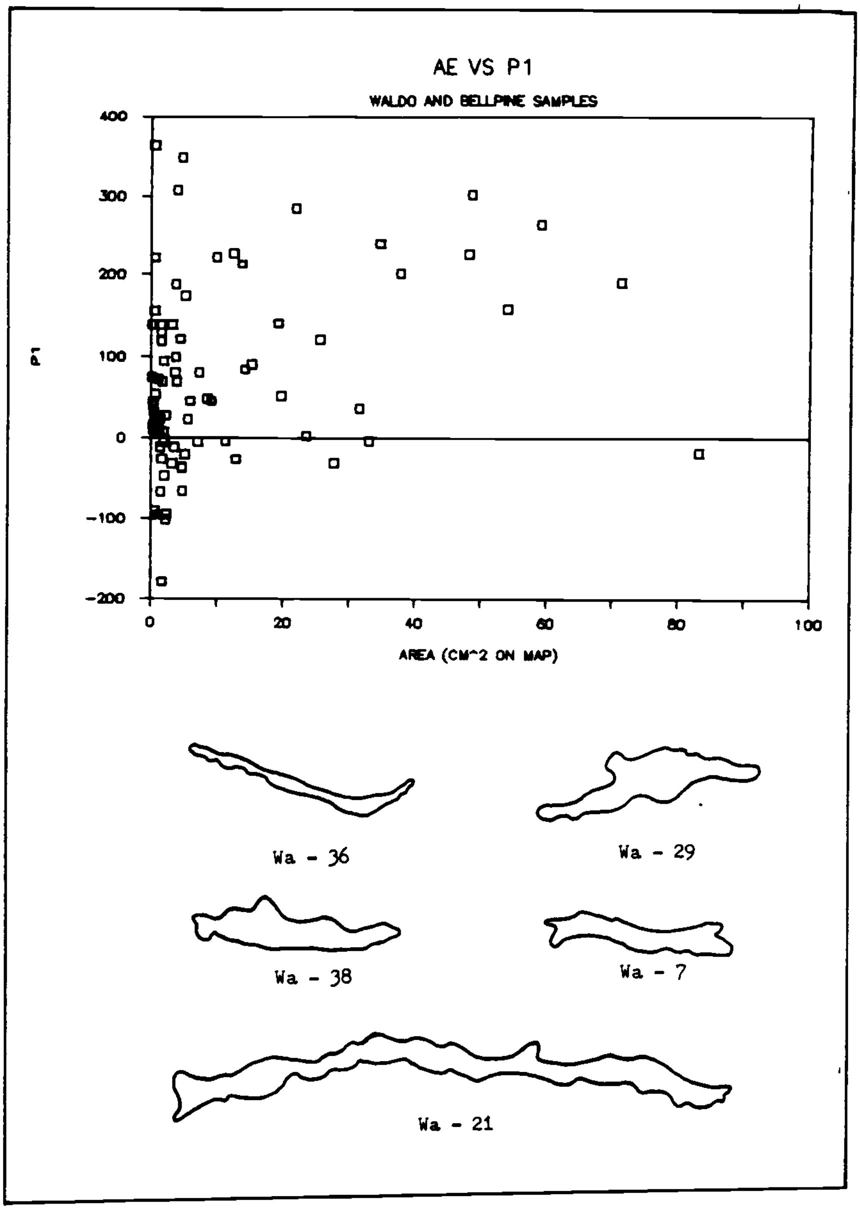


Figure 10. Plot of AE vs P1 and 5 delineations with negative P1 values.

However, the minimum value for Z3 of 0.33 excludes most delineations from being classified in the lowest class of elongation. The results of SIX are not entirely useful because delineations are nearly excluded from having values of 1, 4, 7, etc. The experimental range of Z3 should be divided into 3 classes instead of the potential range. However, the concept of combining a complexity index with an elongation index should provide a useful shape classification for soil map delineations.

Three figure attribute ratios with high loadings are also considered to be defining variables for factor 1. These are the elliptical irregularity index (X7), minimum perimeter index (X5), and Miller's circularity index (X2). The factor 1 loadings for these indexes are -0.78, 0.68, and -0.62, respectively. These three indexes also have fairly high loadings on factor 3 (Table 6).

Of these three attribute ratios, X2 has the most normal distribution of values with a skewness value of only 0.28. Delineations have values for X2 along the entire continuum of potential values beween 0 and 1.0. The mean for X2 was 0.41.

X2 is the ratio of delineation area to the area of a circle with a perimeter equal to delineation perimeter. As boundary complexity increases, the value of X2 decreases. However, the relationships between boundary points on a complex boundary are not preserved. This is why figure attribute ratios can be called shape non-preserving. The same general index for complexity could be assigned to delineations with very different shapes.

A delineation with a perfectly smooth, elliptical boundary would have more boundary than a corresponding circular standard shape. The value of X2 for such a delineation would be less than 1.0. Thus, X2 incorporates elements of form as well as complexity. X2 was not considered for further use because it does not have a very high factor loading and several different shapes can have the same index value.

X7 includes delineation area, perimeter and, long axis values in the index equation. The ellipse is used as a standard shape rather than a circle. A circle is a form of an ellipse, but the equation for a circle uses only two primary measurements at a time. The elliptical standard shape has the same long axis and area as the delineation. X7 is the ratio of delineation perimeter to the perimeter of the elliptical standard. The delineation should always have a larger perimeter than the ellipse. The potential range for X7 is 0 to 1.0.

As the boundary of a delineation becomes more irregular, the value of X7 decreases. The minimum value for the set of delineations is 0.28 and maximum is 1.24. Rounding error and the use of manually measured long axis values causes X7 values for a few delineations to be larger than 1.0. X7 has a fairly high mean value of 0.87, and the coefficient of variation of 17.86 is similar to that of VLI. X7 is fairly difficult to interpret, and like all figure attribute ratios, it does not preserve the spatial relationships between boundary points.

An extremely narrow rod was used as the standard shape for calculating X5. The narrow rod has a perimeter equal to twice the delineation long axis. The ratio of this perimeter to the perimeter of the delineation yields index values for X5 between 0 and 1.0. Again, small inaccuracies in the long axis values compared to the high resolution perimeter values causes an occasional delineation to have a value greater than 1.0.

A circle has an X5 value of 0.63. As a circle is flattened, the index approaches 1.0. Delineations with perimeter values greater than that of circle would have extra perimeter in the form of boundary indentations. X5 values would then be less than 0.63. A skewness value of 1.62 and a mean of 0.72 shows that most of the delineations have roughly circular shapes. X5 incorporates elements of both form and complexity, based on the very general interpretability and the dual loading on factors 1 and 3.

In order to learn more about figure attribute ratios, five additional indexes were constructed. These indexes compare attribute ratio results for a delineation and its convex hull. Q1-Q5 are called attribute ratio comparisons. Four of these comparison indexes have loadings on factor 1 of greater than + 0.50. Three of these also have high loadings on factor 3.

Q3 has the second highest loading on factor 1. Q3 compares X7 for a delineation and its convex hull (Z7) in the form: Q3 = Z7 - X7. According to the discussion of X7, Z7 would be larger than X7. Q3 has a range between 0 and 1.0. Values for delineations tend to be low, with a minimum of 0.01, maximum of 0.52, and a mean of only 0.14. Large numbers for Q3 indicate extreme boundary indentations, i.e. a high overall convex deficiency. According to this index, the sampled delineations do not have highly irregular boundaries. Although Q3 has a high factor 1 loading, it only gives a general picture of complexity and, like X7, is not easily interpreted.

Q4 utilizes Fridland's coefficient of dissection (X8 and Z8) in the form of the ratio of the difference between X8 and Z8, to Z8. Q4 has a factor 1 loading greater than X5 or X2. X8 and Z8 values increase from 1.0 as boundary complexity increases. For a circle, an undissected figure, Q4 would have a value of 1.0. A complex figure would have a value above greater than 1.0, since X8 would be larger than Z8.

The loading on factor 3 for Q4 is nearly as high as the factor 1 loading. Q4 has a mean of 0.40 and a skew of 1.67, which again indicates that delineations do not commonly have a lot of boundary complexity. The interpretation of Q4 refers to complexity only. However, the shape of the complexity influencing the index is unknown.

Three convex deficiency indexes have high loadings on factor 1. The area deficiency index (P2) has a loading of -0.50 on factor 1, and since it has a higher loading on factor 4, it is evaluated with the other factor 4 indexes. The perimeter deficiency index (P3) and the concavity complexity index (MI) have factor 1 loadings of 0.85 and 0.64, respectively.

Values for MI range from 1.0 to an unknown upper limit. A value of 1.0 indicates that a delineation has perfectly smooth triangular indentations. MI has a fairly high loading (0.85), but generated index values for the delineations are not easily interpreted.

P3 is strictly a complexity index. This index is the perimeter of a delineation's concavities expressed as a percent of the delineation's convex hull perimeter. The extra boundary needed for concavities is an expression of complexity. P3 increases as the value for perimeter of concavities (PV) increases. A delineation with no concavities, whether a circle, ellipse, or any convex figure has a P3 value of 0.

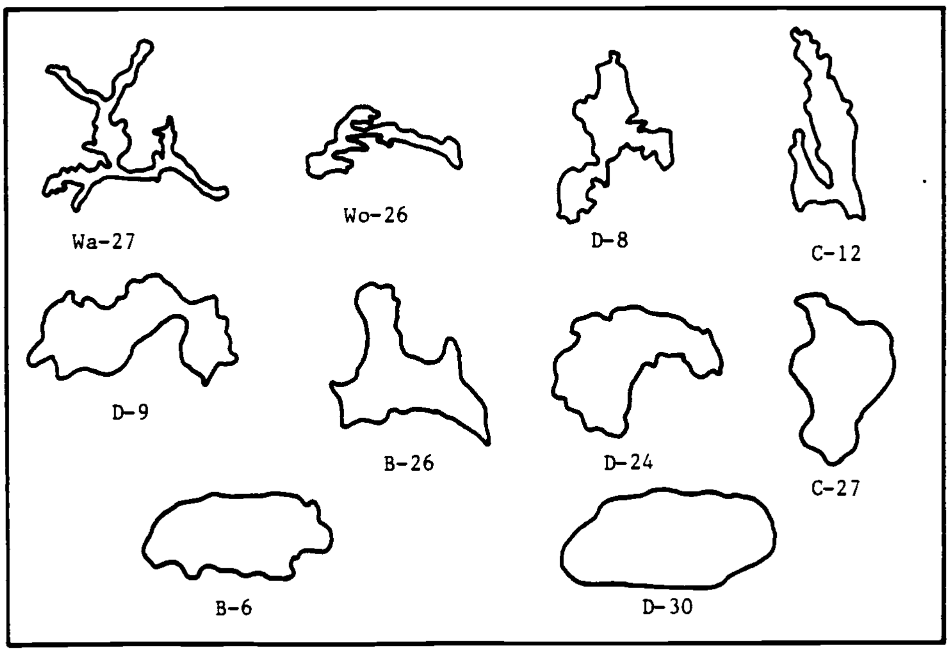
P3 loads higher on factor 1 than any other attribute ratio. The mean value for P3 is 15.5%, and the distribution has a positive skew. These results suggest that delineations tend towards convexity rather than extreme complexity. However, the maximum value of 88.2% shows that delineations can have very complex boundaries. Concavities with the same perimeter can have very different shapes, so P3 is not considered a direct shape measurement. P3 is an excellent measure of overall boundary complexity for soil map delineations, and the index is easy to calculate. However, VLI is preferred over P3 because it is a shape preserving method, and it has a higher loading on factor 1.

Most of the indexes with significant loadings on factor 1 incorporate both form and complexity into their values. Some indexes such as P3 are strictly complexity measurements. Complexity may be the major attribute of shape that delineations possess, which causes the inclusion of these indexes in factor 1. All of the factor 1 indexes have high intercorrelations. The vertex lag index was selected to represent factor 1 indexes for the analysis and comparison of the soil mapping unit samples. VLI has a very high loading, is interpretable, is easy to calculate with the aid of a computer, and is a direct measure of shape.

One problem with the index evaluation is obtaining a visual perspective of which shapes correspond to the various index values. The potential range of VLI values was divided into ten classes. The delineations for the five series map unit samples were placed into their appropriate class. The median delineation for each class was traced. These tracings and their VLI values are included in Figure 11. The tracings are individually scaled down to have approximately the same length long axis. This helped to remove the influence of size from confusing the perception of shape. VLI, compared to P3 and X7, does the best job of placing similar shapes close together and different shapes far apart along the index range and it is a shape preserving method.

Factor 2 Indexes -- All of the indexes that loaded heavily on factor 2 are primary measurements. Five of the indexes are convex hull primary measurements and three of them refer to delineations. The descriptive statistics in Table 10 and intercorrelations shown in Table 11 are used to evaluate the usefulness of the results for each index. Factor 2 represents indexes that account for 24.5% of the variance in the data set (Table 7).

Convex hull area (AH) has a loading of 0.95 on factor 2, which is higher than any other index. AH is useful for creating convex deficiency indexes by allowing the convex hull to be a standard shape rather than a circle or ellipse. The area of a delineation's hull could represent the sphere of influence of the soil and landscape



	WA-27	WO-26	D-8	C-12	D-9	B-26	D-24	C-27	B-6	D-30
VLI -	357	438	501	564	637	701	760	839	905	952
P3 -	67.5	37.1	47.6	12.6	27.1	17.3	14.2	5.1	7.8	1.6
Z7 -	.42	.56	.56	.72	.75	.77	.80	.81	.93	.98

Figure 11. Placement of delineations along the VLI continuum and comparison of three complexity indexes.

TABLE 10. Descriptive statistics for factor 2 indexes based on the combined sample.

	MIN.	MAX.	MEAN	STD. DEV.	SKEW	COEF. OF VARIATION
AH **	0.68	1855.84	57.72	128.72	7.89	221.48
AE **	0.60	1086.32	37.20	81.40	7.33	218.82
PE *	314.00	53694.00	3360.00	4066.00	5.61	121.10
AV **	0.04	769.52	20.92	48.68	7.97	251.82
PV *	108.00	49934.00	3042.00	3886.00	5.33	127.74
PH *	310.00	33878.00	2706.00	2756.00	4.65	101.85
LA *	122.00	6368.00	1044.00	872.00	2.21	83.52
WA **	0.01	108.24	7.72	14.60	3.34	189.12

** Hectares
* Meters

properties represented by a delineation. No immediate use for this information was known.

All of the factor 2 indexes show extremely high correlations with one another. The most interpretable index can be chosen to represent this factor. Because AH by itself is not as interpretable as delineation area (AE), the latter is considered the best index to represent factor 2.

Table 11. Rank order correlation coefficients for factor 2 indexes using the Dixonville series sample.

	AH	AE	PE	AV	PV	PH	LA	WA
AH	---							
AE	.99	---						
PE	.98	.95	---					
AV	.96	.92	.98	---				
PV	.98	.95	.99	.99	---			
PH	.99	.97	.99	.97	.98	---		
LA	.98	.96	.98	.96	.97	.99	---	
WA	.93	.88	.95	.98	.96	.93	.92	---

The values for delineation area range from 0.60 ha to 1086.4 ha.

The mean area of 37.2, the high coefficient of variation (218.8), and very high positive skew of 7.33 all point to an asymmetric distribution of delineation sizes, most of which are small. The precise distribution of these values for each soil map unit in a survey could be informative. A system for classifying soil body area is available and has been previously used by Hole (1978), and Fridland (1972).

The long axis for a delineation (LA), perimeter (PE), and perimeter of the convex hull (PH) have no real advantage over AE for further use by themselves. The concavity information such as area and perimeter of concavities can be used to make indexes, but again, by themselves they provide little interpretable shape information. The weighted average concavity area (WA) is used to understand if a delineation is dominated by one or many concavities. Again, the information provided by WA does not improve on AE for an index to represent Factor 2.

Factor 3 indexes -- All of the indexes with high loadings on factor 3 are figure attribute ratios. Most of the rank order correlation coefficients among these indexes (Table 12) are quite high. The three indexes having the highest loadings on factor 3 are figure attribute ratios for the convex hull. Factor 3 accounts for over 10% of the variance in the factor analysis.

The index with the highest loading on factor 3 is Schumm's elongation ratio for a delineation's convex hull (Z3). This index compares the long axis of a convex hull to the diameter of a circle with the same area. The convex hull represents the general form of a delineation.

Table 12. Rank order correlation coefficients for factor 3 indexes using the Dixonville series sample.

	Z3	Z1	Z5	X1	X3	Z4
Z3	---					
Z1	1.0	---				
Z5	-0.88	-0.88	---			
X1	.92	.92	-0.73	---		
X3	.92	.92	-0.73	1.0	---	
Z4	-1.0	-1.0	.88	-0.92	-0.92	---
Z7	-0.67	-0.67	.91	-0.45	-0.45	.67

The potential range for Z3 is 0 to 1.0. The highest value for any delineation is 0.97, which represents a nearly circular convex hull. The distribution of values is fairly normal with a mean of 0.65, a minimum value of 0.30, and a low coefficient of variation of 18.46 (Table 13). As an index for measuring the form of a delineation, Z3 is interpretable even though it is a figure attribute ratio. This index, coupled with an index of complexity, could provide information about two of the sub-properties of the shapes of soil map delineations with a single index.

Horton's form ratio for a convex hull (Z1) compares the convex hull area (AH) to the long axis squared. The potential range is 0 to 1.0. Because Z1 and Z3 are functionally related, they are perfectly correlated, and each provides essentially the same information. The only difference is that a low value for Z1 indicates a compact, circular shape, and elongation increases as Z1 values increase. Both indexes show that the convex hulls of delineations tend to be on the circular side of either index continuum.

The minimum perimeter index for a convex hull (Z5) has a loading of -0.86 on factor 3. This index has a potential range of 0 to 1.0. The maximum value calculated for Z5, however, is 1.14. (Table 13). Manual measurement and input of long axis values produces small errors in the index calculations. This problem of high X5 and Z5 values is not common in the data set.

TABLE 13. Descriptive statistics for factor 3 indexes based on the combined sample.

	MIN.	MAX.	MEAN	STD. DEV.	SKEW	COEF. OF VARIATION
Z3	0.30	0.97	0.65	0.12	-0.17	18.46
Z1	0.07	0.74	0.35	0.13	0.24	37.14
Z5	0.38	1.14	0.81	0.10	-0.88	12.35
X1	0.02	0.64	0.26	0.12	0.38	46.15
X3	0.15	0.91	0.56	0.14	-0.17	25.00
Z4	1.06	11.05	2.65	1.25	2.25	47.17
Z7	0.48	1.30	0.99	0.08	-2.18	8.08
P5	9.33	96.61	43.77	17.45	0.81	39.87

Z5 is the ratio of two times the long axis to the convex hull perimeter (PH). For a convex figure, PH could only increase from 2 times the long axis value to a value slightly higher than the perimeter of a circle with the same long axis. The value of Z5 for a circle is 0.63. A square is a convex figure with a little more perimeter than a circle. The Z5 value for a square is less than 0.63.

The mean value for Z5 is 0.81, and the skew of the distribution is toward higher values. These results indicate, as do those for Z3 and Z1, that convex hulls for most delineations have a greater ellipticity than a circle. This means that the ratio of minor axis to major axis is lower for delineations than for a circle or square.

Horton's form ratio (X1) and Schumm's elongation ratio (X3) for delineations have loadings on factor 3 of 0.84 and 0.85, respectively. These indexes are functionally related, so they have a rank correlation of 1.0 (Table 12). X1 or X3 might be useful if elements of both form and complexity were needed in a single figure attribute ratio. However, there is no benefit to using these indexes over Z1 or Z3 for representing delineation form.

PE is greater than PH and AE is less than AH because of boundary complexity. Therefore, X1 is less than Z1 and X3 is less than Z3. Both a delineation and convex hull have the same elongation. The best measure of elongation would be one that is the least affected by boundary complexity.

The minimum perimeter index for a delineation (X5) has a factor 3 loading of -0.61. This index has a higher loading on factor 1 and was discussed previously. X5 does incorporate elements of both complexity and elongation, but figure attribute ratios such as this are not easily interpreted in terms of shape.

Stoddart's ellipticity index for a delineation (X4) and convex hull (Z4) have significant loadings on factor 3. Both indexes compare the long axis of a figure to the minor axis of an ellipse with the same area and long axis as the figure. An index value of 1.0 is the minimum possible, which represents a circle. Values increase from 1.0 as elongation of the figure increases.

The mean value for X4 in Table 13 is 1.22, whereas the mean value for Z4 is lower at 1.06. X4 has a much higher maximum than Z4 does. In general, Stoddart's ellipticity index is influenced by the complexity of a figures boundary to some unknown extent. Since the convex hull has little boundary complexity, Z4 would be better for measuring the form of a delineation. Neither Z4 or X4, however, offer any advantage over Z3 for measuring delineation form.

Z2 has a loading of 0.59 on factor 3. This index has a higher loading on factor 4 and is evaluated with the factor 4 indexes.

In summary, Schumm's elongation ratio for a delineation's convex hull is selected as a measure of form representing the factor 3 indexes. It has the highest loading on factor 3, the highest correlations with other factor 3 indexes, and is fairly easy to interpret for a convex hull without the influence of boundary complexity on form measurement.

Factors 4, 5, and 6 Indexes -- Factors 4, 5 and, 6 together account for only 13.9% of the total variance. These latter factors are dominated by figure attribute ratios of various kinds, as is factor 3. Because of the low percent of variance accounted for by these factors, no index was selected from them. Each index with high loadings on these factors is briefly evaluated using the descriptive statistics in Table 14 and the correlation coefficients in Table 15.

The dominant characteristic of the indexes in factor 4 is that they measure shape properties of the convex hull. The convex vertex lag index (CVLI) and classification (CVLC) have the highest loadings. Attribute ratios for this factor include Miller's circularity index (Z2), Fridland's coefficient of dissection (Z8) and, Boundary complexity index (Z9). The elliptical irregularity index using convex hull attributes has a loading of -0.60 on factor 4 and a loading of 0.60 on factor 5. All of these indexes are influenced by

Table 14. Descriptive statistics for indexes with high loadings on factors 4, 5, and 6 based on the combined sample.

	MIN.	MAX.	MEAN	STD. DEV.	SKEW	COEF. OF VARIATION
FACTOR 4						
CVLI	400.70	966.94	806.24	77.34	-2.12	59.21
CVLC	9	46	34	4.65	-1.78	13.68
Z8	1.03	2.22	1.24	0.19	1.84	15.32
Z2	0.20	0.94	0.68	0.16	-0.63	23.53
Z9	1.06	4.92	1.59	0.55	2.48	34.59
FACTOR 5						
Q1	0.01	3.95	0.47	0.51	2.91	108.51
P2	20.21	98.87	72.98	16.41	-0.76	22.49
Q2	0.01	0.43	0.10	0.06	1.25	60.00
Q5	0.02	12.88	1.08	1.20	3.30	111.11
X9	1.12	26.18	3.52	3.12	3.60	88.64
X4	1.22	43.81	4.11	3.44	5.18	83.70
X8	1.06	5.12	1.77	0.63	2.04	35.59
FACTOR 6						
Z6	0.03	1.07	0.29	0.19	1.24	65.52
X6	0.03	1.19	0.32	0.20	1.31	62.50
P4	34.67	99.42	84.70	11.18	-1.45	13.20

the complexity of a figure's convex boundary. Since a convex hull has little if any boundary complexity, Z2, Z8, and Z9 do not provide much usefull information.

CVLI and CVLC are identical to VLI and VLC with respect to their shape preserving qualities and as measurements of boundary complexity. The convex hull is useful for comparisons with delineations in order to measure convex deficiency attributes. The convex hull is also useful for isolating elongation as a sub-property of shape using Z3. However, measuring the shape complexity

Table 15. Rank order correlation coefficients for indexes with high loadings on factors 4, 5, and 6 using the Dixonville series sample.

	CVLI	CVLC	Z8	Z2	Z9	Q1	P2	Q2	Q5	X9	X4	X8	Z6	Z6
CVLC	.99	--												
Z8	-.85	-.81	--											
Z2	.85	.81	-1.0	--										
Z9	-.85	-.81	1.0	-1.0	--									
Q1	-.44	-.40	-.22	-.52	.52	--								
P2	.44	.40	-.52	.52	.52	-1.0	--							
Q2	-.30	-.26	.36	-.36	.36	.97	.97	--						
Q5	-.36	-.32	.45	-.45	.45	.97	-.97	.96	--					
X9	-.60	-.56	.75	-.75	.75	.94	-.94	.86	.92	--				
X4	-.65	-.62	.84	-.84	.84	.62	-.62	.47	.54	.77	--			
X8	-.60	-.56	.75	-.75	.75	.94	-.94	.86	.92	1.0	.77	--		
Z6	.34	.30	-.46	.46	-.46	-.71	.71	-.65	-.76	-.73	-.56	-.73	--	
X6	.34	.30	-.45	.45	-.45	-.65	.65	-.58	-.69	-.68	-.56	-.68	.99	--
P4	-.27	-.23	.25	-.25	.25	.73	-.73	.72	.76	.65	.36	.65	-.78	-.75

of a delineation's convex hull for use alone has very little interpretive value. Further study of the differences between VLI and CVLI may provide a better understanding of the interaction between complexity and elongation of a map unit delineation.

Z2, Z8, and Z9 are all functionally related. They are identical with respect to the properties they measure, and they all rank shapes identically. None are particularly interpretable, so there is no advantage in retaining factor 4 in any further analysis.

The index loading highest on factor 5 is the comparison of Stoddart's ellipticity index between a delineation (X4) and it's convex hull (Z4). This index is labeled Q1, and it represents the difference between X4 and Z4 expressed as a percent of Z4. Z4 represents the form of a delineation's convex hull without the influence of boundary complexity. Q1 measures how much boundary complexity influences the elongation value of X4.

The potential range of Ql is a positive number between 0.01 and infinity. With a mean of 0.47, and a strong positive skew toward lower values, the majority of X4 values are not affected much by delineation complexity. However, the maximum Ql value of 3.95 indicates that X4 can be misleading if used to measure only elongation of a delineation. Ql, as a complexity index has no particular advantage over VLI, and X4 is a typical attribute ratio which is difficult to interpret in shape terms.

The area deficiency index (P2) has the second highest loading on factor 5. This index compares the area of a delineation (AE) to the convex hull area (AH). The index shows the percent of the convex hull area actually filled by the delineation. Values range from 20.2% to 98.9%. The mean value for P2 is 72.98%. P2 is intriguing because the area deficiency for a delineation might indicate the degree of erosion (indentations) or soil creep (lobes) that is actively shaping the soil body. However, P2 is not further considered because several factor 1 indexes measure essentially the same properties.

The numerical difference between Z3 and X3 is the index Q2. This index shows a range of 0.01 to 0.43 for the delineations. A low value is assigned to a delineation if boundary complexity does not have much influence on the value for form. The results of this index are similar to those for Ql. As a whole, delineations do not show high boundary complexity. The mean value for Q2 is only 0.10.

Q5 compares the results for the Boundary complexity index for a delineation (X9) and convex hull (Z9). X9 or Z9 is the ratio of a figure's perimeter squared to the area of a circle with the same perimeter. The perimeter of a convex hull is less than the perimeter of a delineation, so X9 will be greater than Z9. Q5 essentially measures the contributions of complexity and form to X9 values.

A low value for Q5 is assigned to a figure with little boundary complexity. The distribution of values shows a mean of 1.08 and a positive skew of 3.3. Delineations have regular boundaries for the most part with an occasional delineation that is very complex. This same information is provided by several indexes that loaded on other factors.

The figure attribute ratios used to construct Q1, Q2 and, Q5 also have high loadings on factor 5. The Boundary complexity index (X9), Stoddart's ellipticity index (X4), and Fridland's coefficient of dissection (X8) have factor 5 loadings of 0.64, 0.58 and, 0.61, respectively. All three indexes produce values that are influenced by both form and complexity. X4 was previously used as a measure of elongation, but the index is also influenced by complexity. Only very general interpretations are provided by these figure attribute ratios. All three indexes could produce identical values for differently shaped delineations.

Factor 6 appears to include measures of boundary complexity. The concavity perimeter ratio (P4) and the cartographic complexity index for a delineation (X6) and convex hull (Z6) define this factor. Perimeter is the key attribute used to make the complexity ratings. Factor 6 is a weak factor accounting for only 2.6% of the system's variance.

The cartographic complexity index (X6) is an attempt to set a cartographic limit for the amount of perimeter that a delineation could have. The amount of complex perimeter is compared to the actual delineation perimeter. The same index for the convex hull is labeled Z6. Values for Z6 are expected to be lower since PH is always less than PE.

The distribution of X6 values has a mean of only 0.32. The mean value for Z6 is 0.29. The complex perimeter used as a standard is very large, and it is surprising to see values near 1.0 for the maximums. Again, the delineations show a trend of boundary regularity; only an occasional delineation has a very complex boundary.

The concavity perimeter index (P4) compares the perimeter of concavities (PV) to the delineation perimeter (PE) expressed as percent. According to this index, on the average, 84.7% of a delineation's perimeter is concave. However, most of the complexity indexes have shown that the complexity of those concavities is not necessarily high. P4 tells nothing of the shape of concavities and does not provide a great deal of useful information.

Two indexes do not have high loadings on any factor. These are the concavity depth index (MI) and the concavity dominance index (P5). MI measures the average depth of concavities from the midpoint of the mouth to the point on the concavity boundary furthest away. A concavity could be very shallow with a wide mouth and yield a high MI value. The results for this index are not interpretable.

P5 does not load heavily on any factor but does provide some new information. The purpose of the index is to decide whether a delineation is dominated by one or many concavities. This index provides information about the nature of concavities and not just generalized boundary complexity. P5 is calculated by comparing the weighted average sized concavity for a delineation (WA) to the total area of concavities (AV).

A delineation with many small concavities and two large concavities would have a smaller value for WA than a delineation with only one large concavity. A delineation with one large concavity would have a higher value for P5 than a delineation with many small concavities. The dominance of concavities on the shape of a delineation is the property being measured. A large concavity has more influence on shape than indistinct ones. The shape of each concavity is not known.

For the sampled delineations, P5 has a mean of 43.8%. A low value indicates that a delineation has many indistinct concavities that do not influence the shape very much. High values like the maximum of 96.6% represent a delineation with one major concavity. The skew of the distribution is 0.81 and the coefficient of variation is 39.87.

Summary of Index Selection

Factor analysis was the key to the index selection process. The first three factors represent the three dominant, independent spatial properties of the sampled delineations. The properties are complexity, primary measurements, and form. The indexes that were selected to represent each factor are the vertex lag index (VLI), delineation area (AE), and Schumm's elongation ratio for the convex hull (Z3), respectively.

VLI is a shape preserving method of shape measurement and is preferred over the figure attribute ratios that also measure complexity. Of the attribute ratios, the area deficiency index (P3) is the best measure of complexity. The convex hull of a delineation was found to be a more natural and interpretable standard shape of regularity for delineation shape analysis than the circle or ellipse.

VLI comes the closest to meeting the requirements of a true shape index. That is, it provides a range of values that places similar shapes close together and dissimilar shape far apart. It assigns a unique number to differently shaped approximating polygons and is independent of size, orientation, and pattern. The index values can be restated in shape terms, and this ability will improve with experience. Finally, although VLI is used here to measure complexity, the index is designed to be a pure measure of shape and not just the sub-property of shape.

Delineation area is not a shape index, but it is apparently a significant property of delineations due to it's high factor loading. Area has been studied by soil scientists in the past and size classes are available. The size of a body of soil provides some intuitive information about the soil itself and the surrounding soilscape. The distribution of delineation sizes for a map unit might also be informative to map makers and users.

Z3 is a useful index for measuring the form of a delineation independently of boundary complexity. Z3 has the negative properties of all attribute ratios, but it is the relationship of the convex hull to a delineation that makes the index interpretable. Z3 is particularly useful if combined with a pure complexity index into a matrix such as SIX to combine the two primary shape properties of delineations into a single index.

DATA ANALYSIS

Three indexes from the original list of potential shape indexes were selected to make comparisons between the soil map delineation populations that were sampled. Delineation area (AE) was selected, not as a shape index, but as a representative of the primary measurements. The vertex lag index (VLI) is a measure of the shape complexity derived from the 22-sided approximations made for each delineation. Schumm's elongation ratio for the convex hull (Z3) is a quantitative measure of the general form of each delineation. These three indexes represent the three dominant factors in the factor analysis (Table 3), which together account for 87% of the variance in spatial properties for the entire data set.

Not only do AE, VLI and, Z3 represent the significant properties of delineations, they are all relatively unrelated according to the correlation data shown in Table 16. Thus, the selected indexes are not influenced by any spatial property other than the ones they measure. The coefficient of -0.39 for AE and VLI is the highest. Shape complexity increases slightly as delineation area increases. The coefficients for Z3 vs AE and VLI are both very low.

The relationship between shape complexity and area

Table 16. Linear correlation coefficients for the selected indexes based on combined sample.

	AE	VLI	Z3
AE	--		
VLI	-0.39	--	
Z3	0.15	0.08	--

was further examined by looking at the rank order correlations between these variables for each map unit sample. The results are shown in Table 17. Although the overall correlation between AE and VLI is low (Table 16), the magnitude of the correlation does vary from sample to sample. For Bellpine 20-30% delineations, there is

Table 17. Rank order correlation coefficients for AE and VLI for each map unit sample.

	AE vs. VLI
BELLPINE	-0.28
CHEHALIS	-0.52
DIXONVILLE	-0.62
WALDO	-0.21
WOODBURN	-0.55

BELLPINE SLOPE PHASES		DIXONVILLE SLOPE PHASES	
3-12%	-0.57	3-12%	-0.41
12-20%	-0.45	12-20%	-0.27
20-30%	-0.72	20-30%	-0.15
30-50%	-0.32	30-50%	-0.35

some evidence that complexity consistently increases (low VLI) as delineation size increases. Although the other correlations are lower, it is interesting that in every case the relationship is an inverse one.

Index Distributions Within Series

Each of the five soil series samples were evaluated for within-sample characteristics and variations with respect to area, shape complexity, and form. Frequency histograms and descriptive statistics are used for this evaluation. Included in the discussion are brief descriptions of the major soils that the series mapping units represent. These descriptions come from the soil survey for Benton County Area, Oregon (Knezevich, 1975). It is important to note that delineation shape may be affected as much, or more by inclusions and adjacent soil bodies as they are by the dominant soil series. This information is also included in soil surveys, but is not considered here because of the unwanted relationship to soil pattern in this shape analysis project. The distributions of AE, VLI, and Z3 for the five samples provide a better understanding of the populations of delineations mapped as particular soil series.

Woodburn Series -- The Woodburn series consists of soils that formed in lacustrine sediment and silty alluvium on terraces above the floodplain of the Willamette Valley. The surface layer is a silt loam with a depth of about 40 cm. The subsoil extends to a depth of 1.5 m or more, and has a silt loam or silty clay loam texture. Woodburn soils have a well developed argillic horizon. Runoff is slow to medium, and the hazard of erosion is none to slight. Permeability is slow. Most Woodburn delineations are found on slopes of 0-3%. This soil often has a seasonal high water table.

The Woodburn sample has small delineations. The modal class for AE is 5.0 to 10 ha (Figure 12). However, 18% of the delineations are in the 50 to 100 ha class. This pulls the mean up to 18 ha, whereas the median value is only 8.92 ha (Table 18).

VLI values range from very complex (347) to very regular (1000). The modal class for VLI is 800 to 850, and 60% of the sample has values greater than 800. This regularity is the typical situation,

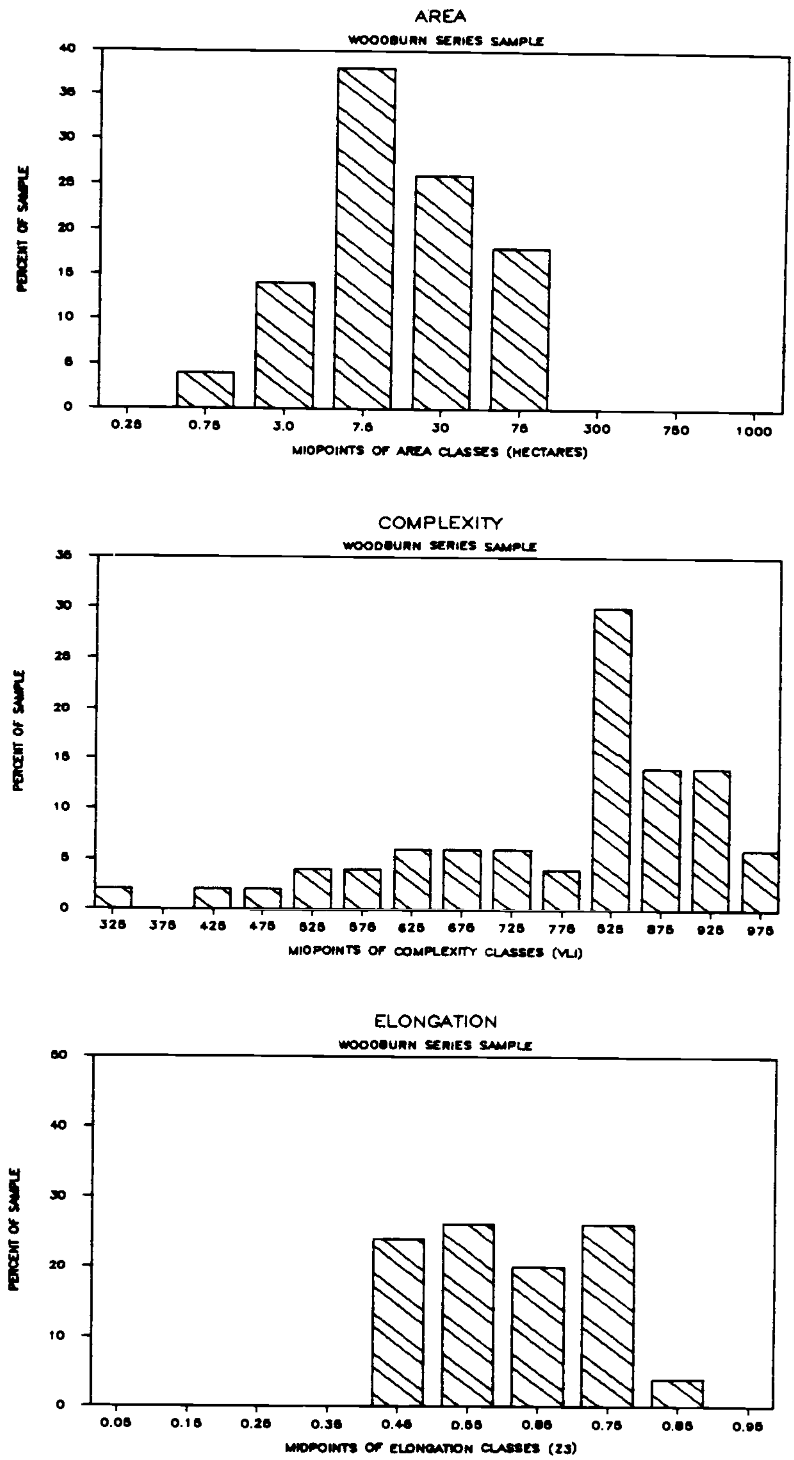


Figure 12. Frequency histograms of AE, VLI, and Z3 for the Woodburn series sample.

although all of the complexity classes lower than 800 include at least one Woodburn delineation except 350 to 400.

Table 18. Descriptive statistics for AE, VLI, and Z3 for the Woodburn series.

	MIN.	MAX.	MEAN	STD.DEV.	COEF. of VARIATION	SKEW	MEDIAN
AE	0.64	83.2	18.0	20.8	114.9	1.63	8.92
VLI	346.8	1000.0	780.9	149.4	19.1	-1.00	825.4
Z3	0.41	0.84	0.61	0.13	20.8	0.08	0.61

Most Woodburn delineations have forms (Z3) that fall anywhere between 0.40 and 0.80 with nearly equal likelihood. The data do not suggest a tendency toward either elliptical or circular forms.

Chehalis Series -- Chehalis delineations occupy areas of alluvial bottom lands. The soils developed in recent mixed alluvium on the floodplain of the Willamette River. This landform is much younger than the foothill areas or valley terraces. In fact, these soils are subject to overflow about once every three to five years. It would seem that the influence of scouring and deposition would have some effect on delineation size and shape.

Chehalis soils are deep and well drained. Runoff is slow, and the hazard of erosion is slight due to slopes being only 0-3%. The surface soil is about 28 cm thick, and the subsoil is about 86 cm thick. Silty clay loam textures dominate to a depth of 1.5 m or more. The surface soil and subsoil are distinguished by a slight structure change in the B horizon, and by color.

The areas of Chehalis delineations vary from extremely small (0.96 ha) to very large (745.2 ha) as shown in Table 19. The

Table 19. Descriptive statistics for AE, VLI, and Z3 for the Chehalis series.

	MIN.	MAX.	MEAN	STD.DEV.	COEF. OF VARIATION	SKEW	MEDIAN
AE	0.96	745.2	63.2	151.2	239.1	3.48	7.2
VLI	365.9	895.6	739.7	150.9	20.4	-0.93	805.0
Z3	0.34	0.96	0.63	0.14	22.2	0.45	0.61

distribution of AE in Figure 13 is highly skewed, so that the modal class is 1.0 to 5.0 ha, and the median size is 7.2 ha. A small number of very large delineations create the skewness and elevate the mean to 63.2 ha, and the CV to 239. Chehalis is characterized by a dominance of small delineations (80% smaller than 50 ha) and the presence of a few very large ones.

The mean VLI for the Chehalis series is 739.7 and the median is even higher at 805.0. Although a few delineations have very intricate boundaries, indicated by a minimum of 365.9, most delineations have very regular boundaries. The skew value of 0.93 and the histogram of VLI show the influence of the few extreme VLI values within the distribution. The modal class of 850-900 which accounts for over 30% of the delineations, is also the highest class present. This class represents delineations with nearly convex boundaries.

The range of Z3 for the Chehalis sample is 0.34 to 0.96. The dominant class is 0.60 to 0.70, and 60% of the delineations fall between 0.50 and 0.70. Some Chehalis delineations have circular forms, but the data indicate a slight tendency toward an elongated nature for the typical form.

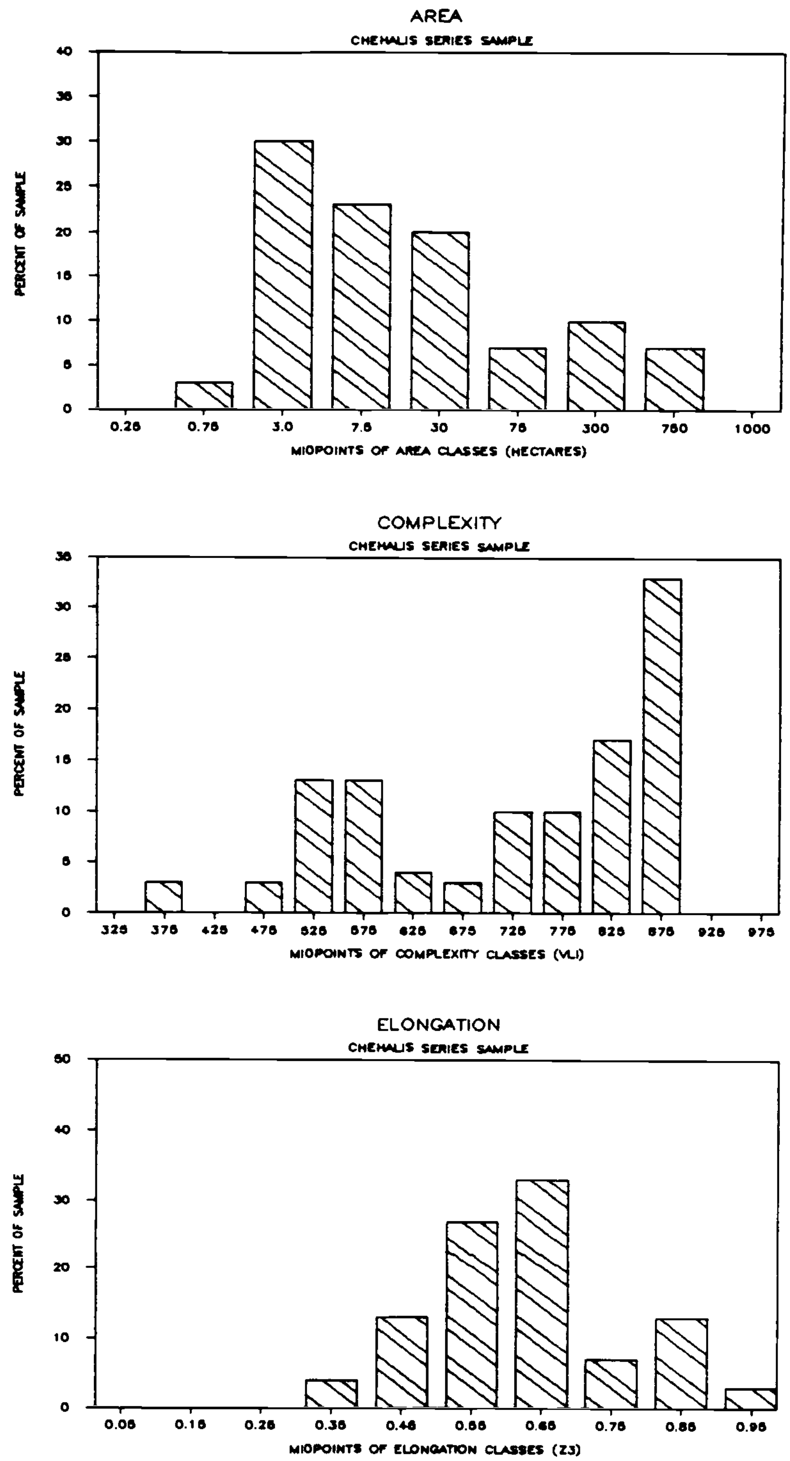


Figure 13. Frequency histograms of AE, VLI, and Z3 for the Chehalis series sample.

The Chehalis series has characteristically small, slightly elliptical delineations with very smooth outlines. The relationship of increased complexity with increased area is illustrated nicely by the Chehalis series. The data for VLI and AE are both influenced by a few delineations with uncharacteristic extreme values.

Waldo Series -- The Waldo series formed in recent alluvium on bottom lands of small streams and drainageways. The source of the alluvium is the highly weathered foothill areas, which contributes to a dominance of clayey textures. Waldo soils are deep and poorly drained. The hazard of erosion is slight, permeability is slow, and runoff is slow due mainly to the dominance of 0-3% slopes. The landscape position of Waldo delineations may have the greatest influence on their spatial properties.

The surface layer is about 28 cm thick and consists of black silty clay loam and silty clay. The subsoil is dark gray and gray clay about 94 cm thick. The substratum is gray silty clay extending to a depth of 1.5 m or more. Distinct mottling is common within the surface layer. This is mainly the result of a seasonal high water table.

Waldo delineations range in size from 0.6 to 94.0 ha. The distribution is positively skewed, but there are no real large delineations that distort the statistics for AE in Table 20. The mean size (16.8 ha) and the median (6.7 ha) are not widely different. The modal class of 1.0 to 5.0 ha shown in Figure 14 includes only slightly more than 25% of the sample.

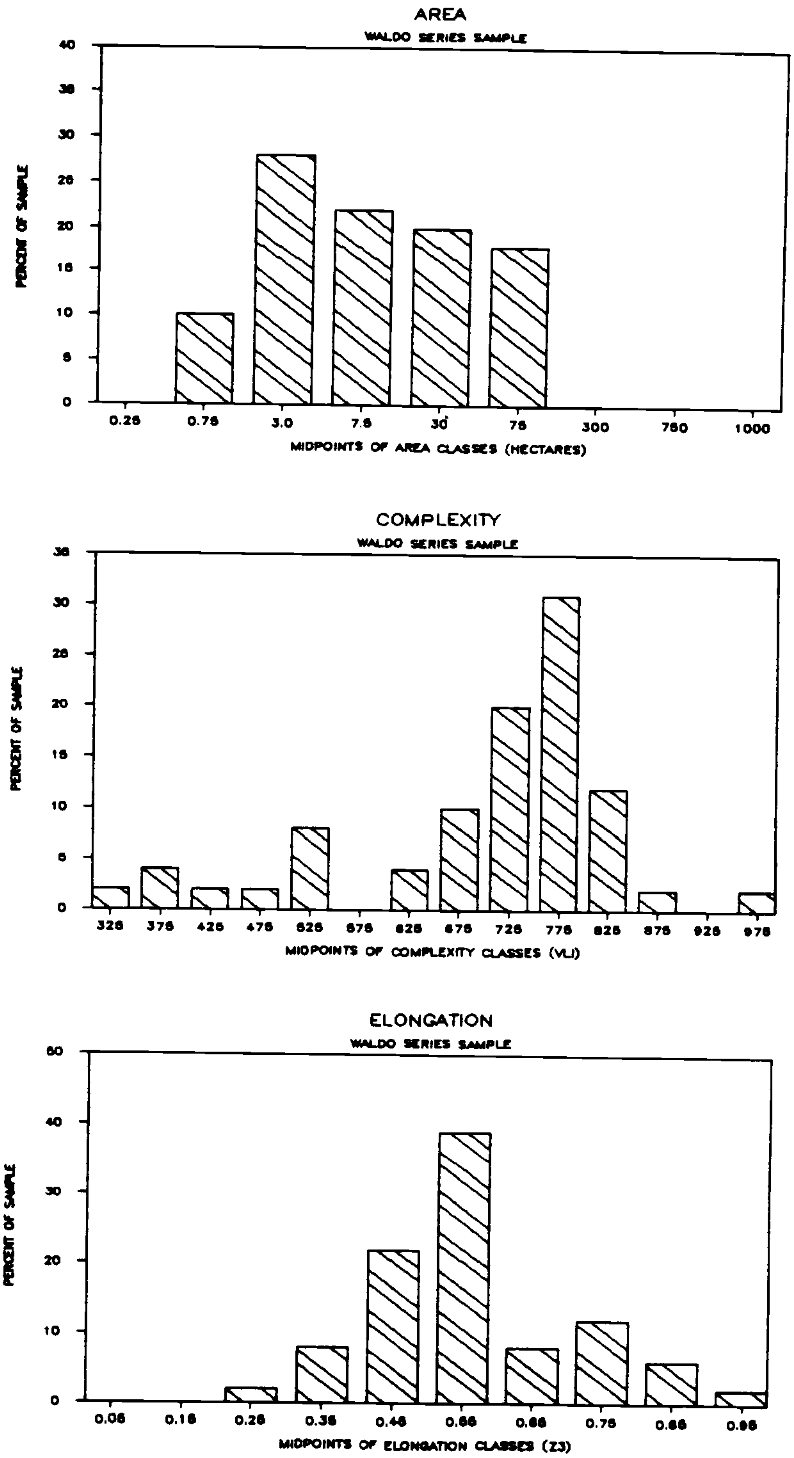


Figure 14. Frequency histograms of AE, VLI, and Z3 for the Waldo series sample.

The histogram for VLI indicates that most Waldo delineations have relatively smooth boundaries. 78% have indexes higher than 650. A subgroup of 18% of the delineations has VLI values less than 550, which indicates a high degree of complexity. These delineations account for the mean complexity (700.7) being a little lower than the median (737.6). For Waldo, however, the modal class (750-800), which by itself includes over 30% of all delineations, provides the most representative value for complexity.

Table 20. Descriptive statistics for AE, VLI, and Z3 for the Waldo series.

	MIN.	MAX.	MEAN	STD.DEV.	COEF. of VARIATION	SKEW	MEDIAN
AE	0.60	94.0	16.8	22.8	134.6	1.87	6.68
VLI	334.5	976.3	700.7	136.7	19.5	-1.09	737.6
Z3	0.30	0.97	0.57	0.14	24.8	0.61	0.54

The distribution of form index values for the Waldo sample is nearly normal. The mean, median, and mode have nearly the same values (0.57, 0.54 and, 0.55, respectively). This is a fairly low index number, indicating a relatively high degree of ellipticity. A few delineations, however, do have indexes above 0.80, indicating that the convex hull of some Waldo delineations can have a very circular form. Circular delineations would not be typical due to the definition of Z3 and the landscape position of the Waldo series. A delineation with a branching nature could yield a high Z3 value even though each branch represents a very narrow area of Waldo along a small tributary stream.

<u>Bellpine Series</u> -- Delineations of the Bellpine series represent

soils found on low foothills and rolling uplands along the Coast range. Bellpine is a residual soil with a colluvial parent material derived from sedimentary material. A representative profile has a dark reddish-brown silty clay loam surface soil about 15 cm thick. The subsoil is about 50 cm thick, with a texture that varies from silty clay loam to silty clay or clay. The substratum consists of fragmented, partly weathered sandstone. Permeability is slow, and the soil profile has common moderately thick clay films indicating a high degree of weathering. The slope phase descriptions for the Bellpine series indicate that runoff increases and the hazard of erosion becomes much more severe as slope increases.

Bellpine series delineations range in size from 1.7 to 332.8 ha (Table 21). The mean size is 74.4 ha, and the distribution has a positive skew of 1.28. The frequency histogram for AE, shown in Figure 15, has a bimodal distribution; 50% of the delineations fall between 5.0 and 50 ha, another group of larger delineations fall within the 100 to 500 ha range. The median size delineation is 18.8 ha, which is much less than the mean value of 74.4 ha.

Table 21. Descriptive statistics for AE, VLI, and Z3 for the Bellpine series.

	MIN.	MAX.	MEAN	STD.DEV.	COEF. OF VARIATION	SKEW	MEDIAN
AE	1.7	332.8	74.4	22.6	121.6	1.28	18.8
VLI	449.7	937.5	734.0	130.3	17.8	-0.15	730.2
Z3	0.42	0.87	0.70	0.12	17.1	-0.73	0.72

VLI has a much more normal distribution than AE for the Bellpine series. The distribution skew is only -0.15 with a mean of 734 and a

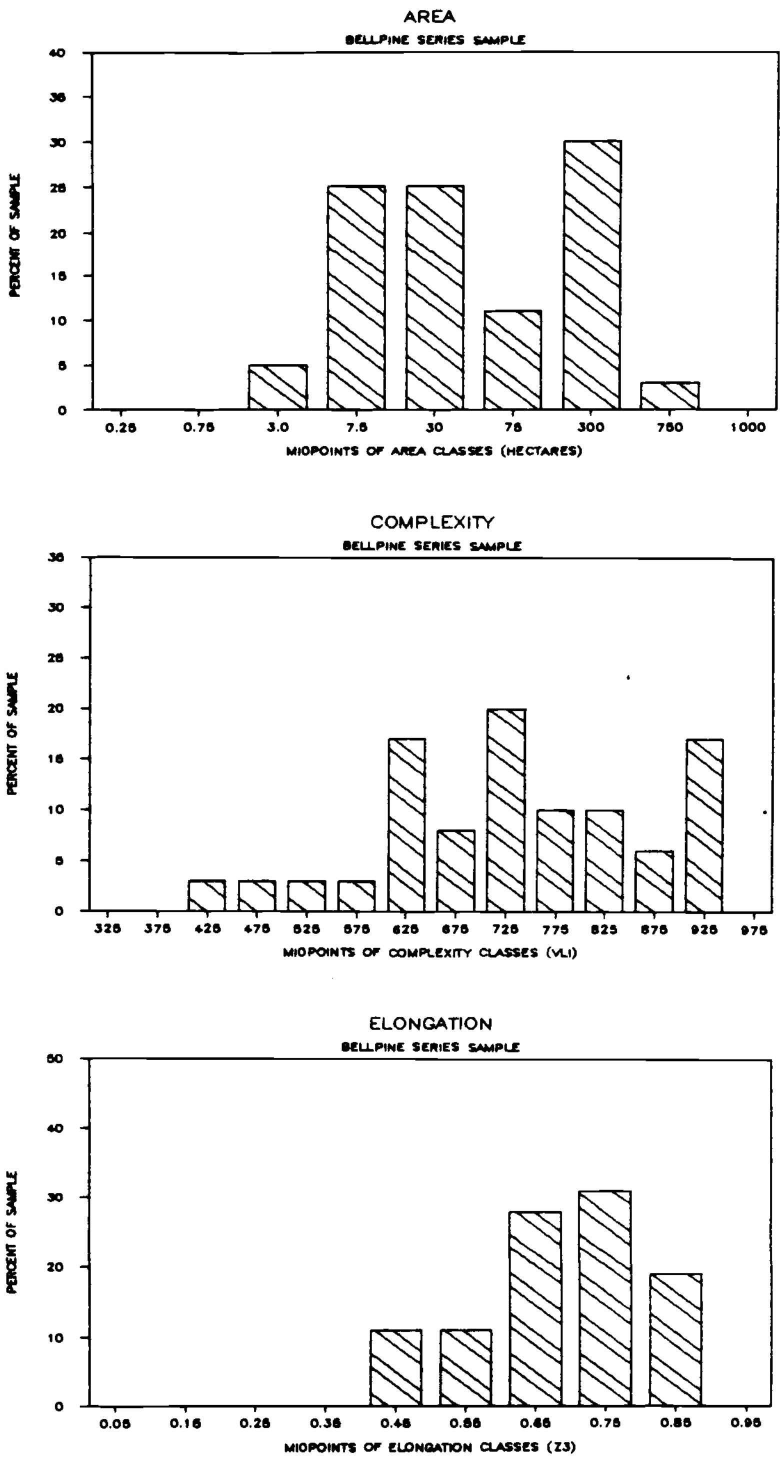


Figure 15. Frequency histograms of AE, VLI, and Z3 for the Bellpine series sample.

median of 730.2. The most complex delineation has a VLI of 449.7. The coefficient of variation is fairly low. No single class for VLI is characteristic for Bellpine delineations. The highest frequency is 22% of the delineations with a VLI between 700 and 750.

The form of Bellpine delineations tends to be nearly circular. The Z3 values range between 0.42 and 0.87; most are greater than 0.60. The distribution has a negative skew around a mean value of 0.70, which indicates a dominance of circular form.

Dixonville Series -- The Dixonville series delineations occupy the low foothills and higher rolling uplands along the coast range. The landform position is very similar to that of the Bellpine series. The surface soil is about 33 cm thick and has a texture of silty clay loam to silty clay. The subsoil is about 60 cm thick with textures of silty clay or clay. Weathered basalt bedrock underlies the subsoil at about 94 cm. Permeability is slow, hazard of erosion is high, and runoff is rapid. As with Bellpine, erosion and runoff would increase with an increase in slope.

The frequency histograms for AE, VLI, and Z3 are shown in Figure 16, and the descriptive statistics are given in Table 22. Delineation sizes range from a minimum of 1.6 ha to a very large delineation of 1086.4 ha, but 64% of the delineations are between 5.0 and 50 ha. Because large delineations are not common, the distribution has a high positive skew. Skewness also accounts for both the large difference between the mean (74.4 ha) and median (20.3 ha), and the high coefficient of variation (246). For the Dixonville series, median size is much more representative than mean size.

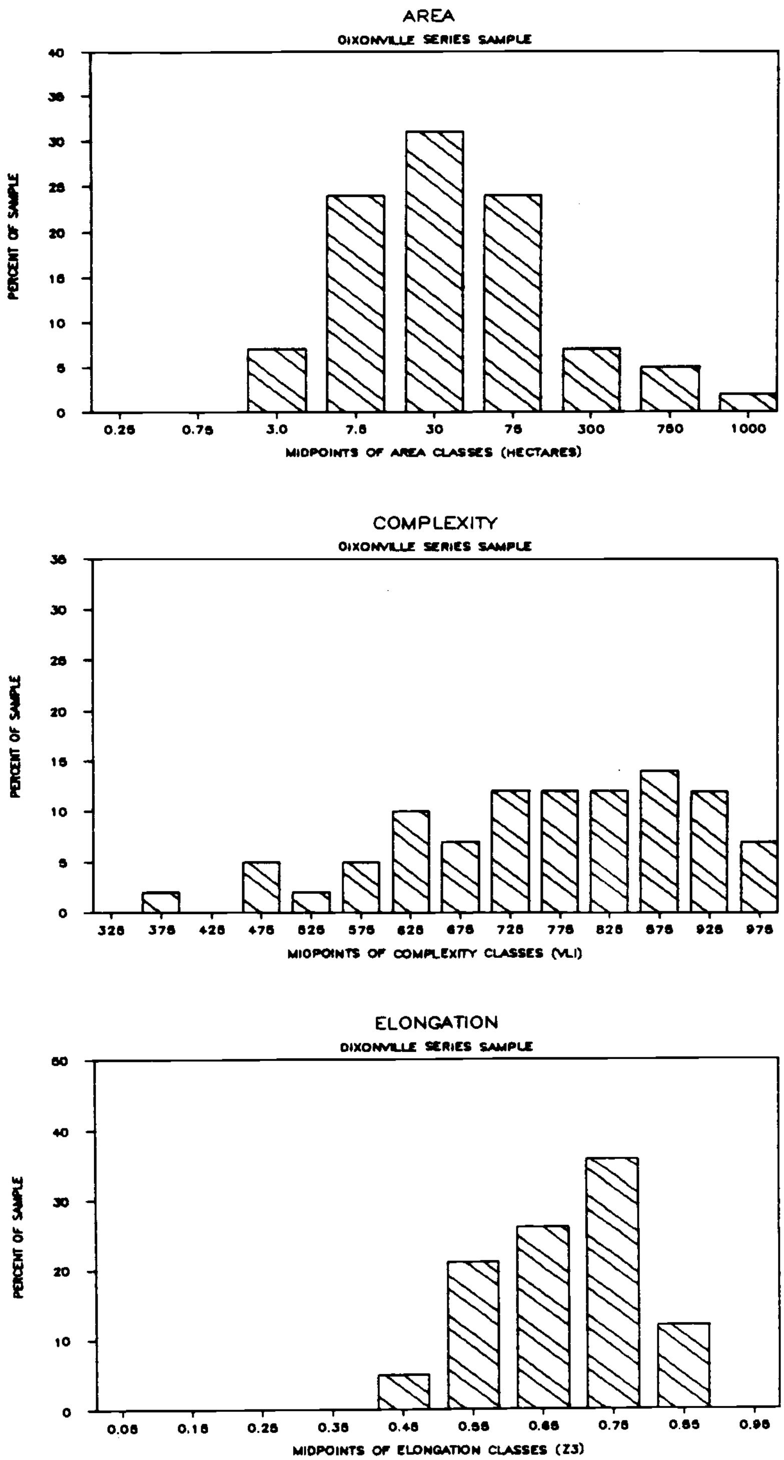


Figure 16. Frequency histograms for AE, VLI, and Z3 for the Dixonville series sample.

Table 22. Descriptive statistics for AE, VLI, and Z3 for the Dixonville series.

	MIN.	MAX.	MEAN	STD.DEV.	COEF. OF VARIATION	SKEW	MEDIAN
AE	1.6	1086.4	74.4	45.7	245.9	4.47	20.3
VLI	359.7	973.3	759.3	149.5	19.7	-0.70	765.7
Z3	0.45	0.86	0.69	0.11	15.9	-0.39	0.69

VLI values are well distributed along the experimental range. The delineations vary from very complex (359.7) to very regular (973.3). The mean is 759.3 and the median is 765.7. At the series level, most Dixonville delineations do not appear to have complex boundaries.

The distribution of delineation form is approximately normal with a slight negative skew. The skewness is caused by the occurence of a few elongated delineations. Otherwise the typical form is relatively circular.

Index Comparisons between Soil Map Units

A nonparametric test was used to compare the map unit samples with respect to area (AE), shape complexity (VLI), and form (Z3). The Mann-Whitney U-test produces calculated Z values for the comparison of each pair of samples based on a single index. The null hypothesis is rejected if the calculated Z value is greater than the absolute value of 1.96. This test criterion is based on a two-tailed test at the 0.05 level of significance. The null hypothesis is stated as: The two samples were taken from the same population. Any observed differences are due to chance. The alternative hypothesis is: The two samples were taken from different populations.

Soil Series Comparisons -- The distributions of AE for the Bellpine and Dixonville samples are similar, as are the group of Waldo, Chehalis, and Woodburn. The latter group represents alluvial and terrace soils with typically small delineations. The medians for AE shown in Table 23 are substantially larger for the Bellpine and Dixonville samples, which are located on the older foothill areas. The Bellpine and Dixonville samples are both significantly different than the Chehalis, Waldo, and Woodburn samples with respect to delineation size (Table 24).

Table 23. Median values for AE, VLI, and Z3 for each soil series sample.

	AE	VLI	Z3
Bellpine	18.8	730.2	0.72
Dixonville	20.3	765.7	0.69
Waldo	6.7	737.6	0.54
Woodburn	8.9	825.4	0.61
Chehalis	7.2	805.0	0.61

The two samples with apparently the most complex delineations are the Bellpine and Waldo. These samples have the lowest medians for VLI shown in Table 23. The Woodburn series stands out as the sample dominated by delineations with the most regular boundaries with a median of 825.4. Comparisons of the Waldo sample with Chehalis, Dixonville, or Woodburn are significant at the 95% confidence level (Table 24). All other comparisons show similarity between samples in terms of complexity.

Although the Waldo and Bellpine samples are similar in terms of complexity, they are quite different when comparison is made on the

Table 24. Mann-Whitney data for the comparisons of soil series samples based on AE, VLI, and Z3 results.

AE

	Dixonville	Waldo	Woodburn	Chehalis
Bellpine	0.80	3.76 *	3.28 *	2.49 *
Dixonville		3.38 *	2.64 *	2.06 *
Waldo			1.15	0.90
Woodburn				0.29

VLI

	Dixonville	Waldo	Woodburn	Chehalis
Bellpine	1.07	0.68	1.91	0.48
Dixonville		2.05 *	0.67	0.55
Waldo			3.55 *	2.11 *
Woodburn				1.06

Z3

	Dixonville	Waldo	Woodburn	Chehalis
Bellpine	0.56	3.90 *	2.86 *	2.18 *
Dixonville		4.09 *	2.78 *	2.00 *
Waldo			1.60	2.03 *
Woodburn				0.54

basis of form. Waldo has the lowest sample median for Z3 in Table 23, and the Bellpine median of 0.72 is the highest.

The grouping and significant comparisons of similar samples based on form is the same as for area. However, the Chehalis sample is significantly different than the Waldo sample based on form (Table 24). The Bellpine and Dixonville delineations tend to have circular form, and the Waldo, Chehalis, and Woodburn samples are similar in that they have typically elongated delineations.

<u>Comparisons between slope phases of the same series</u> --

a) Bellpine. Because delineations of the Bellpine series, as discussed previously, are composites of all adjacent slope phases; the mean, median, and maximum sizes for all four Bellpine slope phase samples (Table 25) are smaller than the same measures for the Bellpine series sample (Table 16). The sample mean and median increases sharply, however, as slope increases above 20%.

Results of the Mann-Whitney tests indicate that the 3-12% and 12-20% samples are not significantly different based on AE (Table 26). The 20-30% and 30-50% samples are also similar in this regard. However, at the 95% confidence level, a significant difference in delineation area does exist between samples of Bellpine soils on 3-20% slopes and those mapped on 30-50% slopes.

The comparison of the shape complexity between Bellpine slope phases shows a trend of increased complexity as slope increases. VLI medians in Table 25 show this trend nicely. Median complexity

Table 25. Descriptive statistics of AE, VLI, and Z3 for the Bellpine slope phase samples.

SLOPE %	MIN.	MAX.	MEAN	STD. DEV.	COEF. OF VARIATION	SKEW	MEDIAN
			AE				
3-12	1.9	68.0	20.0	20.8	104.1	1.29	9.5
12-20	1.4	92.0	17.6	21.6	122.6	2.24	9.7
20-30	3.0	200.0	45.6	52.0	114.9	1.58	30.0
30-50	3.5	237.6	49.6	58.0	117.2	1.80	27.2
			VLI				
3-12	436.7	884.3	737.4	135.6	18.4	-0.87	792.2
12-20	428.2	960.0	737.5	138.4	18.8	-0.53	765.6
20-30	336.8	909.4	679.0	176.6	26.0	-0.46	708.1
30-50	397.5	951.4	671.9	135.1	20.1	-0.0	661.0
			Z3				
3-12	0.53	0.84	0.66	0.07	11.2	0.44	0.65
12-20	0.46	0.87	0.66	0.12	18.6	0.20	0.67
20-30	0.45	0.84	0.67	0.12	17.7	-0.33	0.69
30-50	0.42	0.88	0.66	0.12	18.7	0.13	0.64

increases consistently from 792.2 for the 3-12% sample to 661.0 for the 30-50% sample. The negative skew values in Table 25 also show the tendancy for delineation complexity to increase as slope increases. However, none of the slope phases compared using the Mann-Whitney test (Table 26) are significantly different in terms of complexity.

The Bellpine slope phase samples do not differ a great deal in form. The Bellpine slope phase samples all have means and medians between 0.64 and 0.69. The null hypothesis is accepted for each comparison between samples in Table 26, i.e. each slope phase sample could have been drawn from the same population based on delineation form.

Table 26. Mann-Whitney data for the comparisons of Bellpine slope phases based on AE, VLI, and Z3.

BELLPINE SERIES

AE

	12-20%	20-30%	30-50%
3-12%	0.67	1.93	2.19 *
12-20%		2.79 *	2.83 *
20-30%			0.32

VLI

	12-20%	20-30%	30-50%
3-12%	0.10	1.10	1.79
12-20%		1.34	1.95
20-30%			0.41

Z3

	12-20%	20-30%	30-50%
3-12%	0.12	0.55	0.20
12-20%		0.36	0.00
20-30%			0.44

b) Dixonville. Both the mean and median values for AE in Table 27 increase as slope increases for the Dixonville slope phase samples. The minimum area value (14.2 ha), median (28.5 ha), and mean (57.6 ha) for the 30-50% sample is much larger than for the other samples. The 30-50% sample compared with all other samples has significantly different sized delineations (Table 28). The other slope phase samples are similar with respect to delineation area.

Table 27. Descriptive statistics for AE, VLI, and Z3 for the Dixonville slope phase samples.

SLOPE %	MIN.	MAX.	MEAN	STD DEV.	COEF. OF VARIATION	SKEW	MEDIAN
				AE			
3-12	1.8	71.6	15.2	4.3	113.4	1.96	8.2
12-20	1.6	110.0	18.4	6.2	135.7	2.64	9.4
20-30	3.2	174.4	24.4	7.6	123.7	3.63	15.6
30-50	14.2	269.6	57.6	17.2	119.5	2.26	28.5
				VLI			
3-12	456.4	993.8	798.5	112.4	14.1	-0.74	816.8
12-20	455.3	958.3	783.9	115.3	14.7	-0.87	800.8
20-30	504.9	931.6	798.3	116.0	14.5	-1.05	841.4
30-50	374.2	937.1	741.9	137.6	18.5	-0.69	768.2
				Z3			
3-12	0.50	0.86	0.67	0.10	14.7	0.03	0.67
12-20	0.37	0.90	0.66	0.12	17.9	-0.13	0.68
20-30	0.51	0.87	0.72	0.10	14.0	-0.55	0.75
30-50	0.38	0.88	0.67	0.11	16.0	-0.70	0.68

The only other significant difference between samples based on area is between the 3-12% and 20-30% samples. The 3-12% and 12-20% slope phase samples could have been drawn from the same population. Data for the means, medians, and maximums for AE in Table 27 suggest that the differences are due to a trend of increasing size with an increase in slope.

Table 28. Mann-Whitney data for the comparisons of Dixonville slope phases based on AE, VLI, and Z3.

	Dixonville		
AE			
	12-20%	20-30%	30-50%
3-12%	0.59	2.30 *	4.72 *
12-20%		1.70	4.60 *
20-30%			----
VLI			
	12-20%	20-30%	30-50%
3-12%	0.52	0.40	1.63
12-20%		0.87	1.15
20-30%			1.65
Z3			
	12-20%	20-30%	30-50%
3-12%	0.28	2.19 *	0.35
12-20%		2.15 *	0.46
20-30%			1.73

The shape complexity of Dixonville slope phase delineations is variable. According to the means in Table 27, the 3-12% and 20-30% samples have delineations with generally less boundary complexity than the 12-20% and 30-50% samples have. The median values show a similar trend. Based on VLI, none of the sample comparisons are significantly different (Table 28). Slope phase for the Dixonville series apparently has little if any impact on delineation shape complexity.

The form of Dixonville slope phase delineations appears to be influenced by slope phase. All of the medians and means for Z3, except for the 20-30% sample, cluster near 0.68 and 0.67. The central tendency of delineations for the 20-30% delineations is a circular form. At the 95% confidence level, the 20-30% population of delineations is significantly different than the other three Dixonville slope phases (Table 28). The three samples with the more elongated delineations are very similar.

<u>Comparisons between soil series for equal slope phase</u> --

The slope phases of the Bellpine and Dixonville series were compared by pairing samples with equal slope phase. Results for area and form show that the samples could have been drawn from the same population regardless of which soil series they are classified as (Table 31). The 20-30% slope phases show the greatest difference in central tendency (Tables 25 and 27), but the results are not significant for either area or form (Table 29).

Table 29. Mann-Whitney data for the comparisons between equal slope phases of Bellpine and Dixonville.

	3-12%	12-20%	20-30%	30-50%
AE	1.09	0.23	1.38	0.98
VLI	1.53	1.25	2.86 *	1.86
Z3	0.21	0.22	1.57	0.64

The Bellpine 20-30% sample is dominated by delineations with a different amount of boundary complexity than the Dixonville 20-30% sample. This difference is significant at the 95% level (Table 29). The difference between the 30-50% samples is not as noticeable. The medians and means for VLI given in Tables 25 and 27 are larger, i.e. less complexity, for the Dixonville than for the Bellpine regardless of slope phase. The 3-12% and 12-20% sample comparisons do not show a difference for delineation shape complexity. The impact of increasing slope on increased boundary complexity appears to be greatest for the Bellpine series. This may be do to the difference in erodibility of weathered sandstone parent material for Bellpine and weathered basalt for Dixonville.

CONCLUDING REMARKS

Needed experience, on which future studies may build, was obtained with each step in this shape analysis study. The literature review combined information from a wide range of disciplines concerned with the definition and quantitative study of shape. Perhaps the most useful information included in this section was the availability of modern shape measurements that can be incorporated with computerized geographic data handling systems. Soil surveys include the classification and spatial data needed for these modern systems. Automation of soil survey information might be the key to finding answers to questions about soil body shape and pattern.

Digitization of the individual delineations during the sampling procedure worked well. It was an efficient and accurate method of collecting and storing boundary information. Sampling controlled the amount of shape variability that was used to evaluate and select shape indexes. The results of the potential indexes suggest that the combined sample of 452 delineations tend to be regular, circular, and small. Perhaps this is true for all soil map delineations mapped at 1:20000, or it may only be characteristic of the map units that were selected for sampling. However, there was enough variation represented to meet the research objectives.

The first and primary objective was to develop a quantitative system for describing the shapes of soil map delineations that has interpretative value. Three indexes from a list of potential indexes that represented the common methods of shape measurement were selected as the best. The results of a factor analysis showed that complexity, elongation, and primary measurements are the dominant properties of soil map delineations.

The vertex lag method was expanded by approximating shapes more accurately than in the past and reducing the set of index values to a single value. These were improvments that were suggested by Fridland (1978). VLI was selected as the best measure of boundary complexity. Further experimentation with this method might attempt to find delineations that represent shapes along the entire range of 0.0 to 1000, determine if VLI is a true measure of shape, and use more than 22 edges for approximating delineations.

Delineation area was selected as the best representative of primary measurements. Areal information might be useful for providing better descriptions of populations of delineations defined by soil mapping units in soil surveys. The relationships of area with other spatial and classificational properties, such as complexity and slope, could improve descriptions and interpretations of soil mapping units.

The convex hull, rather than the irregular boundary of a delineation was used to arrive at a suitable measure of form. Schumm's elongation ratio for the convex hull (Z3) removed the influence of boundary complexity that affected the interpretability of common form ratios. The best use of Z3 might be in combination with a complexity index such as VLI. SIX is an example of this type of combination. Although SIX was not realized in time to make map unit descriptions and comparisons, it seemed to provide an interpretable index of both form and complexity.

The convex hull is an interpretable standard shape for soil map delineations that can be used to create several convex deficiency indexes. The relationships of convex deficiency on all or part of a delineation boundary with the processes involved with soil genesis and soil pattern provide another avenue for further research.

Evaluating the variability of spatial properties within and between soil mapping units was the second objective. The variability within the samples was described with descriptive statistics and some frequency histograms. The distributions of VLI, AE, and Z3 for each mapping unit in a soil survey could be described in just this way. This information would be useful for evaluating the uniformity, and making interpretations of map units by the makers and users of soil survey maps.

Results of the Mann-Whitney U test were used to make map unit comparisons. Significant differences between some, but not all samples were observed. The Bellpine and Dixonville series samples, both found on the foothill areas, were similar with respect to VLI, AE, and Z3 results. They both differed significantly from Chehalis, Woodburn, and Waldo based on AE and Z3. The only observed difference based on complexity was between Waldo and all other samples except Bellpine.

Although some comparisons between slope phases of the Bellpine and Dixonville series were significant, no strong trends were noticeable. For both series, the two steeper phases were significantly different from the 3-12% and 12-20% samples.

Complexity did not differentiate between slope phases. Z3 results were only significant when the 12-20% and 20-30% samples of Dixonville were compared. The 20-30% samples for Bellpine and Dixonville were significantly different in terms of complexity. All other comparisons of the two series based on equal slope phases were not significant. All testing was done at the 95% confidence level.

Shape differences and similarities between map units could be due to the classification of the soils themselves, adjacent soils, intensity of the soil survey, landform, or other causes. Slope and landform, as well as the erodibility of the soils, appear to be major factors influencing soil map delineation shape. Hopefully, this study will provide new tools and experience to assist future studies of shape and pattern analysis of soils.

LITERATURE CITED

Akl, S.G. and G. Tousaint, 1978. Efficient Convex Hull Algorithms for Pattern Recognition Applications. Proceedings Fourth International Joint Conference on Pattern Recognition. pp. 483-487.

Attneave, F. and M.D. Arnoult, 1956. The Quantitative Study of Shape and Pattern Recognition. Psychological Bulletin. pp. 452-471.

Bilton, J., 1983. Comparison of soil and vegetation map delineation shapes and areal correspondence. Masters Thesis. Oregon State University.

Blum, H., 1967. A Transformation for Extracting New Descriptors of Shape. Symposium on Models for the Perception of Speech and Visual Form.

Bookstien, F.L., 1978. The Measurement of Biological Shape and Shape Change. Lecture Notes in Biometrics. Number 24. Springer-Verlag: New York. Managing Editor; S. Levin.

Boots, B.N. and L.F. Lamoureux, 1972. Working Notes and Bibliography on the Study of Shape in Human Geography and Planning. Council of Planning Librarians, Exchange Bibliography Number 346.

Boyce, R.R. and W.A.V. Clarke, 1964. The Concept of Shape in Geography. Geographical Review, Volume (54): 561-572.

Bribiesca, E. and A. Guzman, 1979. Shape Description and Shape Similarity Measurement for Two-Dimensional Regions. Geo-Processing, Volume (1): 129-144.

Bunge, W., 1966. Theoretical Geography. Lund Studies in Geography, revised edition, Series C: General and Mathematical Geography 1: 1-210.

Cerney, J.W., 1975. Sensitivity Analysis of the Boyce-Clarke Shape Index. The Canadian Cartographer 12 (1): 21-27.

Chorley, R.J., 1959. The Shape of Drumlins. Journal of Glaciology 3: 339-344.

Chorley, R.J., D.E.G. Malm, and H.A. Pogorzelski, 1957. A New Standard for Estimating Drainage Basin Shape. American Journal of Science 225: 138-141.

Cox, E.P., 1927. A Method of Assigning Numerical and Percentage Values to the Degree of Roundness of Sand Grains. Journal of Paleontology 1 (3): 179-183.

Dent, B.D., 1972. A Note on the Importance of Shape in Cartogram Construction. Journal of Geography 71:393-401.

Duda, R.O. and P.E. Hart, 1973. Pattern Classification and Scene Analysis. John Wiley and Sons, New York.

Freeman, H., 1974. Computer Processing of Line-Drawing Images. Computing Surveys 6:57-86.

Freeman, H. and Shapira, 1975. Determining the Minimum-Area Encasing Rectangle for an Arbitrary Closed Curve. Communications of the ACM 18: 409-413.

Freeman, H., 1977. Shape Description Via the Use of Critical Points. Proceedings of the IEEE Computer Society on Pattern Recognition and Image Processing. pp. 168-174.

Freeman, H., 1979. On the Quantization of Line-Drawing Data. IEEE Transactions on Systems Science and Cybernetics. 5 (1): 70-79.

Fridland, V. M., 1965. Make-up of the Soil Cover. Soviet Soil Science 4:343-354.

Fridland, V.M., 1972. Pattern of the Soil Cover. (Translated from Russian by N. Kraner). Israel Program for Scientific Translations, Jerusalem, 1976. Available from U.S. Dept. of Commerce, Springfield, Va.

Graham, R.L., 1972. An Efficient Algorithm for Determining the Convex Hull of a Finite Planar Set. Information Processing Letters 1:132-133.

Hole, F.D., 1953. Suggested Terminology for Describing Soils as Three-dimensional Bodies. Soil Science Soc. Am. Proc. 17: 131-135.

Hole, F.D., 1978. An Approach to Landscape Analysis with Emphasis on Soils. Geoderma 21:1-23.

Horton, R.E., 1932. Drainage Basin Characteristics. Trans. of the American Geophysical Union 13:350-361.

Hudson, J.C. and P.M. Fowler, 1966. The Concept of Pattern in Geography. University of Iowa, Dept. of Geography, Discussion Paper Series, 1.

Green, et al, 1978. Soil Survey of Wasco County, Oregon: Northern part. National Cooperative Soil Surevey. Washington, D. C.

Jarvis, C.L., 1971. A Method of Fitting Polygons to Figure Boundary Data. Australian Computer Journal 3:50-54.

Jarvis, R.A., 1973. On the Identification of the Convex Hull Set of Points in the Plane. Information Processing Letters 2:13-21.

Haggett, P. and R.J. Chorley, 1969. Network Analysis in Geography. Edward Arnould Ltd. London. pp. 70-73.

Kellogg, C.E., 1963. Why a New System of Soil Classification? Soil Science 96 (1):1-5.

Kendall, sir Maurice, 1975. Multivariate Analysis. Hafner Press, New York. pp.53-54.

Knezevich, et. al., 1975. Soil Survey of Benton County Area, Oregon. National Cooperative Soil Survey. Washington, D.C.

Knox, E.G., 1965. Soil Individuals and Soil Classification. Soil Science Society of America Proceedings 29:79-84.

Lee, D.R. and G.T. Sallee, 1970. A Method of Measuring Shape. Geographical Review 60 (4): 555-563.

Matthews, J. A., 1981. Quantitative and Statistical Approaches to Geography. Pergamon Press Inc. Elmsford, N. Y., pp. 7-10.

Miller, V.C., 1953. A Quantitative Geomorphic Study of Drainage Basin Characteristics in the Clinch Mountain Area, Virginia and Tennessee. Dept. of Geology, Columbia University Technical Report Number 3, pp. 1-30.

Monmonier, M.S., 1974. Measures of Pattern Complexity for Chloroplethic Maps. The American Cartographer. 1:159-169.

Monmonier, M.S., 1982. Computer-Assisted Cartography. Prentice Hall, Englewood Cliffs, N.J..

Muehrcke, P.C., 1978. Map Use : Reading, Analysis and Interpretation. J.P. Publications, Madison, Wis..

National Cooperative Soil Survey, 1980. Guide to Minimum Sized Map Delineations. NSH Notice 58. U.S. Dept. of Agriculture. Washington, D.C.

Nagy, G. and S. Wagle, 1978. Geographic Data Processing. Computing Surveys 11 (2): 139-181.

Nagy, G. and S. Wagle, 1979. Approximation of Polygonal Maps by Cellular Maps. Communications of the ACM 22:518-525.

Ongley, E.D., 1970. Drainage Basin Axial and Shape Parameters from Moment Measures. The Canadian Cartographer 14:38-44.

Pavlidas, T., 1978. Algorithms for Shape Analysis of Contours and Waveforms. Proceedings of the Fourth International Joint Conference on Pattern Recognition. pp.70-85.

Piech, P.A., 1980. Selected Concepts of Soil Map Delineations. Masters Thesis, Cornell University.

Robinson, A.H., R.D. Sale and J.L. Morrison, 1978. Elements of Cartography, 4th edition, Wiley, New York.

Schumm, S.A., 1956. Evolution of Drainage Systems and Slopes in Badlands at Perth Amboy, New Jersey. Bulletin of the Geological Society of America 67:597-646.

Sklansky, J.,1974. Measuring Concavity on a Rectangular Mosaic. IEEE Transactions on Computers, C-21, (12):1355-1364.

Soil Survey Staff, 1951. Soil Survey Manual. Agricultural Handbook Number 18, U.S. Dept. of Agriculture, Washington, D.C..

Stoddart, D.R., 1965. The Shape of Atolls. Marine Geology 3: 369-383.

Stenson, H.H., 1966. The Physical Factor Structure of Random Forms and Their Judged Complexity. Perception and Psychophysics 1:303-310.

Taylor, P. J., 1970. Distances within shapes: an Introduction to a family of finite frequency distributions. University of Iowa, Dept. of Geography, discussion paper Series, 16.

Valentine, K., 1981. How Soil Map units and Delineations Change with Survey Intensity and Map Scale. Canadian Journal of Soil Science 61:535-551.

Van Wambeke, A., 1966. Soil Bodies and Soil Classification. Soils and Fertilizers 29:507-510.

Weast, R. C., et.al., 1964. Mathematical Tables from the Handbook of Chemistry and Physics. 12th Edition. The Chemical Rubber Co.. Cleveland, Ohio.

Zahn, C.T. and R. Z. Roskies, 1972. Fourier Descriptors for Plane Closed Curves. IEEE Trans. Computers, C-21:269-281.