CENTROGRAPHIC MEASURES IN GEOGRAPHY

Aharon Kellerman

CENTROGRAPHIC MEASURES IN GEOGRAPHY Aharon Kellerman

ISBN 0 86094 091 8
ISSN 0306 - 6142

CATMOG

(Concepts and Techniques in Modern Geography)

CATMOG has been created to fill a teaching need in the field of quantitative methods in undergraduate geography courses. These texts are admirable guides for the teachers, yet cheap enough for student purchase as the basis of classwork. Each book is written by an author currently working with the technique or concept he describes.

1. An introduction to Markov chain analysis - L. Collins
2. Distance decay in spatial interactions - P.J. Taylor
3. Understanding canonical correlation analysis - D. Clark
4. Some theoretical and applied aspects of spatial interaction shopping models - S. Openshaw
5. An introduction to trend surface analysis - D. Unwin
6. Classification in geography - R.J. Johnston
7. An introduction to factor analytical techniques - J.B.Goddard & A.Kirby
8. Principal components analysis - S. Daultrey
9. Causal inferences from dichotomous variables - N. Davidson
10. Introduction to the use of logit models in geography - N. Wrigley
11. Linear programming: elementary geographical applications of the transportation problem - A. Hay
12. An introduction to quadrat analysis - R.W. Thomas
13. An introduction to time-geography - N.J. Thrift
14. An introduction to graph theoretical methods in geography - K.J. Tinkler
15. Linear regression in geography - R. Ferguson
16. Probability surface mapping. An introduction with examples and Fortran programs - N. Wrigley
17. Sampling methods for geographical research - C. Dixon & B. Leach
18. Questionnaires and interviews in geographical research - C. Dixon & B. Leach
19. Analysis of frequency distributions - V. Gardiner & G. Gardiner
20. Analysis of covariance and comparison of regression lines - J. Silk
21. An introduction to the use of simultaneous-equation regression analysis in geography - D. Todd
22. Transfer function modelling: relationship between time series variables - Pong-wai Lai
23. Stochastic processes in one-dimensional series: an introduction - K.S. Richards
24. Linear programming: the Simplex method with geographical applications - J. E. Killen

Continued on inside back cover

CONCEPTS AND TECHNIQUES IN MODERN GEOGRAPHY No. 32

CENTROGRAPHIC MEASURES IN GEOGRAPHY

by

Dr. Aharon Kellerman

(University of Haifa, Haifa, Israel 31999)

CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

Acknowledgement

I wish to thank Dr. Stanley Waterman for his valuable comments and programming efforts.

I AN OVERVIEW

Centrography and statistics

Centrography consists of a set of measures and indices for thedes cription and analysis of geographical data (defined as points or areas l ocated in a spatial system), equivalent to similar measures and indices in other branches of statistics dealing with non-geographical sets of data. The uniqueness of the geographical series is its multi-dimensional structure: in its simplest form it will consist of locational information only, and will usually be presented as a map with several points (the series). For the purpose of statistical analysis, however, this simple map is complex from the outset, since each point has two (X,Y) coordinates, for latitude and longitude respectively, so that we actually have two series, or a so called bi-variate series (Fig. 1). Most of our geographical data sets or maps present even further complications, as we attach a size or quality to each of the points, (usually marked 'p' for population, or 'w' for weight). Thus we are dealing with a three-dimensional series. This might become even more complex when altitude ('z') of each point is added as a third or fourth dimension, and when 'p' is computed for several time-periods, thus adding a temporal dimension ('t') to the geographical data set.

Centrography attempts to provide the geographer with equivalents to some of the most basic tools of uni-variate statistics for the description and analysis of spatial data. These include measures of central tendency equivalent to the mode, mean and median, and measures of dispersion such as the variance, standard deviation and analysis of variance transformed to two and three dimensions. In addition to the special equations and models required for the geographical series, centrography is concerned with the unique expression and presentation of geographical data namely the use of maps. Attempts have been made, therefore, to make it possible to draw on the analyzed map most of the centrographic measures and indices in addition to their numerical values.

The potential model has often been considered to be part of centrography (Neft, 1966; Bachi, 1966). This model will not be reviewed here since it has developed separately from centrography, as part of the so-called 'social physics'. Also, in contrast to the other centrographic models, the potential model requires some socio-economic assumptions. The reader interested in the potential model as part of centrography is referred to Neft (1966) and Yam (1968).

Like most branches of statistics, centrography uses a macro approach by trying to present a general picture of the series in one value using all data points. It is also synthetic in nature since it produces measures which are not part of the data itself (Bachi, 1966). One may view centrography as a special branch of general bi-variate statistical theory. As Neft (1966, p.7) has noted, however, there are some differences between the general bi-variate and the geographical cases among which are, for example, the arbitrariness of choice of the coordinate scales in a geographical series (where there is an absolute location for each point) compared to the fixed X and Y coordinates in general bi-variate problems. Also, in the geographical series no assumption is made on the dependence of X on Y and vice-versa while this is a major aspect of general bi-variate theory.

The term 'centrography' for this branch of statistics has been in use for many years (Sviatlovsky and Eells, 1937). The term has some misleading connotations which point to the analysis of centrality only, rather than centrality and dispersion together. Another term for descriptive spatial statistics is 'geostatistics', originally proposed by Hart (1954) followed by Bachi (1962, 1966). Unfortunately, however, the term geostatistics has also become associated with statistical models in geology so that the term centrography will be used throughout this CATMOG.

(ii) Early developments

Published work on centrography dates back as early as 1872 when Hilgard computed the centre of population for the U.S. and started a series of publications using the mean and median centres for the geographical distribution of the U.S. population. These contributions as well as similar other early empirical work done in the Soviet Union, Italy, Germany and Japan are reviewed in the extensive appendix ('Historical résumé') of Sviatlovsky and Eells's (1937) article, while other work is reviewed by Stewart and Warntz (1958). Of all these the most significant was that by the Russian group between 1887-1937, with a peak during the twenties and early thirties. The Russian work, summarized by Poulsen (1959) and Porter (1963), culminated with the establishment of the Mendeleev Centrographic Laboratory in Leningrad in 1925, which was named after a famous Russian chemist who became interested in demography and published his work on centrography in 1906. Led by Sviatlovsky, the Russian centrographers, did not only compute centres of populations but went one step further calculating centres of agricultural crops, natural resources and so on, from which they arrived at conclusions about areas which best fit various agricultural and industrial products. Their work on the optimal regions for grain production contradicted strongly the national Russian goals of geographical expansion of grain production so that the Mendeleev Centrographical Laboratory was closed in December 1934.

There are two interesting characteristics in these early developments. First, there was a relatively rapid international diffusion of the idea of measuring centrality, mainly for countries and regions. This rapid diffusion of centrality studies came before indices of dispersion had been developed, thus making these studies only partially significant. Second, the use of centrography for planning purposes in Russia presents a very early example of 'applied geography'. Unfortunately this application was overly simplistic and, therefore, unacceptable.

(iii) Theoretical developments

The fall of Russian centrography coincided with the independently evolving attempts to formulate dispersion measures to complement centrality indices. Linders in Germany and Italy in 1931, and Lefever (1926) followed by Furfey (1927) in the U.S. developed independently from one another identical centrographic equivalents of the standard deviation. After an interval of a generation a new concern for centrography emerged with the concern for quantitative methods among geographers. Stewart and Warntz (1958) developed the standard distance deviation, as a measure of dispersion, followed by Warntz and Neft (1960) who suggested viewing the point of maximum potential as the harmonic centre. Porter (1963) and Court (1964) discussed various solutions for the median centre problem. Bachi (1962) developed the standard distance, and later (1966) proposed a centrographic analysis of the division of a territory. Neft (1966) summarized many of the centrographic models (excluding the cross of dispersion and sub-division of territories), and provided a useful discussion of inferential statistics and the geographical series. Later, Bachi and Kahn (1976) showed the possible application of centrographic models into the study of probability functions, and Ever-Hadani (1980) has provided an algorithm for the optimal division of territories.

(iv) Centrography and geography

The development of a fully adequate theory of regional analysis by statistical means is one of the outstanding problems of geography today. The map alone is not enough, nor are statistical data alone. They must be brought together. This is the aim of centrography' (Sviatlovsky and Eells, 1937, p. 240). This early call for 'spatial statistics' has not been met by main-stream quantitative geography. Not only has theoretical development of centrography slowed down since the mid-sixties, but their teaching and use have also been limited. Of the many textbooks on quantitative methods in geography published during the last twenty years only a few mention centrography at all, usually restricting their discussion to the mean centre, sometimes adding the standard distance as well (e.g. Yeates, 1974). Two recent texts (Taylor, 1977; Silk, 1979) provide examples and exercises which will be commented upon later in this volume, and Taylor (1977) provides also an historical overview of centrography. Only one text (Ebdon, 1977) provides a detailed discussion of the cross of dispersion, noting the lack of techniques for handling spatial data, and the lack of publicity given to existing techniques. Neft's monograph (1966) which covers many of the centrographic models has not become a standard source for texts on statistics for geographers, perhaps because of its more technical flavour. This situation reflects itself in the use of centrography in empirical geographical studies. The relatively many sub-fields of applications discussed in section five of this CATMOG might be somewhat misleading since in comparison to geographical studies that have used quantitative methods in general, relatively few have been referred to. The situation resembles, therefore, a vicious cycle between a low demand for centrography by researchers and a short supply of full and reliable discussions of the various measures and possible applications.

An exception to this is in the Israeli geographical community which benefited from Roberto Bachi's contributions as a researcher in the theory of centrography (1962; 1966; 1976), and as a teacher of a postgraduate course in this field for geography and statistics students at the Hebrew University. His text, (1966, edited by Y. Yam), includes a discussion of most of the techniques and a variety of demonstrations from several countries, to which a handful of exercises are added. Unfortunately, this text has never been translated from the Hebrew, though it includes an English abstract. Some of Bachi's students in the Department of Statistics of the Hebrew University developed further theoretical aspects of centrography (Shimoni, 1963; Yam, 1969; Ever Hadani, 1975), while the geographers applied the methods in several areas (e.g. Shachar, 1967).

There are several possible reasons for this generally rather disappointing situation. Whilst in the past geographers were more concerned with the 'micro' approach searching for idiographic patterns and leaving the theoretical developments of centrography for statisticians and sociologists, within the single decade of the 1960's they have become involved in the development and use of more sophisticated and 'elegant' statistical models. The many early empirical works and the Russian failure in the application of centrography may in themselves have caused centrography to be ignored. Another reason, noted already by Ebdon (1977) is the cumbersome computation required by some of the models. These two reasons, it is believed, have by now lost much of their relevance.

The last decade has made geographers aware of many 'do's' and 'don'ts' in quantitative geography. On the other hand, our arsenal of spatial statistics is still very limited and the need for basic and applied analysis of geographical data has not changed. The computations and precise mapping efforts required in centrography have disappeared with the rapid diffusion of mini-computers provided with computer graphics which enable the user to receive instant numerical and visual results for any centrographic problem.

II MEASURES OF CENTRALITY

The measures of centrality for a geographical series are extensions of those used for the uni-variate series. As a rule, only the mean centre and the median centre have been used. Use of the modal centre and the harmonic centre has been proposed by Warntz and Neft (1960), and the problems connected with their definitions and computations have been discussed by Bachi (1966).

(i) The mean centre

Assuming that for each point i we have two coordinates X_i and Y_i, (which can be defined using an arbitrary but consistent system) then two uni-variate means must be calculated in order to achieve the two coordinates for the mean of the geographical series, the so called 'mean-centre'. Put another way

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \quad ; \quad \bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} \qquad (1)$$

where n is the number of cases (Fig. 1). This unweighted two-variate mean is sometimes called the 'centroid' (Table 1).

If the points have a certain 'weight' attributed to each (such as the population of each place or point) then the mean centre has to reflect an averaging of these weights as well,

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i P_i}{\sum_{i=1}^{n} P_i} \quad ; \quad \bar{Y} = \frac{\sum_{i=1}^{n} Y_i P_i}{\sum_{i=1}^{n} P_i} \qquad (2)$$

where P stands for population or weight (Table 2).

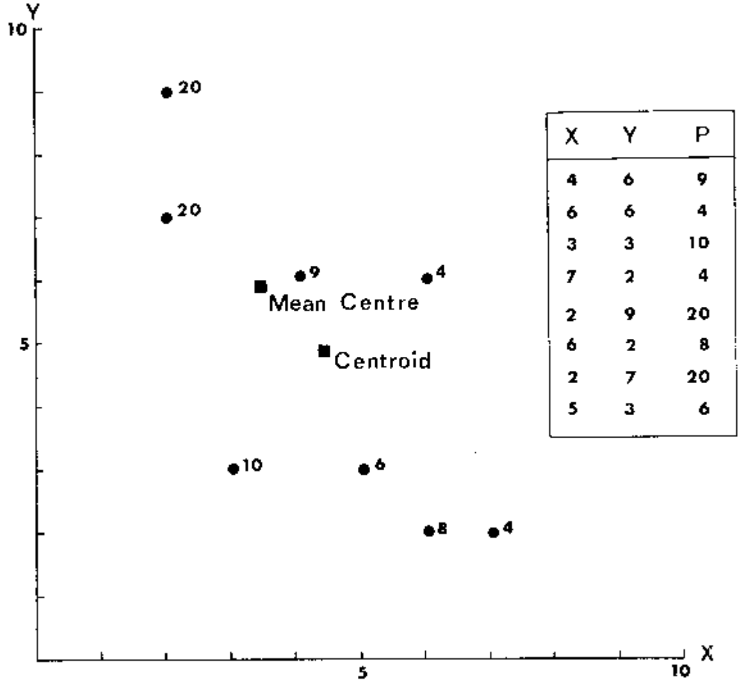


Figure 1. Central points: The Centroid and the Mean Centre

Another method of calculation which can also be of much help in the computation of measures of dispersion calls for the use of relative rather than absolute weights, so that the relative weight of each point is

$$f_i = \frac{P_i}{\sum_{i=1}^{n} P_i} \qquad (3)$$

and, therefore

$$\bar{X} = \sum_{i=1}^{n} X_i f_i \quad ; \quad \bar{Y} = \sum_{i=1}^{n} Y_i f_i \qquad (4)$$

Sometimes the geographical data consist of data spread over continuous regions rather than values at discrete points. In this case we shall have, first, to transform the regions into points. This can be done by various methods such as using the geographical centre of the region as a representative point or using the largest city in the region as a regional point to which the regional weight (however defined) is attributed.

The calculation of mean centres for very large areas requires an adjustment to the sphericity of the earth's surface. This was first calculated by Mendeleev's son (Neft, 1966, p. 28) and the formulae were formally presented by Shimoni (1963).

Table 1. Computation of the Unweighted Mean Centre (Centroid)

Coordinates X	Y
4	6
6	6
3	3
7	2
2	9
6	2
2	7
5	3

$$n = 8$$

$$\sum_{i=1}^{n} X_i = 35$$

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{35}{8} = 4.375$$

$$\sum_{i=1}^{n} Y_i = 38$$

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{38}{8} = 4.75$$

(See Fig. 1)

(ii) Characteristics of the mean centre

The mean centre, as Ebdon (1977) and Silk (1979) have noted, 'represents an average location, not an average of the characteristics of the phenomena to be found at that location' (Silk, 1979, p. 24), (assuming that the location of the mean centre has no weight). This 'average location', presented in the form of one point, provides the researcher with the essence of what can sometimes be a lengthy list of points which comprise the geographical series. This enables the researcher to compare easily several series on a temporal (changes in the same area over time) or inter-regional (by comparing two or more regions) basis. On the other hand, however, the mean centre has no meaning when presented in a numerical form by its two coordinates. It makes sense only when presented graphically on the map relative to the points of the original geographical series. The location of the mean centre is 'synthetic' in that sense that it might be located in the sea when a coastal series is studied. In such a case, its locational characteristics relative to the series pattern and distribution have to be carefully studied. Another result of the 'synthetic' nature of the mean centre is the possibility that two different geographical series may yield the same mean centre. Here too, a careful study of the geographical and point-size distribution of the two series has to be undertaken, in order for the investigator to arrive at logical conclusions regarding an inter-series comparison.

In addition to these geographical qualities, the mean centre also has important statistical characteristics. Being an extension of the uni-variate simple mean it maintains special relations with its 'moments', or in other

Table 2. Computation of the Weighted Mean Centre (Centre of gravity)

Coordinates X	Y	Weight P	Weighted XP	Coordinates YP
4	6	9	36	54
6	6	4	24	24
3	3	10	30	30
7	2	4	28	8
2	9	20	40	180
6	2	8	48	16
2	7	20	30	140
5	3	6	30	18
Total		81	276	470

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i P_i}{\sum_{i=1}^{n} P_i} = \frac{276}{81} = 3.40$$

(See Fig. 1)

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i P_i}{\sum_{i=1}^{n} P_i} = \frac{470}{81} = 5.80$$

words, it minimizes the dispersion of points around it in a special form. The first moment around the mean centre equals zero; so that

$$\sum_{i=1}^{n} (X_i - \bar{X}) = \sum_{i=1}^{n} (Y_i - \bar{Y}) = 0 \qquad (5)$$

In addition, the second moment around the mean centre is minimal; or in other words, the sum of the squared distances from all points to the mean centre is minimal (when the mean centre is compared to any other point drawn on the map). This specific quality of the mean centre causes a greater importance to be assigned to remotely lying points in the determination of the location of the mean centre. (Through the squaring of distances which is done, larger and squared distances have more 'weight' than the smaller distances). This is useful for the definition of the 'standard distance' as a spatial dispersion measure, (developed by Bachi (1962) and discussed in the next section of this paper) and makes it useful in the urban and regional planning problems to be discussed in Section IV. On the other hand, this consideration of remotely lying points may sometimes conflict with the aim of providing a meaningful measure of centrality for the geographical series.

In these cases and others the median centre may be considered a more useful measure of centrality. Another statistical quality of the mean centre proved by Shimoni (1963) is its covariation in case of rotating or transposing the coordinate-system. In such cases the location of the mean centre will not change although the numerical values of its coordinates will change.

(iii) The Median Centre

This measure is the centrographic analogue of the uni-variate median. The simple median has two major characteristics. First, it defines the middle of a series so that half the cases are larger than the median while the remainder are smaller. Second, the median minimizes the first moment; the sum of deviations around the median is minimal (while the first moment around the mean equals zero). It is important to bear in mind these basic characteristics of the simple median, since when attempts are made to extend the median to the spatial two- and three-dimensional problem it becomes apparent that there is no single point for a given spatial distribution which satisfies both qualities of the uni-variate median. The term 'median centre' may refer, therefore, either to a point which defines a 'mid-point' for the studied geographical series, or it may refer to a point which minimizes the sum of distances from all points to it.

In early stages of centrography the first definition of the median centre was used. Its calculation is very simple since a uni-variate median is found for the X and Y coordinates of all points separately, and these two medians define the location of the median centre. This 'median centre' has two major drawbacks. First, it is not a unique point since any change in the coordinates-system may change its location and second, it is highly insensitive to changes in the location of points as long as these changes do not cause movements of points from one side of the median-centre to the other. This version of the median centre was used by the U.S. Bureau of the Census until the 1920's thus raising a long debate among centrographers, the details of which may be found in Sviatlovsky and Eells (1937), Neft (1966, pp. 29-33), and Porter (1963, footnote 16). It was proposed again by Hart (1954) and in the texts of Cole and King (1968) and Hammond and McCullagh (1974), but it has been rejected by most modern centrographers. Ebdon (1977, p. 111) proposed its use for preliminary studies of spatial distribution when speed is important, but it seems that, given the problem connected with the interpretation of its value, it should not be recommended for use at all.

The second version of the median centre, sometimes also called the 'bivariate median', the point of minimum aggregate distance/travel', or the centre of convergence' is, like the mean centre, invariant to changes or rotations of the coordinates-system, so that its location does not change though the numerical values of its X and Y coordinates do. The definition of this version of the median centre is that point to which the sum of distances from all points is at minimum. For the weighted case, the median centre will be the point at which the sum of weights times distance is a minimum (Warntz and Neft, 1960). The major problem with this version of the median centre is the difficulty of computation. The equation to be minimized is

$$\sum_{i=1}^{n} d_{im} = \sum_{i=1}^{n} \sqrt{(X_i - X_m)^2 + (Y_i - Y_m)^2} \qquad (6)$$

where d denotes distances, and m denotes the median centre (Table 3).

Table 3. An Iterative Solution for the Weighted Median Centre

Coordinates		Weight
X	Y	P
4	6	9
6	6	4
3	3	10
7	2	4
2	9	20
6	2	8
2	7	20
5	3	6

Initial Definitions: Median Centre $(X_m;Y_m)$

$$\min \Psi(X_m;Y_m) = \sum_{i=1}^{n} P_i \sqrt{(X_i - X)^2 + (Y_i - Y)^2}$$

$X = \overline{X}$, $Y = \overline{Y}$ Mean Centre (See Table 2)

$$P_{i1} = \frac{P_i}{\sqrt{(X_i - \overline{X}) + (Y_i - \overline{Y})}}$$

First Iteration: $\Psi_1(X_m;Y_m) = 9\sqrt{(4-3.40)^2 + (6-5.80)^2} + 4\sqrt{(6-3.40)^2 + (6-5.80)^2} + \ldots\ldots\ldots\ldots = 228.22$

Second Iteration: $P_{11} = \frac{9}{\sqrt{(4-3.40)^2 + (6-5.80)^2}} = 14.28$, $P_{21} = \frac{4}{\sqrt{(6-3.40)^2 + (6-5.80)^2}} = 1.53$ $P_{31} = \ldots\ldots$

$$\overline{X}_1 = \frac{\sum_{i=1}^{n} X_i P_{i1}}{\Sigma P_{i1}} = 3.35, \quad \overline{Y}_1 = \frac{\sum_{i=1}^{n} Y_i P_{i1}}{\sum_{i=1}^{n} P_i} = 6.04$$

$$\Psi_2(X_m;Y_m) = \sum_{i=1}^{n} P_{i1} \sqrt{(X_i - \overline{X}_1)^2 + (Y_i - \overline{Y}_1)^2} = 79.67$$

'Because of the square root, it cannot be broken into separate sums and cannot be differentiated to give direct solutions for X_m, Y_m' (Court, 1964, p.401). Manually, using direct distances, it can be solved for trivial problems using trial and error to approximate the location of the median centre. However, the problem can be programmed so that iterative solutions produced by computers may be used. An interative solution starts with a tentative definition for x_m, y_m, usually the mean centre $(\bar{x},\bar{y})$. This is then used for a first round of iteration calculating equation 6. In a second iteration the weights are redefined by some ratio of weight to distance, so that new $\bar{x},\bar{y}$ are defined and the median centre recalculated (Table 3). Further rounds of iteration may then be computed with new definitions of $\bar{x},\bar{y}$ and P. The iterative method minimized the trial and error process and enables the user to achieve a rational solution when the differences between two consecutive iterations become small. Porter (1963) proposed another procedure for the computation of the median centre using least-squares and the relationships between the regression lines of the X coordinates on the Y coordinates and vice versa. However, Court (1964) in his comment on Porter's procedure proved, that this solution is not invariant to a change or rotation of the coordinate-system, and is, therefore, unacceptable. This provides a good example of the difficulties encountered when trying to use bivariate statistics for centrographic problems.

(iv) The median centre vs. the mean centre

As mentioned earlier, the major statistical difference between the two centres is the minimization of squared distances (second moment) achieved by the mean centre, and minimization of distances (first moment) achieved by the median centre. The point of minimum aggregate travel, or the median centre could provide, therefore, an optimal location for industries as in the Weber problem (Taylor, 1977) or for such public services as schools and hospitals (Morrill and Symons, 1977). On the other hand, the mean centre, being more sensitive to peripheral points, could be used for more equitable locations of public services. Though the median centre in its second version is more sensitive to changes in the location of points when compared to the first version of the median centre, it is still far less sensitive to changes of location when compared to the mean centre. While the latter one reacts to any change, the first will not react to any change of location of a point along the straight line connecting the point and the median centre (Bachi, 1966).

Technically it is much easier to compute the mean centre rather than the median centre. It is also difficult to measure dispersion around the median centre whereas we shall discuss several measures of dispersion around the mean centre in the next section. However, Bachi (1966) proposed a simple measure of dispersion around the median centre, namely the mean distance of all points from the median centre.

Neft (1966) has, noted that among the similarities of the two centres are their covariation in case of changes of the coordination system, that neither need be located in the sphere of the studied territory, and that the location of each do not reveal anything of the qualities of that location.

III MEASURES OF DISPERSION

Just as it is almost worthless to use the mean for a uni-variate series without a complementary measure of dispersion such as the variance or standard deviation, so the use of a spatial measure of centrality is limited without a companion measure of dispersion around it. In addition to the general need for measures of dispersion for geographical series, there is also a requirement that these measures may be presented graphically, relative to the location of the series-points and relative to the location of any central point, usually the mean centre.

The general rule behind the several measures which follow this discussion is some averaging of distances from the points to the centre. This is directly analogous to averaging deviations from the mean in the uni-variate series. Of the measures presented here, the most useful is standard distance since it is a direct extension of the standard deviation of uni-variate statistics, and since it lends itself easily to graphical presentation and to the construction of indices for inter-map comparison. It is important, however, to provide a discussion of other measures as well in order to identify the statistical relations among them.

Early centrographers did not use measures of dispersion, thus restricting the significance of their work. Extensive discussions of measures of dispersion were provided only a generation later by Warntz and Neft (1960), by Neft (1966) and by Bachi (1958, 1962, 1966). It is unfortunate, however, that 'application of measures of spread are much less common than studies of central tendency in geostatistics. This is a pity, because the two statistical concepts obviously complement one another in the description of frequency distributions in two dimensions as well as one (Taylor, 1977, p.27).

(i) The mean distance

As its name implies the mean distance is the mean distance between all points, so that

$$D = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} f_i f_j d_{ij} \qquad (7)$$

and for the unweighted case

$$D = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{d_{ij}}{n} \qquad (8)$$

where j is any point other than i and d_{ij} is the distance between them (Bachi, 1966) (Table 4). On one hand this measure of dispersion is based on all points, but on the other hand, it requires much computation, and it has no analogue for the uni-variate case.

Table 4. Computation of the Mean Distance for the Weighted Case

Coordinates		Weight			
X	Y	P	f	$f_i f_j d_{ij}$	$\sum_{j=i+1}^{n} f_i f_j d_{ij}$
4	6	9	0.11	$0.11 \cdot 0.05 \cdot \sqrt{(4-6)^2 + (6-6)^2} + 0.11 \cdot 0.12 \sqrt{(4-3)^2 + (6-3)^2} + \ldots$	1.20
6	6	4	0.05		0.20
3	3	10	0.12		0.38
7	2	4	0.05		0.99
2	9	20	0.25		0.44
6	2	8	0.10		0.16
2	7	20	0.25		0.08
5	3	6	0.07		-
Total		81	1.00	$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} f_i f_j d_{ij} =$	3.45

(ii) The mean of squared distances

This measure uses squared distances among points; in other words, each distance is measured once between i and j and then again between j and i. The formula for the mean of squared distances is presented by Bachi (192, 1966) as

$$DS = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} f_i f_j d_{ij}^2} \quad (9)$$

and for the unweighted case

$$DS = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}^2} \quad (10)$$

Bachi (1962, 1966) also showed an interesting relationship between DS and the standard deviations of X and Y, a relationship which makes DS easy to calculate and compare to the standard distance. If we develop the element d_{ij}, then

$$DS^2 = \sum_{i,j}^{n} f_i f_j [(X_i - X_j)^2 + (Y_i - Y_j)^2] \quad (11)$$

If we concentrate now on the X's only, we can open the expression $(X_i - X_j)^2$ using simple algebra, so that

$$\sum_{i,j}^{n} f_i f_j (X_i - X_j)^2 = \sum_{i,j}^{n} f_i f_j X_i^2 + \sum_{i,j}^{n} f_i f_j X_j^2 - 2 \sum_{i,j}^{n} f_i f_j X_i X_j$$

$$= (\sum_i f_i)(\sum_j f_j X_j^2) + (\sum_j f_j)(\sum_i f_i X_i^2) \quad (12)$$

$$- 2(\sum_i f_i X_i)(\sum_j f_j X_j) = 2[\sum f_i X_i^2 - (\sum f_i X_i)^2]$$

because the summations of the square X_i's and X_j's are identical, and because their multiplying each by the other means squaring one of both.

The last equation is twice the variance of X $(2\sigma_X^2)$ since we substracted the squares of all X's from the square mean. If we repeat the same procedure for the Y's, we shall finally achieve

$$DS^2 = 2(\sigma_X^2 + \sigma_Y^2) \quad (13)$$

or

$$DS = \sqrt{\frac{2 \sum_{i=1}^{n} d_{ic}^2}{n}} \quad (14)$$

where c is the mean centre and

$$DS = \sqrt{2 \sum_{i=1}^{n} f_i d_{ic}^2} \quad (15)$$

for the weighted case.

The implication of this is two-fold. First, it is easy to calculate the mean of squared distances using the variances of X and Y. Second, and more important, we can see that the mean of distances squared among all points is related to distances from one specific point, the mean centre. This will help us in understanding and interpreting of the standard distance.

(iii) The standard distance

The standard distance is the extension of the standard deviation of univariate statistics to the two dimensional case. It was introduced by Bachi (1958, 1962), though a similar measure using direct, non-Pythagorean distances had been proposed by Stewart and Warntz (1958) and by Warntz and Neft (1960), and was termed the 'dynamical radius' or the 'standard distance deviation'. Unfortunately, Furfey's (1927) similar measure was not used until re-developed by Bachi.

While the first two measures of dispersion presented here are based on distances of all the points from each other, the standard distance is based, like the standard deviation, on the average of distances of all points from a central point, namely the mean centre. It is defined as

$$d = \sqrt{\frac{\sum_{i=1}^{n} d_{ic}^2}{n}} \qquad (16)$$

for the unweighted case, and

$$d = \sqrt{\sum_{i=1}^{n} f_i d_{ic}^2} \qquad (17)$$

for the weighted case, where

$$d_{ic}^2 = (X_i - \bar{X})^2 + (Y_i - \bar{Y})^2 \qquad (18)$$

The square of the standard distance, d^2, will therefore, be the 'distance variance'.

Ebdon (1977) proposed another version, easier to compute, in which

$$d = \sqrt{\frac{\sum_{i=1}^{n} X_i^2}{n} - \bar{X}^2 + \frac{\sum_{i=1}^{n} Y_i^2}{n} - \bar{Y}^2} \qquad (19)$$

In the weighted case this will become

$$d = \sqrt{\left(\sum_{i=1}^{n} f_i X_i^2 - \bar{X}^2\right) + \left(\sum_{i=1}^{n} f_i Y_i^2 - \bar{Y}^2\right)} \quad \text{(Table 5)} \qquad (20)$$

There are at least three other alternatives to define d which may be used to illustrate some of its qualities and relations to other measures (Bachi, 1962; 1966). First, due to the definition of the mean centre as that point to which the sum of squared distances of all points of the series is minimal, the value of the standard distance measured about the mean centre

Table 5. Computation of the Weighted Standard Distance

Coordinates		Weights			
X	Y	P	f	fX2	fY2
4	6	9	0.11	1.76	3.96
6	6	4	0.05	1.8	1.8
3	3	10	0.12	1.08	1.08
7	2	4	0.05	2.45	0.2
2	9	20	0.25	1.00	20.25
6	2	8	0.10	3.6	0.4
2	7	20	0.25	1.00	12.25
5	3	6	0.07	1.75	0.63
Total		81	1.00	14.45	40.57

$\bar{X} = 3.40$; $\bar{Y} = 5.8$ (See Table 2)

$$d = \sqrt{\left(\sum_{i=1}^{n} f_i X_i^2 - \bar{X}^2\right) + \left(\sum_{i=1}^{n} f_i Y_i^2 - \bar{Y}^2\right)} = \sqrt{(14.44-11.56) + (40.57-33.64)} = 3.13$$

$\sigma_x = 1.69$; $\sigma_y = 2.63$

is minimal. The standard distance is, therefore, the minimal second moment of the point-series. Second, looking at the structure of d (equation 18) and using the simple argument developed in equation 12 for DS it is clear that the standard distance is equal to the square root of the summation of the variances of X and Y, or

$$d = \sqrt{\sigma_X^2 + \sigma_Y^2} \qquad (21)$$

which relates d directly to the deviations along the two axes, and which makes d even easier to calculate. Third, since it has been shown that

$$DS = \sqrt{\frac{2 \sum_{i=1}^{n} d_{ic}^2}{n}} \qquad (14)$$

and d is defined as

$$d = \sqrt{\frac{\sum_{i=1}^{n} d_{ic}^2}{n}} \qquad (16)$$

it is clear that the relation between the average of squared distances among all points and the average of distances from all points to the mean centre is $\sqrt{2}$, or

$$d = \sqrt{\frac{\overline{DS^2}}{2}} \qquad (22)$$

(iv) Charactersitics of the standard distance

Being a complementary measure of the mean centre, the two measures share most of their characteristics, such as invariation in case of rotation or transposing of the coordinate system (Shimoni, 1963), sensitivity to changes in the location of any point, giving more weight to peripherally located points, being synthetic in the sense that its location on the map (see sub-section v) does not reveal any information about that area and that two series may yield equal values of d. All these qualities have been discussed already in detail in Section II(ii)

While the mean centre has little significance as a numerical value until presented graphically, the standard distance yields only a numerical value. This means that it has to undergo further processing in order to be presented on a map (using the 'cross of dispersion'), but that values of d for several series may be easily compared using measures of relative dispersion (heft, 1966; Taylor, 1977). If two series are compared so that the geographical distribution of the same variable is studied in two regions, then differences in the size of the two regions will yield different values of d reflecting differences in territorial size rather than differences in the distribution of the studied variable. This may be solved by using a measure of relative dispersion rd_z for which

$$rd_z = \frac{d_z}{A_z} \qquad (23)$$

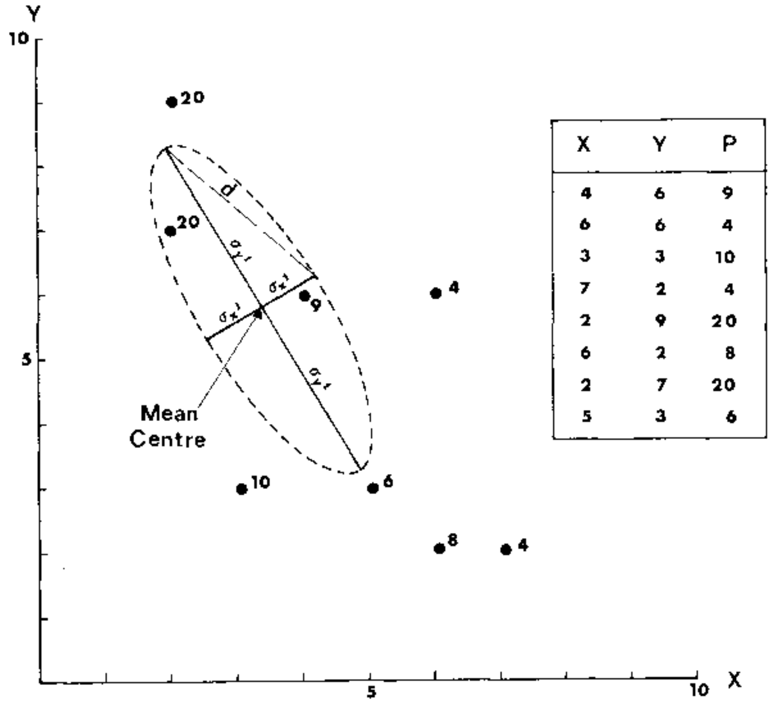
where z is the studied variable and A is any measure representing territorial size, such as the radius of the territory if transformed to a circle, or just the area of the territory. Similarly, when the distribution of a certain variable is related to the distribution of another, such as where the location of any urban service is related to population distribution, and we wish to compare a few cities, then the standard distance of the urban service of each city is divided by the standard distance of population of that city, so that inter-city comparison becomes possible. In this case, therefore,

$$rd_z = \frac{d_z}{d_p} \qquad (24)$$

where p stands for population.

(v) The cross and ellipse of dispersion

We have now established the mean centre as an analogue for the mean and the standard distance as an analogue for the variance and standard deviation for spatial data. The relationship between these two measures, however, is not straightforward. First, it is difficult to relate the value of the standard distance to the mean centre, since the mean centre does not have much meaning as a value, when expressed by its two coordinates $(\bar{X},\bar{Y})$, but is a point on a map, while the standard distance is expressed as one non-



X	Y	P
4	6	9
6	6	4
3	3	10
7	2	4
2	9	20
6	2	8
2	7	20
5	3	6

Figure 2. The Cross and Ellipse of Dispersion.

cartographic value. Second, the distance-deviation of the points around the mean centre is not equal in every direction, so that it has to be summarized and expressed at least in two directions along the X and Y axes. These two axes of dispersion could eventually be plotted on a map, and thus serve as a complementary graphic expression to the mean centre. Thus, the cross of dispersion will serve as a graphic presentation of d. It consists of two arms of lengths $2\sigma max$ and $2\sigma min$ or $2\sigma X'$ and $2\sigma Y'$, centered on the mean centre, and rotated through an angle α at which the two arms become $2\sigma max$ and $2\sigma min$ (Fig. 2). This cross will illustrate, therefore, the direction of the dispersion of the data points around the mean centre and the direction and magnitude of the minimum and maximum dispersion. The sum of the squared distances from the data points to the longer axis will be minimal while the sum of squared distances to the shorter axis will be maximal. In addition since

$$d^2 = \sigma max^2 + \sigma min^2 = \sigma X^2 + \sigma Y^2 \qquad (25)$$

by definition of the standard distance, it is easy to plot d on the cross of dispersion, so that d becomes a diagonal (Fig. 2).

The ellipse of dispersion (or standard deviational ellipse) drawn around the cross of dispersion will delimit an area with a radius of one standard deviation in each direction around the mean centre. The calculation and presentation of the cross and ellipse of dispersion involve the calculation of the angle of rotation (α), the arms (σmax, σmin), the ellipse and various indices that can be calculated from these measurements.

1) The angle of rotation: The angle of rotation (α) is the angle by which a cross centred on the mean centre at 90° parallel to the map coordinate axes must be rotated so that the arms will present the major directions of the point series (maximum and minimum dispersion). This is the angle at which σX and σY are uncorrelated, and thus

$$\sigma max \cdot \sigma min = \sigma X \sigma Y \sqrt{1-r^2} \quad (26)$$

where r is the Pearson correlation coefficient between σX and σY,

and

$$\sigma max \cdot \sigma min < \sigma X \cdot \sigma Y \qquad r \neq 0 \quad (27)$$

Lefever (1926) followed by Ebdon (1977) suggested the following formula for α.

$$\tan(\alpha) = \frac{-(\Sigma X_i^2 - \Sigma Y_i^2) + \sqrt{(\Sigma X_i^2 - \Sigma Y_i^2)^2 + 4(\Sigma X_i Y_i)^2}}{2\Sigma X_i Y_i} \quad (28)$$

While this equation involves considerable computation, the solution proposed by Linders (1931) as quoted by Shimoni (1963), Bachi (1966), and Kadmon (1968, 1971) is much easier to calculate and yields the same results, (equation 29). (The difference between equations (28-29) stems from a slightly different equality to zero of the derivatives of equations (34-35) with regard to a. See Lefever (1926),Shimoni (1963)).

$$\tan(2\alpha) = \frac{2cov(X;Y)}{\sigma X^2 - \sigma Y^2} \quad (29)$$

where

$$cov(X;Y) = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{n} \quad (30)$$

for the unweighted case, or

$$cov(X;Y) = \sum_{i=1}^{n} [f_i(X_i - \bar{X})(Y_i - \bar{Y})] \quad (31)$$

for weighted data points (Table 6).

Clearly, since the relative dispersion of the X and Y values can be quite different, and the dispersal of one can thus be larger or smaller than the other, then equation (29) can produce both positive and negative values for $\tan\alpha$. If the value is negative, then when scanning tangent tables for α, the negative sign is ignored, but the correct value for α will in fact be $(90^o - \alpha)$.

Table 6. Computation of the Weighted Cross of Dispersion

Coordinates		Weights		Covariation		
X	Y	p	f	$X-\bar{X}$	$Y-\bar{Y}$	$f(X-\bar{X})(Y-\bar{Y})$
4	6	9	0.11	0.6	0.2	0.013
6	6	4	0.05	2.6	0.2	0.026
3	3	10	0.12	-0.4	-2.8	0.134
7	2	4	0.05	3.6	-3.8	-0.684
2	9	20	0.25	-1.4	3.2	-1.120
6	2	8	0.10	2.6	-3.8	-0.988
2	7	20	0.25	-1.4	1.2	-0.420
5	3	6	0.07	1.6	-2.8	-0.313
Total:		81	1.00			-3.352

$\bar{X} = 3.4$ (See Table 2)

$\bar{Y} = 5.8$

$\sigma x = 1.69$ (See Table 5)

$\sigma y = 2.63$

$$\tan(2\alpha) = \frac{2cov(X;Y)}{\sigma X^2 - \sigma Y^2} = \frac{-6.704}{-2.57} = 2.60$$

$\alpha = 34^o30'$ $\cos(\alpha) = 0.82$; $\sin(\alpha) = 0.56$

$$\sigma x' = \sqrt{\sigma x^2 \cos^2\alpha + \sigma y^2 \sin^2\alpha + 2cov(x;y)\sin\alpha\cos\alpha} = 1.01$$

$$\sigma y' = \sqrt{\sigma x^2 \sin^2\alpha + \sigma y^2 \cos^2\alpha - 2cov(x;y)\sin\alpha\cos\alpha} = 2.91$$

(See Fig. 2)

An iterative graphical (and hence inexact) solution has been described by Hyland (1970; 1975), and noted or used also by Caprio (1970), Buttimer (1972), Herbert and Raine (1976), Raine. (1978) and by Matthews (1980). 'Employing an arbitrary cartesian grid, the mean centre of the points is calculated. A new grid, composed of orthogonal axes in the same scale as the first, is centred on the mean centre. The deviation of each point, from the X-axis, is employed to calculate the standard deviation. The axes are then rotated, say by 10 degrees, and the procedure repeated. If all the standard deviations along the X-axes are joined, an ellipse will be described'. (Hyland, 1975, p. 263).

2) The arms: Any point formerly defined as X_i, Y_i will now be redefined according to the rotated cross-axes, so that

$$X'_i = Y_i \sin\alpha + X_i \cos\alpha \qquad (32)$$

$$Y'_i = Y_i \cos\alpha - X_i \sin\alpha \qquad (33)$$

Therefore, the new axes $\sigma X'$ and $\sigma Y'$ to be allocated on the cross-axes, and defined as σmax and σmin, are

$$\sigma X' = \sqrt{\sigma X^2\cos^2\alpha + \sigma Y^2\sin^2\alpha + 2cov(X;Y)\sin\alpha\cos\alpha} \qquad (34)$$

$$\sigma Y' = \sqrt{\sigma X^2\sin^2\alpha + \sigma Y^2\cos^2\alpha - 2cov(X;Y)\sin\alpha\cos\alpha} \qquad (35)$$

This is Shimoni's (1963) solution, used also by Bachi (1966) (Table 6). Similar, though more cumbersome solutions have been offered by Lefever (1926) and Ebdon (1977). Another solution by Kadmon (1968; 1971) was found by Gordon, Kellerman and Waterman (1980) to be misleading since it was based on some non-existant relations between the X;Y axes and the X';Y' axes.

3) The ellipse of dispersion: The cross-arms $\sigma X'$ and $\sigma Y'$ may serve as radii for an ellipse (Fig. 2). The analysis of this ellipse may assist in the analysis of dispersion, since the ellipse delimits an area of one standard deviation in each direction (or an area which includes 68% of the distribution in case of a normal distribution, or at least 37% of the weighted series in all other cases). Lefever (1926) gives the area of the ellipse as

$$F = \pi\sigma X'\sigma Y' \qquad (36)$$

Furfey (1927) showed that an ellipse will be created in many cases, but that the general formula is

$$A = \frac{1}{2}\pi d^2 \qquad (37)$$

An ellipse will not result if certain specific relations occur between σX and σY on the one hand and the sets of point coordinates X and Y on the other. Thus, if $\sigma Y=2\sigma X$ and $r(X;Y)=1$, then two circles will result rather than the ellipse. Also, if $\sigma Y=2\sigma X$ and $r(X;Y)=.5$ then a collapsing ellipse will result. Finally, if $\sigma Y=\sigma X$ and $r=0$, then a circle will replace the ellipse (Fig. 3).

A few writers, such as Schneider (1968) and Yuill (1971) have fitted ellipses of dispersion around the median centre (defined as the point of minimum aggregate travel). This is an incorrect usage, since it is analogous

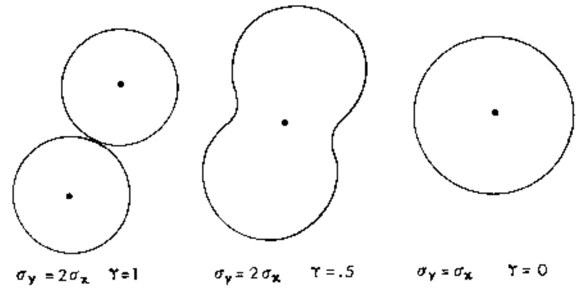


Figure 3. Furfey's Various Shapes of Standard Deviation Areas.

to the calculation of standard deviation for the median rather than the mean in uni-dimensional statistics. The reader is reminded that the mean centre is defined as the point for which the second moment (or standard distance) is minimal. By the same token a few writers (Taylor, 1977; Silk, 1979; Jones, 1980) have suggested presenting the standard distance graphically by drawing a circle with a radius of d around the mean centre. That this is misleading is clearly understood from the earlier discussion.

4) Indices: Using the cross-arms and the ellipse of dispersion several indices have been proposed which enhance the analysis of dispersion in addition to the comparison of areas of ellipses.

Lefever (1926) suggested a concentration index Ca in which

$$Ca = \frac{Ne}{F} \qquad (38)$$

where

Ne = Number of data points within the ellipse

F = Area of the ellipse (which may be replaced by Furfey's A)

Ca, or the number of points per square kilometer may be divided by the total number of points in the study area (n) so that the concentration of points within the ellipse will be evaluated relative to the total population.

Bachi (1973, 1976) developed a series of indices based on $\sigma X'$ and $\sigma Y'$ which enable one to understand some of the factors pertaining to the dispersion, or shape, of a territory, when an area is studied, and the relationships between a dispersed variable and the given territory on which it is dispersed when the spatial distribution of a variable is studied.

The shape of a territory consists of its oblongitude and structure. The oblongitude of a territory is defined by

$$L = \frac{1}{2}\left[\frac{\sigma X'}{\sigma Y'} + \frac{\sigma Y'}{\sigma X'}\right] = \frac{d^2}{2\sigma X'\sigma Y'} \qquad (39)$$

This measure varies between 1 and ∞.

The structure of a territory can be defined by a parameter which will indicate how much larger the 'spread per unit of area in a given territory is in comparison with that of an ellipse, so that

$$S_r = \frac{4\pi\sigma X'\sigma Y'}{A} \qquad (40)$$

where A is the area of the territory. Combining the two parameters together Bachi T1973; 1976) defined the 'relative spread of a territory' as

$$d_r^2 = LS_r = \frac{(2\pi)d^2}{A} \qquad (41)$$

In the study of the distribution of a variable over a given territory it is possible to define 'eccentricity index'

$$E = \sqrt{\frac{(\bar{X}_v-\bar{X}_t)^2 + (\bar{Y}_v-\bar{Y}_t)^2}{d_t}} \qquad (42)$$

where v denotes the variable and t the territory.

We may, of course, compare the parameters formerly defined for the territory with those of the variable distributed on it. This will assist us in analyzing the pattern of distribution of the variable over the territory. Assuming the denominators in the following ratios to represent uniform distributions we can evaluate the departure of each distribution of a variable from such a uniform distribution, and compare several variables over different territories, or same variables over different territories. Thus, it is possible to compare distance variances by

$$D_{pt} = \frac{d_p^2}{d_t^2} \qquad (43)$$

oblongitude ratios

$$L_{pt} = \frac{L_p}{L_t} \qquad (44)$$

and principal axes by

$$P_{pt} = \frac{\sigma X'_p \sigma Y'_p}{\sigma X'_t \sigma Y'_t} \qquad (45)$$

The last ratio (45) will be close to zero for a concentrated variable, it is 1.0 for a variable uniformly distributed, and it may exceed 1.0 for a highly concentrated variable at the periphery of a territory.

(vi) Further developments

The preceding sections have proposed several measures and indices which enable the description and anlysis of a given geographical point set either in itself or in comparison to other such sets. It is possible to extend these measures for the case in which we assume that a territory or a variable spread over a territory consist of a number of 'groups' or clusters of points which define sub-territories in the case of a variable or which define sub-areas in case of analysis of a shape of a territory. Measurements for this splitting of a territory have been developed by Bachi (1962; 1966), while the algorithm of an optimal division of a territory has been proposed by Ever-Hadani (1980). Also, Bachi (1962, pp. 121-3) developed several indices which allow for the comparison of point location in either two sub-territories or in two adjacent territories.

IV APPLICATIONS

In the following section applications of centrography in several sub-fields of geography will be reviewed. While reference has been made to over twenty works, the reader should be aware that this is a relatively small number, remembering that during the last generation quantitative methods have been routinely used in geographical work, while centrographic measures have been mostly ignored for reasons discussed already.

The use of centrographic measures in geographical work is quite diverse. Early Russian centrographic work concentrated on applications in economic-geography looking for centres of various agricultural crops for planning purposes. An example of this work is provided by Sviatlovsky and Eells (1937). A second generation of work in applied centrography is characterized by its use for the study of population distribution as in the work of Neft (1960; 1966) and those of Bachi (1958; 1962) stemming from his interest in demography. The late sixties and early 1970's could be regarded as a third phase in applied centrography in which the methods have been used mainly in studies of urban geography. This period was also marked by considerable use of centrographic measures in theses and dissertations written in the Department of Geography of the Hebrew University, a result of the postgraduate course in geostatistics given by Bachi in that department. The late 1970's have witnessed a rediscovery of centrography by Anglo-American geographers, mainly in empirical studies in the newly developing sub-fields of behavioural and social geography. While this marks an interesting blend of the old and the new it is believed that centrography is a powerful tool for the description and analysis of data in all branches of geography, with a special promise for another new specialization in geography, namely optimal location. Before going into a more detailed discussion of the studies using centrographic measures, it is also interesting to note that a survey of the literature failed to find a single study using centrography in the field of sociology, though Lefever (1926) and Furfey (1927) made their early contributions with the sociological community in mind.

(i) Population geography

This is the most straightforward application of centrography in which the distribution of population over a region or a country is studied using settlements as points and their populations as weights. At least two points of view may be adopted in such studies. The first is a temporal one comparing the distribution of population in the same area over a period of time, resulting in the plotting of mean centres of the population for each year, thus studying possible movements of the population over the area. This study is then complemented by a comparative temporal study of dispersion by comparison of the standard distances of the population for each year looking for changes in the dispersion or concentration of the population. This is then complemented by a comparison of crosses of dispersion, thus concluding on changes in the direction of population dispersion. Another possible viewpoint is the comparison of concentration and dispersion of population distribution in one country or region with those of others either directly or through the use of any of the indices already presented in the preceding sections. By the same token several age or ethnic groups of the same population may be compared. Sometimes both approaches have been used in the same study.

Mean centres of many countries such as the U.S.A., the U.S.S.R., Germany, France and others were presented by Sviatlovsky and Eells (1937) who provide also an annotated bibliography of early works on other countries such as Japan. Bachi (1962) and Neft (1966) provide detailed analyses of the U.S.A. and Israel, and both compare these countries with others such as Brazil and Italy. Another study of Israel is that of Shachar's (1970). A study of Ireland has been provided by Waterman (1969b), Tanzania has been analyzed by Hirst (1971) and an analysis of the State of Iowa was demonstrated by Taylor (1977, pp. 24-5).

(ii) Urban Geography

Centrographic methods have been used in urban geography for three purposes: 1) Analysis of changes in the distribution of population in the city over time, 2) Analysis of the location and distribution of urban activities, (both industrial and commercial), 3) Analysis of urban internal migration.

The analysis of changes in the distribution of population inside the city is similar in its approach to that of regional or national study of population. Thus, Waterman (1969a) studied the location of Jewish and Arab population in Acre, Israel and showed changes in the mean location and dispersion over time. A similar analysis of the Jewish-religious population of Jerusalem has been provided by Bukshpan (1972) who used, in addition to mean centres, standard distances and crosses of dispersion, an index of inter-territorial comparison (minD). Comparing the distribution of the religious population over the territory of Jerusalem and that of the general population over the same territory, she showed that the religious population is more concentrated than the general one. Shenar (1968; 1970) studied the urban development of Nablus and showed the various centres of development of the city over time and its gradual elongation by using mean centres and crosses of dispersion.

The location and distribution of economic activities in the city lends itself easily to centrographic analysis. Shachar (1967) studied the mean location and dispersion of over fifty services in Tel-Aviv. He claimed that the mean centre is especially useful for analysis of urban land-uses since most of the changes occur in the urban periphery, and that the sensitivity of the mean centre to remotely located points makes it an excellent tool for a comparative study. On the other hand, the median centre was found to be very helpful in the locational analysis of CBD functions which require high accessibility. The median centre which minimizes distances is, therefore, a good measure for those functions. Shachar also compared the Tel-Aviv findings with those of Jerusalem and Rome, and used the cross of dispersion for identifying the directions of dispersion of several urban functions. Waterman (1969a) provided a similar analysis of Acre and studied the mean location and dispersion of urban services over time and separately for the Jewish and Arab populations. Davies (1972) performed a centrographic analysis of the retail pattern of Coventry using mean centres and standard distances.

A study of industrial location and dispersion in Tel Aviv was performed by Gonen (1968) who mapped the mean location and dispersion of several industries over time, which enabled him to study further, their location factors and possible inter-industry linkages.

The third use of centrography in urban geography was demonstrated by Goldberg (1975) who studied the pattern of internal migration in Jerusalem. He calculated mean centres and crosses of dispersion for the maps of in-migration and out-migration of the several sections of Jerusalem, so that he identified the directions of both migrations. Doing so for the years 1966 and 1970 he could trace the changes in the geographical pattern of the internal migration in the city.

In a recent study, Jones (1980) made use of all three applications of centrography to urban study mentioned here. His study of Atlanta presents a centrographic mini-atlas of the city.

(iii) Agricultural and rural geography

Unfortunately only a few studies in these areas have used centrographic methods. Bar-Gal (1971) calculated mean centres of various crops for two Arab villages in Israel in order to study the adjacency of several crops. The standard distances of these crops provided information on both their dispersion and the relationship between dispersion and size of area devoted to each crop when the two variables were compared. Ever-Hadani (1980) used his model of optimal division of a territory for a rational regionalization of cotton growing areas in the U.S.

Kellerman (1972) studied the location and dispersion of the inter-village service and industrial centres in Israel. Computing mean centres and crosses of dispersion for the several levels of such centres he was able to show those which tended to be located and concentrated in the north or south of the country. In other words, it was possible to show numerically and graphically that as the centres' levels rise they become more concentrated.

(iv) Behavioural and social geography

As already noted the use of centrography in these sub-fields has developed in the last few years and promises to be of much further use in the future.

Ebdon (1977) has demonstrated the possible use of centrography in the study of mental maps. He asked a class of students at the University of Nottingham to locate on a base-map twenty major landmarks in Nottingham, while the base-map in itself consisted of only three landmarks for orientation purposes. The students' mental maps were analyzed by calculating mean centres and ellipses of dispersion for each of the twenty landmarks and by drawing them on a 'true' map, thus comparing the students' mean responses with the true locations of the landmarks.

Further hypotheses were tested by dividing the students into several sub-groups according to their origin (urban and rural) and according to their transportation usage. This is an excellent example of the possible use of centrography as an analytical tool, where the mean centres served as indicators for differences in landmark-perception, standard distances provided information on the dispersion of the mean perception, and the ellipses demonstrated the direction of spatial bias in those perceptions. A similar approach was used by Matthews (1980) in his study of children's mental maps of Coventry. Delimiting ellipses around the mean centres of mental maps of several age groups, he was able to demonstrate the spatial expansion of the children's mental maps, when several age groups were compared. In addition, the several shapes and locations of the ellipses assisted in the analysis of the spatial cognition of the city centres by the children.

Another possible use of centrographic measures in behavioural-social geography is the study of activity-spaces. The pioneering study in this area is probably that of Bashur et al., (1967), as quoted by Caprio (1970) and Yuill (1971). They examined where people moved in going to work or visiting friends thereby employing the ellipse dynamically to indicate social interaction' (Caprio, 1970, p. 18). Hyland (1970) delimited ellipses for data on social interaction of Appalachian migrants in Cincinnati as indicated by visits to friends and organizations. The ellipses demonstrated the activity-spaces of these migrants. A similar approach was used by Raine (1978) in his study of Cardiff. Buttimer (1972) used standard deviational ellipses for the delimitation of planned and unplanned activity-spaces, and for the comparison of activity-spaces of several types of households and several territories. She recommended the use of ellipses for the study of level of social interaction, its concentration, shape, direction, and differences when a few services are compared. Herbert and Raine (1976) went one step further when they drew ellipses around activity-spaces of social interaction in several neighbourhoods in Cardiff, and proposed using ellipses produced from these data and data on neighbourhood service areas, for the delimitation of community council boundary-lines.

An interesting use of centrographic measures for crime-geography was done by Rose and Deskins (1980) in their study of Detroit. Using ellipses of dispersion they showed the less dispersed pattern of black robbery-homicide and the more dispersed character of white robbery-homicide. Also when the distribution of homicides sites is compared with the residential location of the victims, the victim residences are more dispersed.

(v) Further possible use of centrography

Another possible use of centrographic measures is for the optimal location of services. It has already been proposed (Kellerman, 1971) to use the median centre for the location of regional agricultural service industries owned by Kibbutzim in Israel and many writers have proposed the median and mean centres for the location of public services (Bachi, 1962, p. 93; Smith, 1977, pp. 182-187). Morrill and Symons (1977) compared the median and mean centres as possible solutions, and proposed viewing the median centre as an efficient and, therefore, optimal location since it minimizes the sum of distances of all users to the facility. On the other hand, however, the mean centre is an equitable solution since it is sensitive to remotely located users (by squaring the distances from users to the facility), while still weighting the size of each users-location. Schneider (1968) proposed using the standard deviational ellipse for hospital location but his model, unfortunately, cannot be used, since he suggested drawing ellipses around the median centres and his method of division into sectors is not fully developed. The use of mean centres as equitable locations for public services such as hospitals could be enhanced, however, with the complementary use of the standard distance (Kellerman, 1981). If a certain value of standard distance is decided upon as a critical value for the distances to be overcome by users when going to the hospital, then the mean centre may be evaluated in light of this value, so that if it is being exceeded at the mean centre, then the region may be sub-divided using Ever-Hadani's (1980) algorithm, so that two or more facilities will have to be built.

V SUMMARY AND CONCLUSION

This CATMOG has presented one of the earliest attempts of quantitative geography, though most of the contributions to its theory have not been made by geographers. The various measures and procedures offered by centrography have been discussed, and though the order of presentation has been from centrality to dispersion, it has also gone from the easier to the more difficult techniques.

It has also been the purpose of this CATMOG to propose several uses of centrography in geographical research and it is hoped that this volume will make its contribution in encouraging geographers to use these techniques which have been developed specifically for use in geographical research.

BIBLIOGRAPHY

A. CENTROGRAPHIC MEASURES

Bachi, R., (1957), Statistical analysis of geographical series. *Bulletin de l'Institut International de Statistique,* 36, 229-40. Also in *Spatial Analysis* (1968), edited by B.J.L. Berry and D.F. Marble, (Englewood Cliffs, N.J., Prentice Hall), 101-9.

Bachi, R., (1962), Standard distance measures and related methods for spatial analysis. *Papers and Proceedings of the Regional Science Association,* 10, 83-132.

Bachi, R., (1966), *An introduction to Geostatistics.* Compiled by J. Yam, (Jerusalem, Academon) (Hebrew).

Bachi, R., (1973), Geostatistical analysis of territories. *Bulletin of the International Statistical Institute,* 45, 121-33.

Bachi, R. (1976), Geostatistical analysis of Internal migrations. *Journal of Regional Science,* 16, 1-19.

Bachi, R. and E. Samuel-Kahn, (1976), Applications of parameters of shape to statistical distributions. *Regional Science and Urban Economics,* 6, 205-27.

Cole, J.P. and C.A.M. King, (1966), *Quantitative Geography.* (London, Wiley).

Court, A., (1964), The elusive point of minimum travel. *Annals of the Association of American Geographers,* 54, 400-3.

Ebdon, D., (1977), *Statistics in Geography.* (Oxford, Basil Blackwell).

Ever-Hadani, S., (1975), *An optimal partition of geostatistical and multivariate series - methodology and application.* (Hebrew University, unpublished Ph.D. dissertation) (Hebrew with an English abstract).

Ever-Hadani, S., (1980), Applications of cluster analysis algorithm to geo-statistical series. *Regional Science and Urban Economics,* 10, 123-51.

Furfey, P.H., (1927), A note on Lefever's 'standard deviational ellipse'. *American Journal of Sociology,* 33, 94-8.

Gordon, D., A. Kellerman and S. Waterman, (1980), On the computation of the cross of dispersion. Unpublished paper.

Hammond, R. and P. McCullagh, (1974), *Quantitative techniques in geo-graphy.* (Oxford, Clarendon).

Hart, J.F., (1954), Central tendency in areal distributions. *Economic Geography,* 30, 48-59.

Kadmon, N., (1968), The mapping of distribution parameters. *The Cartographic Journal,* 64-9.

Kadmon, N., (1971), Comparing economic maps with the aid of distribution-parameter graphics. *International Yearbook of Cartography,* 11, 278-87.

Kellerman, A., (1981), The use of standard distance as an equity criterion for public facility location. *Area,* 13, (forthcoming).

Lefever, D.W., (1926), Measuring geographic concentration by means of the standard deviational ellipse. *American Journal of Sociology,* 32, 88-94.

Linders, F.J., (1931), Uber die berechnung des schwerpunkt and der trägheits-ellipse einer bevölkerung. *Atti del Congresso Internationale per lo Studio dei Problemi della Popolzione,* 10.

Morrill, R.L. and J. Symons, (1977), Efficiency and equity aspects of optimum location. *Geographical Analysis,* 9, 215-25.

Neft, D.S., (1966), *Statistical analysis for areal distribution.* (Philadelphia, Regional Science Research Institute) Monograph series, 2.

Porter, P.W., (1963), What is the point of minimum aggregate travel? *Annals of the Association of American Geographers,* 53, 224-32.

Poulsen, T.M., (1959), Centrography in Russian geography. *Annals of the Association of American Geographers,* 49, 326-7.

Shimoni, Z., (1963), *Theoretical Research in statistical-geographical series.* (Hebrew University, unpublished Ph.D- dissertation) (Hebrew with an English abstract).

Silk, J., (1979), *Statistical concepts in geography.* (London, Allen and Unwin).

Smith, D.M., (1977), *Human geography: a welfare approach.* (London, Arnold).

Stewart, J.Q. and W. Warntz, (1958), Macrogeography and social science. *Geographical Review,* 48, 167-84.

Stewart, J.Q. and W. Warntz, (1959), Some parameters of the geographical distribution of population. *Geographical Review,* 49, 270-3.

Sviatlovsky, E.E. and W.C. Eells (1937), The centrographical method and regional analysis. *Geographical Review,* 27, 240-54.

Taylor, P.J., (1977), *Quantitative methods in geography: an intro-duction to spatial analysis.* (Boston, Houghton Mifflin).

Warntz, W. and D. Neft, (1960), Contributions to a statistical methodology for areal distributions. *Journal of Regional Science,* 2, 47-66.

Yam, J., (1968), *Concentration and segregation of geostatistical distributions: new methods and exemplary illustrations.* (Jerusalem, Szold Institute) publication 478.

Yam, J., (1969), *The analysis of divergence in geostatistical dis-tributions: methodology and applications.* (Hebrew Univer-sity, unpublished Ph.D. dissertation) (Hebrew with English abstract

Yeates, M., (1974), *An introduction to quantitative analysis in human geography.* (N.Y., McGraw-Hill).

Yuill, R.S., (1971), The standard deviational ellipse: an updated tool for spatial description. *Geografiska Annaler,* 53B, 28-39.

B. APPLICATIONS

Bar-Gal, Y., (1971), *The distribution of agricultural land-uses in Makr and Jdeida (the Western Galilee).* (Hebrew University, unpublished M.A. thesis) (Hebrew).

Bashur, R.L., G.W. Shannon, S.E. Flory, and R.V. Smith (1967), *Community interaction and racial integration in the Detroit area: an ecological analysis.* Report for project No. 2557, U.S. Office of Education (Ypsilanti, Michigan, Eastern Michigan Uni-versity).

Bukshpan, S., (1972), *The spatial organization of the religious population in Jerusalem 1959-1971.* (Hebrew University, unpublished M.A. thesis) (Hebrew).

Buttimer, A:, (1972), Social space and the planning of residential areas. *Environment and Behavior,* 4, 279-318.

Caprio, R.J., (1970), Centrography and geostatistics. *Professional Geo-grapher,* 22, 15-9.

Davies, R.L., (1972), The retail pattern of the central area in Coventry. *The Institute of British Geographers, Occasional Pub-lications,* 1, 1-32.

Goldberg, R., (1974), *The spatial organization of intra-urban migration in Jerusalem.* (Hebrew University, unpublished M.A. thesis) (Hebrew).

Gonen, A., (1968), *Industrial location in Tel-Aviv.* (Hebrew University, unpublished Ph.D. dissertation) (Hebrew with an English abstract).

Herbert, D.T. and J.W. Raine, (1976), Defining communities within urban areas. *Town Planning Review,* 47, 325-38.

Hilgard, J.E., (1872), The advance of population in the United States. *Scribner Monthly,* 4, 214.

Hirst, M.A., (1971), Tanzania's population distribution 1948-67: some macroscopic comments. *Area,* 3, 2-7.

Hyland, G.A., (1970), Social interaction and urban opportunity: the Appalachian in-migrant in the Cincinnati central city. *Antipode,* 2, 68-83. Also in *Readings in Social Geography,* (1975) edited by E. Jones (London, Oxford University), 250-63.

Jones, B.G., (1980), Applications of centrographic techniques to the study of urban phenomena: Atlanta, Georgia 1940-1975. *Economic Geography,* 56, 201-22.

Kellerman, A., (1971), The application of economic location theories for the location of service-centers for Kibbutzim. *Environmental Planning,* 16, 12-20 (Hebrew).

Kellerman, A., (1972), *The spatial characteristics of* the *inter-village centers in Israel.* (Hebrew University, unpublished M.A. thesis) (Hebrew).

Matthews, M.H., (1980), The mental maps of children. *Geography,* 65, 169-79.

Raine, J.W., (1978), Summarizing point patterns with the standard deviational ellipse. *Area,* 10, 328-33.

Rose, H.M. and D.R. Deskins, Jr., (1980), Felony murder: The case of Detroit. *Urban Geography,* 1, 1-21.

Schneider, J.B., (1968), A new approach to areawide planning of metropolitan hospital systems. *Hospitals,* 42, 79-83.

Shachar, A., (1967), Some applications of geo-statistical methods in urban research. *Papers and Proceedings of the Regional Science Association,* 18, 197-206.

Shachar, A., (1970), The impact of the new towns on population distribution in Israel. *Studies in the Geography of Israel,* 7, 25-63, (Hebrew).

Shenar, A., (1968), *The urban geography of Nablus.* (Hebrew University, unpublished M.A. thesis) (Hebrew).

Shenar, A., (1970), The development of the built area in an Arab city. *Studies in the Geography of Israel,* 7, 96-124, (Hebrew).

Waterman, S., (1969a), Some aspects of the urban geography of Acre, Israel. (University of Dublin, unpublished Ph.D. dissertation).

Waterman, S., (1969b), Some comments on standard distance: a new application to Irish population studies. *Irish Geography,* 6, 51-62.

25. Directional statistics - G.L. Gaile & J.E. Burt
26. Potential models in human geography - D.C. Rich
27. Causal modelling: the Simon-Blalock approach - D.G. Pringle
28. Statistical forecasting - R.J. Bennett
29. The British Census - J.C. Dewdney
30. The analysis of variance - J. Silk
31. Information statistics in geography - R.W. Thomas
32. Centrographic measures in geography - A. Kellerman

This series, Concepts and Techniques in Modern Geography is produced by the Study Group in Quantitative Methods, of the Institute of British Geographers.
For details of membership of the *Study Group, write to the Institute of British Geographers, 1 Kensington Gore, London, S.W.7.*
The series is published by Geo Abstracts, University of East Anglia, Norwich, NR4 7TJ, to whom all other *enquiries should be addressed.*