Geografiska Annaler: Series B, Human Geography

ISSN: 0435-3684 (Print) 1468-0467 (Online) Journal homepage: http://www.tandfonline.com/loi/rgab20

Compactness of Geographic Shape: Comparison and Evaluation of Measures

Alan M. Maceachren

COMPACTNESS OF GEOGRAPHIC SHAPE: COMPARISON AND EVALUATION OF MEASURES

BY
ALAN M. MACEACHREN*

ABSTRACT. Geographic shape compactness is of concern in such contexts as urban morphology, political districting, and accuracy of enumeration unit values. The purpose of the present paper is to compare, categorize, and evaluate the various methods suggested for measuring compactness. Review of suggested compactness indices resulted in the identification of four categories, based upon: 1) perimeter-area measurement, 2) single parameters of related circles, 3) direct comparison to a standard shape, and 4) dispersion of elements of a shape's area. In all, eleven indices were calculated for a sample of U.S. counties and their frequency distributions were compared. Additionally, similarities and differences in the units of high, median, and low compactness were identified. Within each category, indices exhibited considerable similarity while significant differences were evident from category to category. These differences suggested variation in accuracy with which the indices measure compactness. It was hypothesized that measures based upon dispersion of elements of a shape's area would provide the most accurate measure because they consider the shape as a whole while the other indices do not. Determination of the accuracy with which each index measures compactness supported this hypothesis. Indices based upon direct comparison to a standard shape were found to be of similar accuracy and, therefore, also judged to be suitable compactness measures. Those measures based upon perimeter-area measurement and single parameters of related circles were determined to be considerably less accurate measures of compactness.

Shape is a fundamental spatial characteristic of concern in geographic investigation. Geographical description often makes reference to the shape of regions or features, with analogies sometimes being made to the shape of common objects, as in the Grain "Belt" or the "boot" shape of Italy. More often, specific shape descriptors have been employed (e.g. symetrical, elongated, linear, sinuous, or compact). In recent years there have been attempts to go beyond verbal description of geographic feature shape to more precise mathematical description.

Two approaches are apparent in attempts to quantify shape. The first is to develop a measure of shape uniqueness by which any shape can be distinguished from all other shapes and similar shapes result in similar descriptions. Bunge's (1966) internal distance sum measure and application of dual axis fourier analysis to shape description by Moellering and Rayner (1981) are representative of this approach. Alternatively, measures have been developed of individual aspects of shape thought to be relevant in geographic analysis (e.g., elongation, dissection, and compactness of regions, indentation of borders, sinuosity of linear features, and symmetry of networks). Of these, region compactness has been given the greatest attention due to its potential applicability to a broad range of geographic problems. Relationships have been hypothesized between region compactness and such characteristics as transportation system efficiency, accessibility of services, and homogeneity of regions. These hypothesized relationships have led to consideration of compactness in geographic investigations of such topics as urban morphology (Lo, 1980), village comparisons (Lee and Salle, 1970), political districting (Morrill, 1981; Johnston and Rossiter, 1981; Austin, 1981), and accuracy of enumeration unit values and resultant maps (Robinson and Hsu, 1970; Coulson, 1978; Fairchild, 1981).

Measures of compactness were suggested as early as 1822 by Carl Ritter (Frolov, 1975). A variety of procedures have since been employed in attempts to develop an acceptable measure of compactness. These procedures vary in the underlying approach used, in the ease of calculation, and presumably in the accuracy with which compactness can be characterized across the range of possible geographic shapes. The purposes of the present paper, therefore, are to examine characteristics of compactness measures used by geographers and others, compare values obtained by these measures for a variety of geographic shapes, and evalutate the relative accuracy of each as a compactness measure.

A large variety of methods by which region shape can be measured have been presented in the literature. Often, authors fail to make a distinction concerning the aspect of shape (e.g., compactness, elongation, indentation, etc.) a measure is designed to represent. In some cases measures of specific shape characteristics are actually presented as being representative of shape unique-

* Associate Professor Alan M. MacEachren, Department of Geography, 302 Walker Building, The Pennsylvania State University, University Park, PA 16802

3.54 : 1 3.72 : 1 4.0 : 1 4.56 : 1

Fig. 1. Perimeter to area rations of simple geometric shapes.

ness (e.g., Pounds, 1972; Boyce and Clark, 1964). Because of this lack of distinction among region shape measures, those selected for consideration here include both those specifically intended as compactness measures and those which have been applied to analysis in which compactness is a major concern.

Description of measures

There exist a multitude of measures that have been developed to, or have potential to, represent shape compactness. For ease of comparison and evaluation, these measures can be grouped into four general categories: perimeter and area measurements, single parameters of related circles, direct comparison to a standard shape, and dispersion of elements of area around a central point.

The intent here is to consider basic approaches taken to compactness measurement not to trace development of all published measures. Similarities and differences among related measures will be examined and a set of measures representative of the range of approaches taken will be identified for comparison and analysis. Emphasis will be placed on compactness measures that have been applied to geographic investigation, however, those developed for use in other fields of inquiry will also be considered.

Perimeter-area-measures

For simple geometric shapes a readily apparent relationship exists between the perimeter and area of a shape and its compactness (Figure 1). It is not surprising, therefore, that some of the earliest attempts to develop a compactness index relied on perimeter to area ratios. In geography the simple ratio of perimeter of a shape to its area was suggested as early of 1822 by Ritter (Frolov, 1975). Initial attempts by psychologists to analyze visual perception of form (Bitterman, Krauskopf, and Hochberg, 1954) also relied on this simple measure.

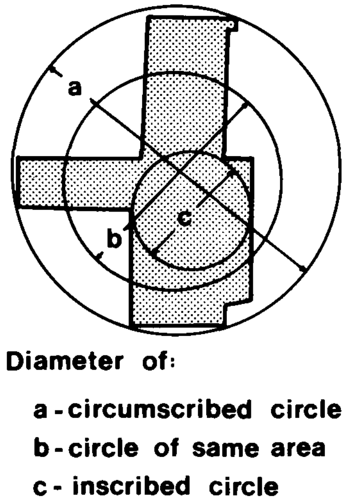

Fig. 2. Inscribed circle, circumscribed circle, and circle of same area for a sample unit.

A problem with the perimeter/area ratio is that it varies with size. Squaring the perimeter value or taking the square root of the area eliminates this difficulty. The latter approach, modified by a constant, produces a value range of zero to one. Nagel suggested the form

$$\frac{\sqrt{\text{Area}}}{.282 \times \text{perimeter}} \tag{1}$$

which yields values of zero for a circle increasing to one as the shape becomes less and less compact (Frolov, 1975). For convenience most subsequent authors have modified the index by subtracting the value from one (Attneave and Arnoult, 1957) or reversing the numerator and denominator (Pounds, 1972; Rimbert, 1976). In either case, values increase to one as a shape approaches a circle.

The alternative to taking the square root of the area, squaring the perimeter value, has been applied to measuring shape in a variety of contexts such as drainage basin analysis (Miller, 1953), form perception (Attneave, 1957), evaluating political units (Haggett, 1973), and computer form identification (Duda and Hart, 1973). Generally presented in a form such as

$$\frac{\text{Area}}{(.282 \times \text{perimeter})^2} \tag{2}$$

this measure is simply the square of that described above. Differences among relatively compact shapes, therefore, will be made more readily apparent while those among extremely irregular shapes will be deemphasized.

Parameters of related circles

A second group of compactness measures are those based on specific parameters of circles having some direct relationship to a shape. Parameters measured are the diameter of the smallest circle that can circumscribe the shape (i.e., the shape's longest axis), the largest circle that can fit inside the shape (i.e., the largest inscribing circle), and the circle having the same area as the shape (Figure 2). Variations between the shape's longest axis or area and the area or diameter of the circle selected, constitutes a measure of the departure from the most compact possible shape.

Comparison of the area of a shape to that of the smallest circumscribing circle was suggested as early as 1892 by Ehrenberg (Frolov, 1975) and has been employed more recently in relation to topics such as political districting (Reock, 1961) and urban form (Gibbs, 1961). Area of the circumscribing circle can be calculated from the diameter, or the shape's longest axis, according to the formula

$$\pi \cdot (.5 \times \text{longest axis})^2 \qquad (3)$$

To produce an index that increases from zero to one with increasing compactness, the area of the shape is divided by (3).

$$\frac{\text{Area}}{\pi \cdot (.5 \times \text{largest area})^2} \qquad (4)$$

The related measure,

$$\frac{\text{diameter of a circle having the same area as the shape}}{\text{diameter of the circumscribing circle}} \qquad (5)$$

was suggested by Schumm (1956) as a measure of drainage basin shape. This index has the form

$$\frac{2 \cdot \sqrt{\text{area}/\pi}}{\text{Longest axis}} \qquad (6)$$

It is apparent that (6) is simply the square root of (4). As with equations (1) and (2), differences among compact units will be emphasized by (4) while differences among non-compact units will be brought out with (6).

Rather than comparing the circumscribed circle to the shape itself, it has also been compared to the largest inscribed circle (Frolov, 1975).

$$\frac{\text{area of inscribed circle}}{\text{area of circumscribed circle}} \qquad (7)$$

This procedure has the practical advantage that the investigator need only determine the area of two circles rather than a circle and an irregular shape.

The square root of the above measure

$$\frac{\text{diameter of inscribed circle}}{\text{diameter of circumscribed circle}} \qquad (8)$$

has also been employed (Haggett, 1966). Differences between these indices should be similar to those cited above for similar pairs of measures.

Using parameters of the shape and its corresponding inscribed and circumscribed circles, an additional measure can be devised.

$$\frac{\text{Area of the inscribed circle}}{\text{area}} \qquad (9)$$

This form was also suggested in 1892 by Ehrenberg (Frolov, 1975), however, there seems to have been little or no application of it in geography. The square root version of the measure has apparently not been used at all.

Direct comparison to standard shape

Rather than simply comparing the area or perimeter of a shape to that of a related circle, it is possible to measure differences between a shape and some standard shape more directly. One such approach is an index presented by Lee and Salle (1970). To compute the index, the shape and standard unit are first superimposed. Lee and Salle suggest an arrangement to maximize the common area. A more easily replicable method, however, is to align the shape and standard unit so that their centroids correspond. Generally when compactness is of concern, the standard unit used will be

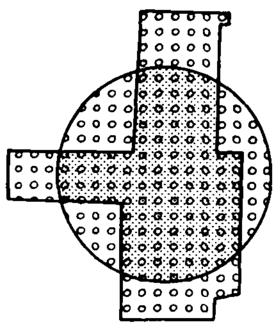

Fig. 3. Union and intersection of circle having the same area as a unit.

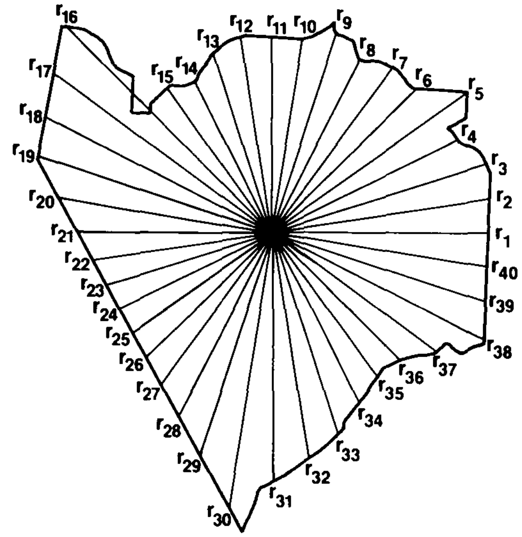

Fig. 4. Calculation of Boyce-Clark index with radials plotted from centroid to parameter.

a circle, making orientation of the standard irrelevant.

Once the shape and standard circle are aligned, the areas of union and intersection are calculated (Figure 3). The index is determined by

$$\frac{\text{area of intersection of the shape and circle of same area}}{\text{area of union of the shape and circle of same area}} \qquad (10)$$

A second method for comparing a shape to a standard directly is the radial line method developed by Boyce and Clark (1964). The measure consists of drawing a set of equally spaced radials from a location within a shape to its perimeter. If the shape were a circle, and the point of origin for radials the center of that circle, all radials would be of equal length. Compactness of a shape is then measured by the variation in lengths of the radials (Figure 4). Boyce and Clark (1984) present the measure in the form

$$\Sigma \left| \frac{ri}{\Sigma ri} \cdot 100 - \frac{100}{n} \right| \qquad (11)$$

where *r* is the length of radial *i* and *n* is equal to the number of radials. Values produced by the index range from zero for a circle to 200 [rather than 175 as first cited by Boyce and Clark (Blair and Biss, 1967)]. Slight modification of the equation produces a more practical range decreasing in value from one for a circle to zero for a straight line

$$1 - \left[\frac{\Sigma \left| \frac{ri}{\Sigma ri} \cdot 100 - \frac{100}{n} \right|}{200} \right] \qquad (12)$$

Being somewhat easier to calculate than the index proposed by Lee and Salle, the Boyce-Clark index has seen wide application in geographic and related analysis. Examples include investigations of urban morphology (Lo, 1980; Wilkins and Shaw, 1971), political districts (Austin, 1981), atoll shape (Stoddart, 1965), sand grain shape (Ehrlich and Weinberg, 1970), isopleth map accuracy (Fairchild, 1981), and even cartogram units (Dent, 1972).

Dispersion of elements of area

Both the Lee-Salle and Boyce-Clark measures are attempts to consider any shape as a whole rather than so focus on a single parameter of that shape. In both cases, however, there remains a focus on

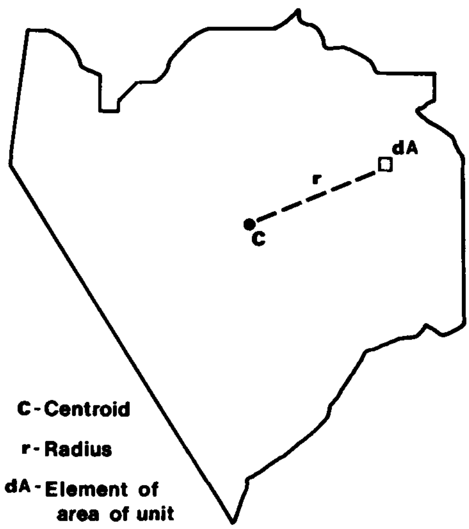

Fig. 5. Dispersion of elements of area (dA) around centroid (c).

limited characteristics; area of mismatch between the shape and a standard and distances to the perimeter respectively.

Two related measures have been developed to overcome limitations inherent in compactness measures that emphasize particular aspects of shape. For each, a shape is considered to be composed of a series of infinitesimally small elements of area dA. Dispersion of these elements around a shape's centroid is the basis of the measure (Figure 5).

Each index seems to have been derived independently by a number of authors in relation to geographic problems (e.g., Blair and Biss, 1967; Massam and Goodchild, 1971; Whittington, Beavon, and Mabin, 1972; Bachi, 1973), as well as by psychologists searching for objective measures which relate to subjective judgement of form (e.g., Zusne, 1965). In some cases the index is explained in terms of the concept of moments used in physics (e.g., Zusne, 1965; Massam and Goodchild, 1971). Alternatively, the measure has been presented from a purely statistical viewpoint in which the variance or standard deviation for locations of elements dA serve as the basis.

In terms of moments of area, the measure takes the form

$$\frac{A^2}{2 \cdot \pi \cdot \int \cdot r^2 \cdot dA} = \frac{A^2}{2 \cdot \pi \cdot \int \cdot r^2 \cdot dx \cdot dy} \tag{13}$$

where A is the area of the shape and r the distance from the shape's centroid to a particular element of area (dA). The terms dx and dy represent the length and breadth of infinitesimally small rectangles. Division by 2 results in a range of values for the index from one for a circle decreasing to zero for decreasingly compact shapes.

The equivalent index emphasizing statistical measures of variance takes the form

$$\frac{A}{2 \cdot \pi \cdot (\sigma_X^2 + \sigma_Y^2)} \tag{14}$$

where A is the area of the shape and $\sigma^2 x$ and $\sigma^2 y$ represent the variance of the x and y locations for each element of area dA. Bachi (1973) has demonstrated that when used in this form, the index can be calculated directly from a series of $x - y$ coordinate pairs used to define the boundary of the shape. The measure was refered to by Bachi as the Relative Distance Variance.

The second measure in this category was derived by Blair and Biss (1973). The actual measure derived, however, represents the square root of (14) and can be expressed as

$$\sqrt{\frac{\text{Area}}{2 \cdot \pi \cdot (\sigma_x^2 + \sigma^2{}_y)}} \tag{15}$$

It has been suggested that this index will exhibit a somewhat smaller range of values than will index (14) (Whittington, Beavon, and Mabin, 1972). Actually, while this will be true for relatively compact shapes, the opposite will be true for the least compact shapes.

A particular advantage of these measures of dispersion has been pointed out by Massam and Goodchild (1971). An ability exists to distinguish between the geometric shape of the unit being measured and the shape of a specific distribution within that unit. The latter is accomplished by simply weighting each element of area, dA, according to the distribution of the variable being examined.

Comparison of measures

In all, eleven distinct, though not necessarily independent, measures have been identified (Table 1). Due to apparent independent development of several indices by different authors, indices will be

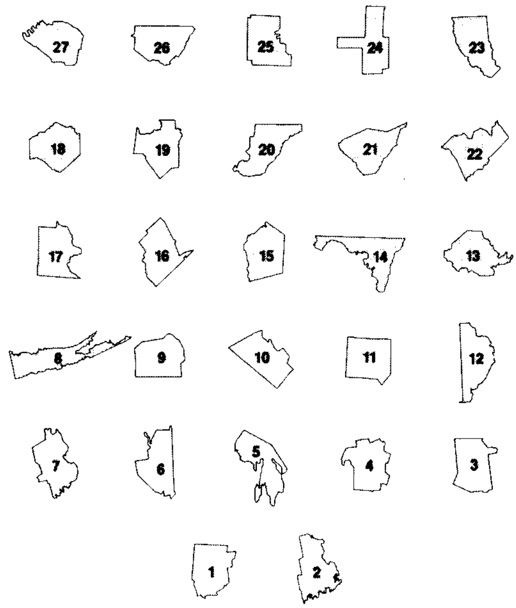

Fig. 6a. Sample counties, eastern U.S.

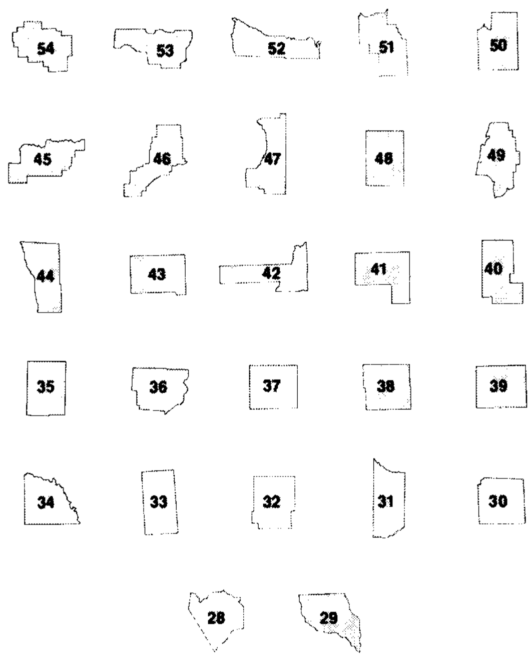

Fig. 6b. Sample counties, western U.S.

referred to by the letters assigned in the table rather than by the name of their developer.

Although there have been previous attempts to categorize shape measures in general (e.g., Frolov, 1975) and compactness measures in particular (Whittington, Beavon, and Mabin, 1972), little effort has been directed to systematic comparison of compactness indices.

In one of the few such attempts, Stoddart (1965) compared means and frequency distributions for four of the indices considered here (B, C, D, and I). For 99 variously shaped atolls, he demonstrated that the indices were significantly different in range and distribution of values as well as in rank ordering of atolls. Stoddart's results emphasize the need to systematically consider the nature and magnitude of differences among compactness measures.

Sample units and calculation of indices

The initial step in comparison of the indices under consideration was to select a sample of units for which the indices could be calculated. Compactness is a relevant factor in description and analysis of many kinds of geographic shapes. Its use in relation to administrative units, however, has been most wide spread, therefore, a sample of administrative units was selected for use in the present study. To insure an adequate range in size and shape of units, a stratified random sample of six counties from each of the nine census divisions of the conterminous United States were selected (Figure 6a and 6b). The stratification by region results in a range from quite irregular units from the eastern United States, where a metes and bounds system of land division prevails, to the regularly shaped units in central United States, where a township and range system was dominant.

Each of the indices was calculated for all 54 sample shapes. The diameters of the largest possible inscribed circles were determined by hand. All other computations were accomplished using a FORTRAN computer program based on the equations presented in Table 1. Complete descriptions of the steps necessary in calculating indices J and K can be found in publications by Bachi (1973) and Blair and Bliss (1967) respectively.

For all indices except index I (the Boyce-Clark index) only one value is possible for any given shape. This index, however, has four variable components that must be considered: the number of radials used, the orientation of these radials,

Table 1. Compactness Measures Evaluated

A. $\dfrac{.282 \times \text{perimeter}}{\sqrt{\text{area}}}$

B. $\dfrac{\text{area}}{(.282 \times \text{perimeter})^2}$

C. $\dfrac{\text{area}}{\pi \cdot (.5 \times \text{longest axis})^2}$

D. $\dfrac{\sqrt{\text{area}/\pi}}{(.5 \times \text{longest axis})}$

E. $\dfrac{\text{area of inscribed circle}}{\text{area of circumscribed circle}}$

F. $\dfrac{\text{diameter of inscribed circle}}{\text{diameter of circumscribed circle}}$

G. $\dfrac{\text{area of inscribed circle}}{\text{area}}$

H. $\dfrac{\text{area of intersection of the shape and circle of same area}}{\text{area of union of the shape and circle of same area}}$

I. $1 - \left[\dfrac{\Sigma \left| \dfrac{ri}{\Sigma ri} \times 100 - \dfrac{100}{n} \right|}{200} \right]$

J. $\dfrac{A}{2\pi\,(\sigma_x^2 + \sigma_y^2)}$

K. $\sqrt{\dfrac{A}{2\pi\,(\sigma_x^2 + \sigma_y^2)}}$

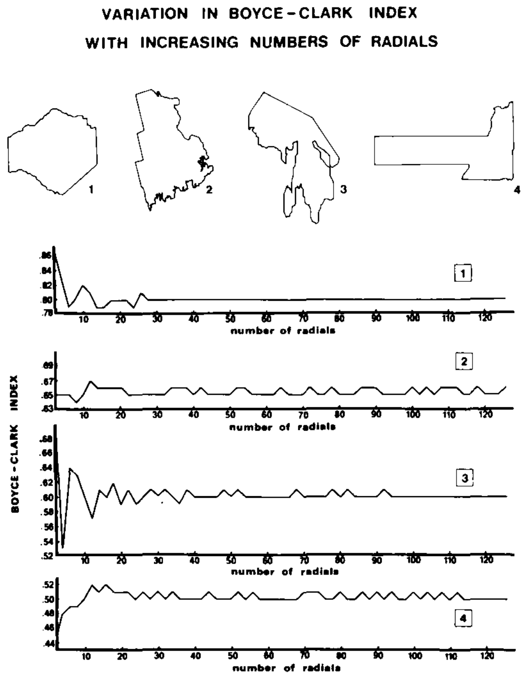

Fig. 7. Influences of the number of radials on index I (modified Boyce-Clark index).

the location of the point from which radials are projected, and the procedure for dealing with radials that intersect the perimeter more than once.

In the present case, the origin for radials has been defined as the center of gravity of the unit. For radials that cross the shape's outline more than once, the length of the longest radial is first determined. To compensate for indentations or gaps in the shape that the radial crosses, the length of any segments of the radial not in the shape is added to the longest radial.

In relation to the number of radials and their orientation, evidence has been presented by Cerny (1975) that, with the center of gravity as an origin, increasing the number of radials will cause a decrease in the index value to some point at which the index will stabilize. Given sufficient radials, the index will be independent of their orientation.

For one shape, the state of New Hampshire, Cerny found that the index seemed to stabilize at 32 to 40 radials. To determine whether this range is consistent across shapes or varies with complexity of the shape, four units from the present sample were tested. They represent a range from the most compact to the least compact (according to index J) As can be seen by the graphs in Figure 7, the index stabilizes between 15 and 40 radials depending on characteristics of the shape. There is, however, a slight fluctuation of ± .01 up to 120 radials for the shape with the most intricate outline. For purposes of accurate comparison with other indices, therefore, 120 radials were used to calculate index I in the present study. In practice, however, it appears that 40–50 radials should be sufficient for most shapes.

Variation among indices

A correlation matrix calculated among all eleven indices demonstrates the variability existing among these measures (Table 2). While it is obvious that all of the indices are related, with some being nearly identical, correlation coefficients as low as 0.70 are found between certain measures. These differences indicate that the choice of a compactness index could lead to significant differences in the outcome of any geographic analysis in which compactness was a major factor (e.g. political districting, evaluation of map accuracy, etc.).

Examination of frequency distributions for each of the measures further emphasizes their variability. Considerable differences are found in both the range of values obtained for the sample units and the shape of the distributions themselves (Figure 8 a-k). Based on individual indices, the compactness of sample U.S. counties could be characterized as normally distributed, bimodal, generally compact, generally not compact, or having little discernable pattern whatsoever.

Differences that exist are, as would be anticipated, greatest between categories of measures, and somewhat less within each of the four categories. Values derived by the indices from the perimeter-area category (A and B), for example, result in a similar rather variable distribution for both indices. Each indicates a small number of noncompact units and an irregular distribution of values spanning about one-half of the entire range of compactness. According to both indices, the least compact unit is unit 8 and the most compact is unit 37.

In comparison with indices A and B, unit 37 cited as the most compact for both, ranks between fourth and ninth on indices C through G. The most compact unit on four of these five measures, unit

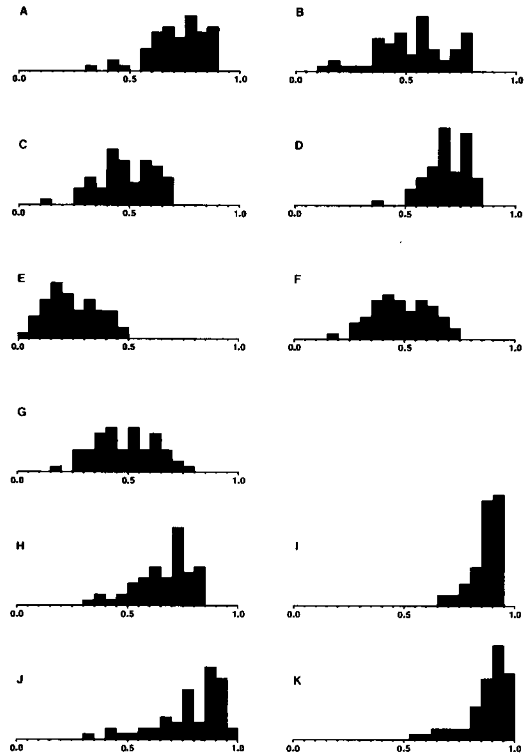

Fig. 8. Frequency of compactness values according to indices A through K.

18, is ranked twelfth and fourteenth on indices A and B respectively.

The remaining four measures, those involving direct comparison to a standard shape and the two measures of dispersion of elements of area, all produce similar frequency distributions. The units rated at the extremes of the range are similar as well. A distribution skewed toward compact units

Table 2. Correlations Among Index Values.

	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
A	1.00	0.99	0.74	0.77	0.75	0.77	0.70	0.74	0.75	0.77	0.77
B	0.99	1.00	0.76	0.78	0.77	0.79	0.72	0.74	0.75	0.77	0.76
C	0.74	0.76	1.00	0.96	0.92	0.93	0.78	0.86	0.86	0.89	0.86
D	0.77	0.78	0.96	1.00	0.90	0.93	0.78	0.92	0.90	0.93	0.93
E	0.74	0.77	0.92	0.90	1.00	0.99	0.95	0.89	0.85	0.88	0.85
F	0.77	0.79	0.93	0.93	0.99	1.00	0.95	0.94	0.90	0.93	0.91
G	0.70	0.72	0.78	0.78	0.95	0.95	1.00	0.87	0.83	0.84	0.82
H	0.74	0.73	0.86	0.92	0.89	0.94	0.87	1.00	0.98	0.98	0.97
I	0.75	0.75	0.86	0.90	0.85	0.90	0.83	0.98	1.00	0.98	0.98
J	0.77	0.77	0.89	0.93	0.88	0.93	0.84	0.98	0.98	1.00	1.00
K	0.77	0.76	0.86	0.92	0.85	0.91	0.82	0.97	0.98	1.00	1.00

Table 3. Units in the Median Range of Compactness

A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
44	44	43	43	17	17	14	50	50	25	25
40	40	13	13	14	14	7	41	41	24	14
50	50	30	44	49	49	6	21	23	14	29
25	25	22	30	50	50	27	14	21	29	21
		14	22	25	25		7	19	21	
			14				3	17		
			10					14		
								7		
								3		

is indicated in each with measures I and K resulting in rather narrow ranges. Similarity of distributions H and I to J and K respectively, together with the high correlations found between these measures (0.98 in both cases), suggests that indicies H and I may provide practical approximations to J and K; the latter being more direct measures of compactness but more difficult to calculate.

Endpoints of the range for indices H through K agree with those indicated by indices based on parameters of related circles (C through G). Only the least compact unit, unit 8, is the same across measures from all four categories. A look at values in the middle of the range indicates even less agreement. The median four units were selected from the range produced by each index. With tied values included, there were actually between four and nine units selected in each case (Table 3). Although some agreement is apparent, particularly within categories, considerable variation exists in units identified as being of "average" compactness. Greatest agreement is found for unit 14; included by nine of the eleven indices. Two other units (25 and 50) were in the median range for slightly over half of the indices, each being selected six times, but only one additional unit was selected in more than three cases.

The potential for variability in interpretation identified is obviously undesireable. If compactness is to continue to be used as an analytical as well as a descriptive geographic variable, it is necessary to determine the relative accuracy with which each measure represents compactness and to employ only those measures having sufficient accuracy for a particular application.

It can be anticipated that the indices based on dispersion of elements of area will provide the most accurate measures of compactness due to their consideration of the unit as a whole. The question becomes how much better are these measures than the other more easily calculated measures that have had considerable use in geographic analysis?

Based on comparisons above, one hypothesis addressed here is that measures H and I, using direct comparison to related circles, approximate measures J and K and will exhibit a comparable level of accuracy in measuring compactness. Due to differences demonstrated, it is further hypothesized that indices based on parameters of related circles and perimeter-area measurement will be significantly less accurate in representing compactness, with the perimeter-area measures being the least effective.

Evaluation of indices

Although a number of authors have discussed relative advantages and disadvantages of specific pairs or groups of compactness indices (e.g., Frolov, 1975; Whittington, Beavon, and Mabin, 1972) little attention has been given to their systematic evaluation. Evaluation that has been done has not been directed to how well they measure compactness. Griffith (1982), for example, compared indices C, F, G, I, and J in terms of their correspondence to the geometry of origins and destinations. With the exception of G, all exhibited an identifiable relationship, although C, F, and J seemed to be somewhat redundant in their contribution.

In what may be the only attempt to actually evaluate the extent to which indices represent compactness, Kimerling, Morrison, and Yatebe (1973) evaluated indices E, F, G, H, and I in terms of their correspondence to visual judgement of compactness. On this basis, indices F and H were judged the best. The approach, however, does not determine the accuracy with which an index represents compactness, but the ability of the index to predict perception.

Method of evaluation

From a geographic perspective, the importance of compactness is most often related to various enumeration unit characteristics. Compactness is often considered to be indicative of homogeneity within units. The more compact a unit is, the shorter the average distance between any two locations, therefore, the more similar characteristics of those locations are likely to be. Compactness, as a result, has been used as a representation of homogeneity for such applications as political

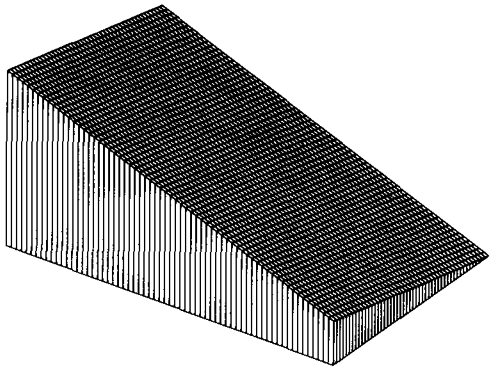

Fig. 9. Test distribution.

districting and evaluation of the potential accuracy of maps based on aggregate values derived for enumeration units.

As pointed out by Massam and Goodchild (1971), however, compactness of the geographic boundaries of a unit does not necessarily correspond to the compactness of individual distributions within that unit. They, therefore, proposed a modification of index J which includes weighting factors corresponding to the specific variable under consideration.

Recognition of the effect of distribution variability on the degree to which compactness measures homogeneity, suggests a method by which compactness indices can be evaluated. Given a distribution that varied at a fixed rate, the homogeneity of the distribution within any unit placed on that distribution would be a function of size, compactness, and orientation of the unit. If size and orientation are controlled for, homogeneity within a unit will be a direct function of compactness. This principle is employed here in evaluating the eleven indices under consideration.

A distribution was generated having a continuous linear trend along one diagonal (Figure 9). The surface was represented by a matrix of z-values, 112 rows by 75 columns 1/10th of an inch apart. All sample units were rescaled to identical areas of 1 sq. in. Each was then mathematically superimposed on the grid matrix representing the distribution. Orientation and location were random and the procedure was repeated thirty times for each unit. These repetitions served to eliminate any potential influence of orientation in relation to the direction of the distribution.

Points of the grid matrix within a unit at a given position were determined. The z-values associated with these points were summed and a mean value calculated. The standard deviation of values around this mean represents the variability of values within a unit or the extent to which the distribution is not homogenous within that unit. The mean standard deviation for the thirty placements of that unit was calculated as a measure of lack of homogeneity for the unit. With size of units as well as surface variability constant and unit orientation controlled for, the unit's compactness should determine the mean standard deviation for the unit. The accuracy of an index as a representation of compactness can, therefore, be measured in terms of its correlation with these mean standard deviation values. The closer the correlation is to -1.0, the more accurate the measure of compactness.

Analysis and results

Pearson correlation coefficients were calculated between mean standard deviations and the compactness values obtained by each index (Table 4). As hypothesized, indices J and K, the two indices directly measuring dispersion of elements of area, exhibit the highest correlations (0.92 and 0.93 respectively). From a theoretical perspective, the correlations would be expected to be closer to one. The fact that they are not can be attributed to representation of a continuous surface as a discrete matrix of z-values. This results in variation in the number of locations included in standard deviation calculations of 93 to 104.

Having correlations with mean standard deviation of 0.91 and 0.90 respectively, indices H and I are, as hypothesized, comparable to indices J and K. Correlations for indices based on comparison to individual parameters of related circles, on the other hand range from 0.74 to 0.85. They are thus considerably less effective measures of compactness, with index D (having a correlation of 0.85) the best measure from this category. As hypothesized, indices A and B, based on perimeter-area measurement, provide the poorest representations of compactness with correlations to mean standard deviation of only 0.70 and 0.68 respectively.

Examination of index values for four sample units spanning the compactness range illustrates the nature of errors resulting for indices based on individual shape characteristics (Figure 10). Units selected are relatively evenly spaced according to indices J and K, the most accurate measures.

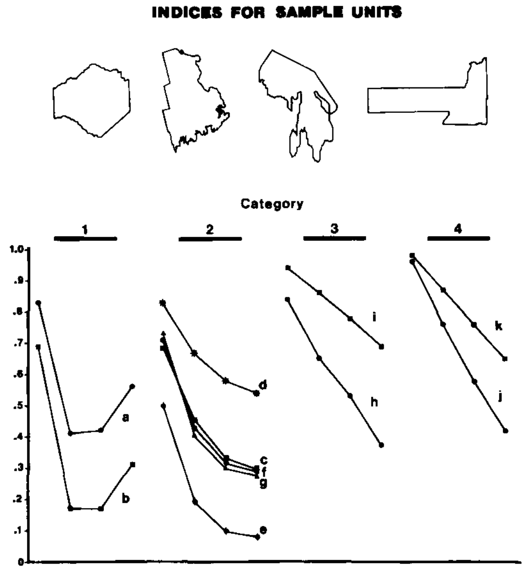

Fig. 10. Plot of index values for selected counties.

Through this somewhat diverse range of shapes, indices H and I result in quite similar linear trends although actual value ranges differ considerably. Indices C through G, however, being based on circumscribed or inscribed circles, are overly sensitive to indentations and extrusions. The problem is apparent for units two and five and results in lower than warrented compactness values.

Use of perimeter as a major component of a compactness index overemphasizes the shape of the unit's boundary. The extreme influence of this factor is readily apparent in values obtained by indices A and B for units two and five. Again values indicate that units are less compact than is actually the case.

It can thus be concluded that similar units may be misrepresented by both perimeter-area based measures and measures using individual parameters of related circles. Those measures using perimeter, however, are more sensitive to boundary shape and will generally produce a much less adequate measure of compactness. Schwartzberg's (1960) contention that these measures are suitable only for characterizing indentation of borders is clearly substantiated.

At the other end of the spectrum, indices based on measurement of dispersion of elements of area must, on both empirical and theoretical grounds, be considered the most accurate measures of compactness. While index K (used by Blair and Biss, 1967) actually had the highest correlation with mean standard deviation values, its effective range is quite narrow. The square of this index, index J (suggested by Bachi, 1973 and by Massam and Goodchild, 1972 among others, provides a clearer distinction among units as seen previously in the frequency graphs produced for each (Fig. 9 J and K). Index J, therefore, can be considered the most suitable measure of compactness.

If computer facilities and a digitizer are not available, indices J and K are exceedingly cumbersome to calculate (see Coulson, 1978 for one suggested method). In these situations, measures H (developed by Lee and Sallee, 1970) and I (the Boyce-Clark index modified to range from zero to one), afford quite acceptable alternatives. Index I, however, produces the narrowest effective range of any index evaluated and, therefore, may, for many purposes, not result in adequate differentiation among units. This difficulty could be overcome to some extent by squaring values produced. A problem not addressed here, however, would remain; a questionable applicability to fragmented units. On the basis of these drawbacks to I, index H must be considered somewhat superior although it is slightly more difficult to calculate by hand.

For descriptive rather than analytical purposes, a less accurate measure of compactness may be justifiable on the basis of ease of calculation. In such a case, index D (suggested by Schumm, 1956) would seem to be the most appropriate in terms of accuracy, and requires measuring only the area and longest axis of a shape.

As a final illustration of the potential for misinterpretation when index D or any other of the simple measures are used, however, plots of all values for each index were produced. For these

Table 4. Correlation Between Mean Standard Deviation and Compactness.

Index	Standard Deviation
A	–0.70
B	–0.68
C	–0.75
D	–0.85
E	–0.75
F	–0.81
G	–0.74
H	–0.91
I	–0.90
J	–0.92
K	–0.93

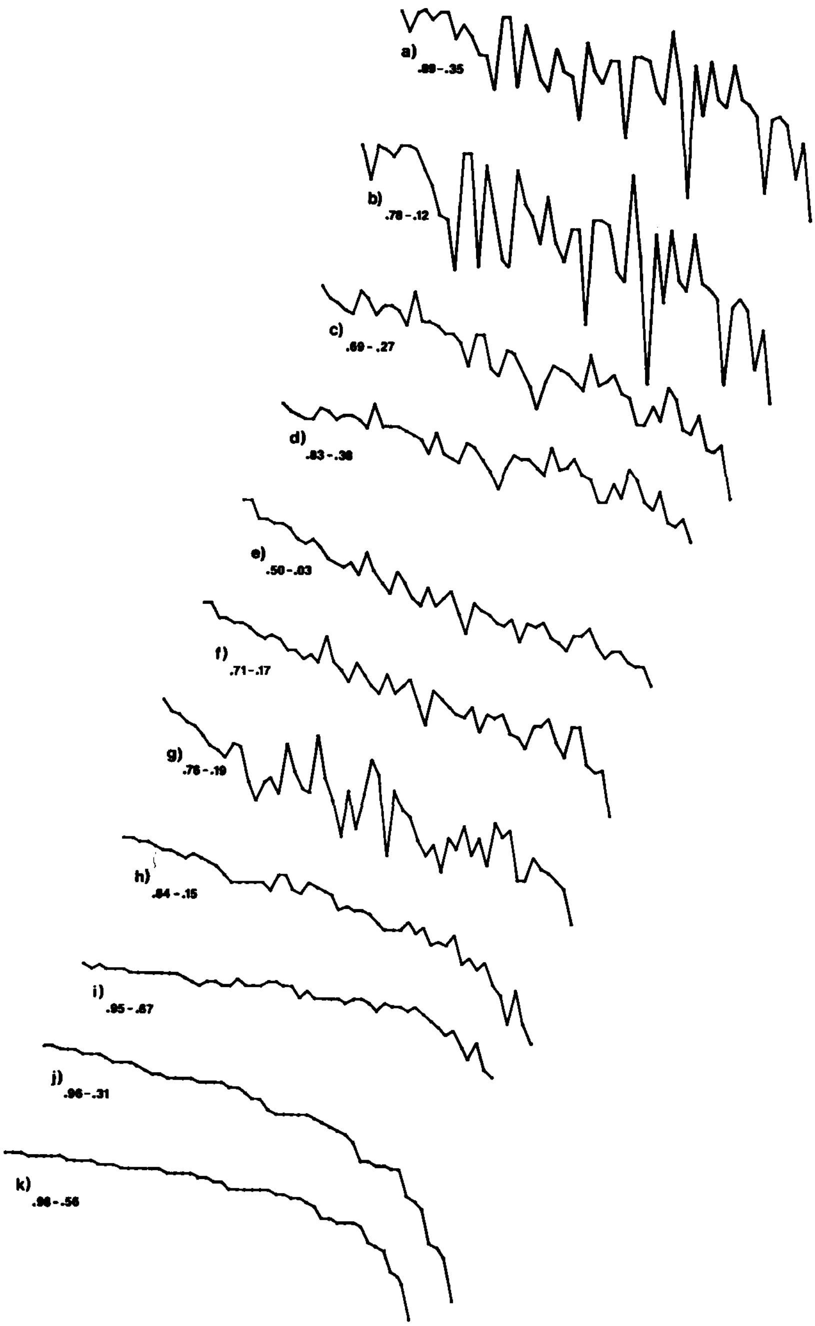

Fig. 11. Plots of compactness values on each index arranged in rank order according to indices J and K.

plots, units are arranged in rank order as determined for indices J and K. Unit values were plotted in this order for all indices (Figures 11 A-K). The extent of variation represented reinforces conclusions made thus far. Deviations from the general downward trend presented for J and K can be interpreted as an indication of the extent to which compactness is misrepresented by each of the other indices. It can be seen, for example, that while index D produces a similar overall pattern, rank ordering of units will be much different than for index J.

An interesting feature of plots J and K is their general shape. They indicate that, for U.S. counties, a large percentage of units are relatively compact with a lesser number of non-compact units. Comparison of this plot to plots of political districts, which are much more subject to gerrymandering, could be one method of indicating the extent to which gerrymandering is a significant problem in specific cases.

Summary and conclusions

Compactness can probably be considered the single most important aspect of geographic shapes. The proliferation of compactness measures in the geographic literature certainly supports this contention. In the present paper four categories of compactness measures have been identified based upon: perimeter-area measurements, single parameters of related circles, direct comparison to standard shapes, and dispersion of elements of area around a center point. A review of suggested measures resulted in selection of two measures representing the first category, five representing the second, and two each representing the third and fourth categories.

All eleven compactness measures were calculated for each of 54 sample U.S. counties. The initial stage of the study was comparison of values produced by the indices in an effort to identify similarities and differences in compactness descriptions obtained. Comparison was based on correlations among the measures, analysis of frequency distributions, and examination of the extent to which indices agreed on counties identified as being of high, median, and low compactness.

While all indices exhibited some degree of correspondence, with correlations ranging from 0.70 to 0.99, substantial differences were identified as well. As expected, indices within each category produced the most similar results with the largest differences occurring between indices in categories one and four. These categories could be characterized as using the most limited and the most complete parameters in their respective calculations. All indices did agree on the least compact shape and indices in categories two, three, and four were largely in agreement on the most compact unit, with two units ranked first or second on all indices in these categories. Considerable differences were, however, identified in the intermediate ranges. This resulted in very different rank orderings of units as well as a variety of general representations of the compactness range as illustrated by the frequency distributions (Figures 8 a-k).

Following an analysis of variability among indices, the relative accuracy with which each index represents compactness and the nature of errors that do occur were examined. The two indices measuring dispersion of elements of area around a unit's centroid were judged to be most accurate. Index J, the relative distance variance, having a greater range of values than K, its square root, can be considered the more suitable measure. The two indices based on direct comparison to related circles, H and I developed by Lee and Salle (1970) and Boyce and Clark (1964) respectively, were found to be close approximations to J an K.

With the possible exception of index D, developed by Schumm (1965), indices based on perimeter-area measurement and on individual parameters of related circles are found to be of insufficient accuracy in representing compactness to warrent their continued use for this purpose. Index D, (the radius of the circle of the same area divided by the radius of the circumscribing circle), is not as accurate as the measures cited above. It can, however, be calculated easily by hand, results in the most accurate compactness representation of indices in the first two categories, and (of indices in categories one and two) produces a frequency distribution most similar to that of index J.

Results of the present study have applicability to evaluation and reinterpretation of past research in which compactness was a major factor as well as to selection of a compactness index for future geographic analysis. In relation to past applications, for example, compactness has long been identified as a significant variable for analysis in political geography.

For political geographic analysis at a national or international level, Pounds (1972, pp. 54–55) identified compactness of countries as second only to size in geographic significance. This contention is based on the relationship of compactness to such factors as ease of travel, communication efficiency, and homogeneity of the population. According to Pounds (1972, pp. 55), "the only possible measure" of compactness is the perimeter-area measure presented here as index A. Present results indicate that there is little or no justification for this conclusion. Pounds seems to consider length of border that must be defended as an important variable which he confuses with compactness of the country.

On a more local scale, compactness has long been seen as a factor to be considered in formation of political districts. Hess (1967) among others, was an early advocate of measures based on dispersion of elements of area for application in districting problems. The present study certainly supports these measures as being the most accurate. The ability to include weighting factors reflecting population distribution characteristics as suggested by Massam and Goodchild (1971) would seem to make them even more suitable.

Morrill (1981), however, has recently pointed out that compactness may have been given too high a priority in many districting problems. He, therefore, advocates use of a simpler measure to reduce the perceived importance of district compactness. If there is need for a more easily calculated measure and a limited range of error is deemed acceptable, it seems that the most accurate measure within these criteria would still be advisable. Index D, therefore, might be appropriate to the districting problem rather than index E, the alternative suggested by Morrill.

Results of the present study may have somewhat different implications in relation to other geographic investigations in which compactness is a factor. In examining accuracy of values aggregated to political enumeration units and the subsequent accuracy of maps derived from these values, compactness is considered a critial factor (e.g., Coulson, 1978; Fairchild, 1981). In contrast to applications in political districting, it is not enough to identify units that are seriously outside specified compactness limits. For assessing data accuracy of enumeration units, compactness is used as a predictive variable that must be measured precisely if it is to be used at all. This application, therefore, requires selection of the most accurate compactness measure. Index J would be the obvious choice.

Fairchild (1981) in her study of isopleth map accuracy used the Boyce-Clark index I, rather than index J, to measure compactness. While this index has proved to be fairly accurate, it has an extremely limited effective range. The lack of a strong relationship between unit compactness and isopleth accuracy in Fairchild's study may, therefore, be the result of the index selected rather than the lack of a definite relationship.

In general, results of the present study should be viewed as a guide for evaluating previous applications of compactness indices and for selection of indices for future applications. Although examples cited above identify the most accurate compactness index, results do not necessarily support the use of this index in all cases. Also, the fact that the indices based on perimeter-area measurement and individual parameters of related circles are considerably less accurate representations of compactness than those based on dispersion of elements of area does not mean they are of no use in geographic investigation. There may well be cases in which irregularity and length of a political border or dissection of a political unit may be of more significance than the unit's compactness.

In addition to providing a basis for selection of the most appropriate measure of compactness for a particular problem, the review of compactness presented points to a lack of precise definition of spatial parameters of significance in geographic analysis. Shape has traditionally been used in geography as a single descriptive variable. As a result, no clear distinction has been made among characteristics of shape such as compactness, elongation, symmetry, etc. This lack of distinction has carried over to mathematical description of geographic shapes, resulting in confusion concerning the appropriate use of the various measures that have been developed. An effort, therefore, is needed to more specifically identify the characteristics of geographic shape that are relevant to particular problems and to determine appropriate methods for the measurement of each.

References

Attneave, F, 1957: Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology,* 53: 221–227.

Attneave, F. and *Arnoult, M. D.,* 1956: The quantitative study of shape and pattern perception. *Psychological Bulletin,* 53, 6: 452–471.

Austin, Robert, F., 1981: The shape of West Malaysia's districts. *Area,* 12, 2: 145–150.

Bachi, Roberto, 1973: Geostatistical Analysis of Territories. *Bulletin: International Statistical Institute* (Proceedings of the 39th Session), 45, Book 1: 121–131.

Bitterman, M. E., Krauskopf, J. and *Hochberg, J. E.,* 1954: Threshold for visual form: a diffusion model. *American Journal of Psychology,* 67: 205–219.

Blair, D. J. and *Bliss, T. H.,* 1967: The Measurement of shape in geography: an appraisal of methods and techniques. *Bulletin of Quantitative Data for Geographers,* 11, 45 pp.

Boyce, R. R. and *Clark, W. A. V.,* 1964: The concept of shape in geography. *The Geographical Review,* 54: 561–572.

Bunge, W., 1966: *Theoretical Geography,* Lund Studies in Geography, Series C, General and Mathematical Geography, no. 1, Lund: C. W. K. Gleerup, 289 pp.

Cerny, J. W., 1975: Sensitivity analysis of the Boyce-Clark shape index. *Canadian Cartographer,* 12: 21–27.

Coulson, Michael, R. C., 1978: Potential for variation, a concept for measuring the significance of variation in size and shape of areal units. *Geografiska Annaler,* 60B: 48–64.

Dent, B. D., 1972: A note on the importance of shape in cartogram construction. *Journal of Geography,* 71: 393–401.

Duda, Richard, O. and *Hart, Peter, E.,* 1973: Chapter 9, Descriptions of Line and Shape, in *Pattern Analysis and Scene Analysis,* John Wiley & Sons, New York, 327–378.

Fairchild, Dana, 1981: The effect of enumeration unit shape on isopleth map accuracy. *Technical Papers,* American Congress on Surveying and Mapping, 41 st Annual Meeting, Washington, D. C., 423–432.

Frolov, Y., 1975: Measuring of shape of geographical phenomena: a history of the issue. *Soviet Geography: Review and Translation,* 16: 676–687.

Gibbs, Jack, P., 1961: A method for comparing the spatial shapes of urban units, in Jack P. Gibbs, *Urban Research Methods,* D. Van Nostrand Company, Inc., Princeton, New Jersey, 99–106.

Griffith, Daniel, A., 1982: Geometry and spatial interaction. *Annals of the Association of American Geographers,* 72, 3:332–246.

Haggett, P., 1966: *Locational Analysis in Human Geography,* St. Martin's Press, New York, 339 pp.

Hess, S. W., 1969: Compactness-What shape and size? *Conflicts Among Possible Criteria for Rational Districting,* National Municipal League: 15–23.

Hsu, Mei-Ling, and *Robinson, Arthur, H.,* 1970: *The Fidelity of Isopleth Maps: An Experimental Study,* University of Minnesota Press, Minneapolis, Minnesota, 92 pp.

Johnston, R. J. and *Rossiter, D. J.,* 1981: Shape and the definition of parliamentary constituencies. *Urban Studies,* 18: 219–223.

Kimerling, A. J., Morrison, J. L. and *Yatabe, P. M.,* 1973: Compactness in legislative redistricting, a comparative study of several shape measures. *Proceedings of the American Congress on Surveying and Mapping,* 33 rd Annual Meeting, Washington, D.C.: 275–289.

Lee, D. R. and *Sallee, G. T.,* 1970: A method of measuring shape. *The Geographical Review,*60, 4: 555–563.

Lo, C. P., 1980: Changes in the shapes of Chinese cities, 1934–1974. *The Professional Geographer,* 32, 2: 173–183.

Massam, Bryan, H. and *Michael F. Goodchild,* 1971: Temporal trends in the spatial organization of a service agency. *Canadian Geographer,* 15, 3: 193–206.

Miller, V. C., 1953: *A. Quantitative Geomorphic Study of Drainage Basin Characteristics in the Clinch Mountain Area, Virginia and Tennessee,* Columbia University, Department of Geology, Technical Report No. 3, New York, 125 pp.

Moellering, Harold, and *Rayner, John, N.,* 1981: The harmonic analysis of spatial shapes using dual axis fourier shape analysis (DAFSA). *Geographical Analysis,* 13, 1: 64–77.

Morrill, Richard, L., 1981: *Political Redistricting and Geographic Theory,* Resource Publications in Geography of the Asciation of American Geographers, 76 pp.

Pounds, Norman, J. G., 1972: *Political Geography,* 2nd edition, McGraw-Hill Book Company, Inc., New York, 453 pp.

Reock, E. C., 1961: Measuring compactness as a requirement of legislative appointment. *Midwest Journal of Political Science,* 5, 1.: 70–74.

Rimbert, C. Cauvin, 1976: *La Lecture Numerique des Cartes Thematiques, Les method de la cartographic thematique fascicule 1,* Laboratoire de Cartographic Thematique de l'ERA 214, Strasbourg, Institut de Geographie. 172 pp.

Sampson, Robert, J., 1975: *Surface II Graphics System,* Kansas Geological Survey, Lawrence, Kansas, 240 pp.

Sanders, Ralph, A. and *Porter, Philip, W.,* 1974: Shape in revealed mental maps. *Annals of the Association of American Geographers,* 64, 2: 258–267.

Schumm, S. N., 1963: Sinuosity of alluvial rivers on the Great Plains. *Bulletin of the Geological Society of America,* 74: 1089–1100.

Schwartzberg, J., 1966: Reapportionment, gerrymandering, and the notion of compactness. *Minnesota Law Review,* 50: 443–457.

Stoddart, D. R., 1965: The shape of atolls. *Marine Geology,* 3: 368–383.

Whittington, G., Beavon, K. S. O. and *Mabin, A. S.,* 1972: Compactness of shape: review, theory, and application. *Environmental Studies,* Occasional Paper No. 7, of the Department of Geography and Environmental Studies, University of the Witwatersrand, Johannesburg, 40 pp.

Wilkins, C. A. and *Shaw, J.,* 1971: Measures of shape distortion in urban geography. *The Australian Geographer,* 11, 6: 593–595.

Zusne, Leonard, 1965: Moments of area and of the perimeter of visual forms as predictions of discrimination performance. *Journal of Experimental Psychology,* 69, 3: 213–220.