4. REPRINT

Imperfections in the Uniform Plane

Michael F. Dacey

with Forewords by John D. Nystuen, The University of Michigan

In this section, *Solstice* Board member, John D. Nystuen, selects a paper from the collected papers of the Michigan Inter-University Community of Mathematical Geographers (MICMOG) (of which he is Editor) to reprint here, some 30 years after its initial presentation. In addition to the reprint of work of Michael Dacey, Nystuen's original Foreword, and introduction of Dacey and his work to the assembled MICMOG group, is also reprinted. In addition, a new Foreword by Nystuen takes a look at the Dacey paper in retrospect. The paper is reprinted with permission of Nystuen, on behalf of the Michigan Inter-University Community of Mathematical Geographers.

Foreword, December, 1994

John D. Nystuen

Thirty years ago Michael Dacey contributed to the development of spatial statistics in highly original ways. Many of the ideas he used and introduced to the literature in the 1960s are now part of generally accepted spatial theory. For example, he was one of the first to use the idea of a dimensional transformation to permit evaluations of the spatial association of point and area phenomena. The transformational approach proved useful as a general concept as Keith Clarke has demonstrated in his interesting book (Clarke, 1990). Arthur Getis, a colleague of Dacey's, and Barry Boots used many of Dacey's ideas in their book (Getis and Boots, 1978) about modelling spatial process.

Today, vigorous effort is being expended on incorporating spatial analysis functions into Geographic Information Systems (GIS) software. We are re-issuing one of Dacey's seminal works to bring to the attention of contemporary scholars an important source of many of the concepts now becoming accessible to general uses of GIS technology. Dacey's work now speaks to another generation.

References

Clarke, Keith C. 1990. *Analytical and Computer Cartography*, Prentice-Hall, Englewood Cliffs, NJ.

Getis, A. and Boots, B. 1978. *Models of Spatial Processes*, Cambridge University Press, Cambridge.

Foreword, May, 1964

John D. Nystuen

We are pleased to present to our readers a paper by Professor Michael F. Dacey. Many of us are aware, if only vaguely, of his provocative and voluminous writings. Professor Dacey has penetrated deeply into realms where few, if any, have gone before. He travels alone and has left but a thin trail of mimeographed papers as scent. The track is now long and difficult to follow and he does not rest. He has allowed one of his works to become discussion paper #4 of our series. We hope this will expose his activities to a wider audience. Some may be inspired to join him in the new work that he is doing. I hope so. Certainly we must keep in contact with him. Regrettably many of his results depend upon his previous statements

now difficult to obtain. I will attempt in this foreword a short review of the pertinent ideas by way of a summary of this paper. I have also added, with his permission, a glossary of symbols at the end of the paper.

Michael Dacey has for several years explored abstract spatial patterns using probabilistic methods. This paper is one of a series of such studies. Most of the work provides empirical examples of the concepts. The contrast in methodologies displayed between discussion paper #3 (W. Bunge, "Patterns of Location") and this one is marked. Professor Bunge turns away from probabilistic formulations (see page 3 of "Patterns of Location") and Professor Dacey rejects deterministic models (see page 1 below). I believe the relative worth of these two broad approaches to abstract geography will receive increasing attention in the literature. There is much precedent for concern over this question in other disciplines. Clearly Dacey accepts the value of a probabilistic approach.

It may aid the reader if the paper is viewed as consisting of six parts.

1. Professor Dacey first describes an abstract model of imperfections in a uniform plane. The characteristics of this model are specified in a general way. I believe that Professor Dacey is the first to suggest models where non-random patterns are disturbed by random variables (see Dacey and Tung, 1962).
2. The point pattern which results from the above mentioned model is to be summarized quantitatively in such a fashion that it can be compared with some actual geographic point pattern. Professor Dacey calls upon his previous extensive investigations of nearest neighbor statistics to do this job [1]. He specifies how measures of the distances to the 1st nearest, 2nd nearest, ... kth nearest neighbors of a sample of points in the point pattern may be used to describe the point pattern by probability distributions of these lengths. The strategy is to then compare the probability distributions of the model with a geographic pattern using a simple χ^2 statistic. Professor Dacey is aware that nearest neighbor methods may be used to compare point-to-area relations as well as point-to-point relations. A point pattern is not simply a set of points. The points occupy a space for which a metric is defined. The metric makes possible distance measures between the points. The fact that there is a space creates the boundary problems mentioned in the text. The original purpose of these statistics was to test if points were more clustered or more even than random. Imagine a study area which is mostly empty but has in one small region an even distribution of points. Measuring distances between points and using the nearest neighbor test would indicate a point pattern more even than random. In one sense, however, they are clustered for they occupy only a small section of the study area. There is a strategy for this situation. Use another point set to represent the area. This may be done by using an even distribution of points in the area or by assigning points to the area at random. The second set of points now represents the study area. The area has been abstracted into a point pattern and the nearest neighbor method may be used. Measures between the two point sets now reveals the original point pattern to be clustered. The decision concerning which method to employ depends upon whether the phenomenon studied has a postulated interaction of point-to-point or point-to-area. The text indicates the procedure for using either method.
3. Theoretical order distances are specified by equations (16) and (17). The probability functions are made more explicit and operational by assuming each lattice point is disturbed by the same two dimensional normal variate. Professor Dacey has ample evidence

that these particular probability distributions are useful for this purpose.[2]

4. Solutions of the equations in the previous section would yield an analytic solution regarding expected order distances for various disturbance models. However, these equations prove very difficult to evaluate. Recourse to a simulated solution is sought. An *almost periodic disturbance model* is postulated. Its parameters are estimated from data on an actual pattern of urban places in Iowa. Using these parameters, a set of points conforming to the structure of the theoretical model is generated with random digits and tables of normal deviates. This artificial pattern is one of many possible representations of the theoretical pattern. It is presumed to display the type of pattern expected from an analytic solution if one could be found.
5. The author now has two patterns: one, a simulated theoretical pattern which conforms to the structure of the model; and the other, an actual urban place pattern in Iowa. He also is able to make the appropriate nearest neighbor measures which characterize each pattern. The frequency distributions are then compared using the χ^2 statistic.
6. In an addendum, the author presents further testing of his model by taking advantage of a computer program which generates the distance measures required. The paper ends.

It must be clear to the reader from the contents of this paper that Michael Dacey has indeed traveled over much ground. He has previously developed many of the results needed in this study. Many of his solutions and applications are ingenious. He exhibits an understanding of the theoretical implications of his work. He has a wide knowledge of the literature on probability and is able to adopt simulation methods and computer technology to his purpose. All he lacks is someone to talk to.

Endnotes

1. Examples of his statements on nearest neighbor measures include: "Analysis of Central Place Patterns by Nearest Neighbor Method," Seattle, May 1959, mimeographed; "Analysis of Central Place and Point Patterns by a Nearest Neighbor Method," *Proc. of IGU Symposium in Urban Geography*, Lund, 1960, pp. 55-75; "Identification of Randomness in Point Patterns," (with Tze-hsiung Tung), Philadelphia, June 1962; mimeographed. (Dacey and Tung is now forthcoming in the *Journal of Regional Science*, v. 4.

2. See references at the end of the paper and also: "Order Neighbor Statistics for a Class of Random Patterns in Multidimensional Space," *Annals*, Association of American Geographers, v. 53 (Dec. 1963): 505-515, "Certain Properties of Edges on a Polygon in a Two Dimensional Aggregate of Polygons Having Randomly Distributed Nuclei," Philadelphia, June 1963, mimeographed.

Imperfections in the Uniform Plane

Michael F. Dacey

Wharton School of Finance and Commerce
University of Pennsylvania
See end of article for additional information

A statistical formulation of the spatial properties of central place system is proposed. Currently, the theoretical locations of central places are specified by geometric or algebraic quantities. This type of statement leads to certain rejection of central place models, for it is inconceivable that any observed pattern of central places corresponds exactly to the specified geometry. A probabilistic formulation is preferred for empirical analysis because deviations from the precise locations are contained within the statement of the model.

In the classical theory of Christaller (1933) and Lösch (1939) central places form a honeycomb pattern or hexagonal lattice on the undifferentiated, unbounded plane. A probabilistic statement of this location pattern incorporates deviations from the precise lattice locations, and the deviations are subject to stochastic processes. This initial formulation of a probabilistic central place distribution uses the concept of imperfections in the uniform plane to define these deviations. Imperfections may be combined with the central place geometry in many ways. Here one basic formulation and two closely related models are proposed. The models possess some properties of the Christaller-Lösch system and evidently are not inconsistent with the spirit of central place theory.

This report has two purposes. First, a general model of imperfections in the uniform plane is constructed. Second, the application of a particular model to a map pattern is evaluated.

The map pattern of urban places in Iowa has been selected for an initial examination of the imperfection concept. The empirical test involves interpretation of parameters of the model in terms of phenomena commonly studied by geographers and estimation of these parameters from the Iowa map pattern. Because the formal statement of the model contains equations that are difficult to evaluate analytically, this initial study has used a simulation technique to obtain summary measures on theoretical patterns. Properties of a fabricated pattern are compared with the Iowa map pattern, and the level of agreement is found acceptable to the first approximation.

The Christaller Spatial Model

The theoretical distribution of central places may be expressed in terms of a plane lattice. Let P represent a plane symmetry lattice. Choosing any arbitrary point of this lattice as an origin point O, the location of any other given lattice point can be defined with respect to this origin by a vector T

$$T = ut_1 + vt_2 \tag{1}$$

where u and v are integers. The vector notation implies that the plane is constructed as a linear lattice having a translation period t_1 which is repeated periodically at an interval t_2. The translation periods t_1 and t_2 may be regarded as vectors separated by the angle g. Using K to denote a collection, the lattice points of P are defined by

$$P = KT = K(ut_1 + vt_2). \tag{2}$$

Central place theory conventionally uses a hexagonal lattice for which the translations t_1 and t_2 are of the same unit length and the angle of periodic rotation is $g = \pi/3$.

A more general discussion is obtained by not restricting attention to the hexagonal lattice. In this report P represents any plane lattice which may have a three-, four-, or six-fold axis. In applying the lattice to a particular problem, the translation periods t_1 and t_2 and the angle of rotation g need specification.

Types of Imperfections in the Uniform Plane

Three types of imperfection in the uniform plane are studied in this report. These imperfections are closely related to certain kinds of imperfections found in nearly perfect crystals. An introduction to crystal imperfections is found in Van Bueren (1961, especially Chapters 2-4) and an excellent synthesis of the concept of imperfection in the solid state is given by Seitz (1952). The basic principles of our formulation draw heavily upon concepts used in the study of crystals and the solid state; the mathematical formulation is, however, quite different.

The imperfections under consideration are identified as (i) dislocations or disturbances, (ii) vacant lattice sites and (iii) interstitial points. These three types of imperfections are most easily defined by considering two maps containing point symbols. For the present purposes assume the maps have identical area and number of points. One map represents a finite domain of the lattice P. The other map, called S, may show fabricated locations or the positions of actual objects. Figure 1 is "good" map S overlaid on a square P.

i. The term dislocation is more descriptive of the first imperfection, but it has a definite meaning in crystallography and solid state physics; so we shall call this imperfection a disturbance. A disturbance occurs when the location of a point is not exactly at a theoretical lattice site but is 'sufficiently' close so that with high degree of certainty a disturbed point is correctly associated with its theoretical location.

ii. A vacant lattice site occurs where no point is 'close' to a theoretical lattice site. Where two or more points occur in the vicinity of a lattice site, it is not called a vacant lattice site even though the one point correctly associated with that theoretical location may not be identifiable.

iii. An interstitial imperfection occurs in the uniform plane where a point is not identified with any lattice site. Interstitial locations occur where a point is too distant from a theoretical location to be associated with high degree of certainty with a particular lattice site, or where two or more points are located 'close' to a lattice site and the one point correctly assigned to that theoretical location is not identifiable.

These imperfections are not given precise definitions. In constructing the imperfection model more precise definitions are given.

A Model of the Imperfect Plane

One basic formulation and two modifications are described. All imperfections under consideration are the result of stochastic processes, in the space rather than the more common time dimension. The principal feature of an imperfection model is the imperfection in pattern related to disturbances or shocks from geometrically exact locations (Figure 1). While this single type of imperfection is adequate for many physical systems, it is probably too restrictive to encompass patterns formed by economic, social or cultural systems. To

handle complex map patterns two additional types of two dimensional stochastic processes were studied. One type of imperfection generates interstitial points and is defined by a two dimensional, uniform, random variable. The other type of imperfection generates clusters of points and is defined by spatially contiguous probability distributions. Because the pattern of urban places in Iowa is relatively homogeneous and contains no examples of large metropolitan centers, it was not necessary to incorporate a contagious process in a model for the Iowa map pattern. For this reason, only the first two types of imperfections are discussed in this report.

The Disturbance Effect

Each lattice point of P is associated with a stochastic variable ξ. The ξ is the disturbance variable and defines the realized location of a point with respect to its theoretical lattice site. It is convenient to separate ξ into its two polar components: a distance ρ and a rotation angle θ. So, $\xi \equiv (\rho, \theta)$.

The displacement of the point s_{ab} from its equilibrium position $(at_1 + bt_2)$ is given by the random variable ξ_{ab}. So, the disturbed position of this point is

$$s_{ab} = at_1 + bt_2 + \xi_{ab}. \tag{3}$$

It is assumed that the same stochastic variable is associated with each lattice site. Then, if a point is disturbed from each lattice site the collection of randomly disturbed points is

$$S_1 = K(ut_1 + vt_2 + \xi_{ab}), \tag{4}$$

u and v integers. This notation indicates that ξ has translation period t_1 which is repeated periodically at an interval t_2. In this sense the stochastic variable is carried through space and is associated in turn with each lattice site. Accordingly, in point set S_1 each lattice site $(at_1 + bt_2)$ has exactly one corresponding disturbed point s_{ab}.

Vacant Lattice Sites

It is not necessary to apply a disturbance to each lattice site. Instead a lattice site and the variable ξ_{ab} may be taken in conjunction with a binary or on-off operator which nullifies the vectors defining some disturbed points so that the corresponding lattice sites are vacant. As a consequence, there is a sparser network of disturbed points than lattice sites. Because a disturbed point is not associated with each lattice site, the disturbance term is said to be repeated almost periodically. A more precise definition of the almost periodic disturbance is given.

A binary operator to produce vacant lattice sites is defined for $(at_1 + bt_2)$, denoted in symbols by β_{ab}, such that for $0 \leq \lambda \leq 1$,

$$\beta_{ab} = 1, \quad \text{with probability } \lambda$$

$$\beta_{ab} = 0, \quad \text{with probability } 1 - \lambda. \tag{5}$$

The vectors defining location of the disturbed point s_{ab} are multiplied by β_{ab} so that the disturbed point is realized with probability λ and is not defined with probability $(1 - \lambda)$. In more precise form, the location of the disturbed point having equilibrium position $(at_1 + bt_2)$ is

$$s_{ab} = \beta_{ab}(at_1 + bt_2 + \xi_{ab}) \tag{6}$$

Figure 1. Map of imperfection model. Most symbols show disturbance effect on a square lattice. There are two vacant lattice sites, and two examples of interstitial points. Most map patterns are, of course, not this regular. This figure shows a six by four square lattice which has been altered as suggested.

with the usual convention that $s_{ab} = 0$ does not define a point at the lattice site 0. So, for $\beta_{ab} = 0$ the disturbed point s_{ab} does not exist, while for $\beta_{ab} = 1$ location is found precisely in the manner for the period disturbance.

Each lattice site is associated with the same stochastic variable and with the same binary operator. Accordingly, the relation (6) is carried through space with translation period t_1 repeated periodically at interval t_2. The collection of points generated by the almost periodic

disturbance is

$$S_2 = K(\beta_{uv}(ut_1 + vt_2 + \xi_{uv}) \tag{7}$$

u and v integers. The S_2 is completely identified by the underlying lattice P, the probability λ, and the parameters specifying the components ρ and θ of the stochastic variable ξ. It is summarized by the parameter set $S(t_1, t_2; \lambda, \xi)$.

Uniform Random Disturbance

This collection of points, denoted by R, is a random point set. To make the definition explicit, an arbitrary origin is selected and the lattice point O of P is convenient. The R is specified by the theoretical frequency of points within distance r of the origin. Where the parameter γ is the expectation that a unit area contains a point belonging to R, put

$$p = \pi\gamma r^2 \tag{8}$$

where $\gamma > 0$. The frequency p describes any arbitrary disk of radius r, so that the distribution ξ is independent of the specified origin. It is a property of R, Feller (1957) that the distribution conforms to a Poisson process. The probability of finding exactly j points of R within any disk of radius r is $p^j e^{-p}/j!$.

Definition of the Basic Model

The model to be considered in this report is defined by the combination of an S and the R point sets; call this model M and

$$M = S \cup R. \tag{9}$$

This model is summarized by the parameter set $M(t_1, t_2; \lambda, \xi; \mu)$, where $\mu = (\lambda + \gamma)$. For a model containing S and R points only, μ is the mean density of points per unit area.

Several interesting formulations of M are defined by special values of the parameters λ and γ.

The *periodic disturbance model* M_1 is given by $\lambda = 1$, for one disturbed point is associated with each lattice site. A *complete periodic disturbance model* also has $\gamma = 0$, for each point is disturbed from a lattice site and there are no random points from R.

The *almost periodic disturbance model*, called M_2, is given by $0 < \lambda < 1$. The magnitude of γ determines if M_2 has a one-to-one correspondence of points to lattice sites or if M_2 has more or less points than lattice sites. If $\gamma = 1 - \lambda$ the theoretical density of points belonging to S_2 and R equals the density of lattice sites. If $\gamma > 1 - \lambda$ the expected number of points exceeds the number of lattice sites, while the expected number of points is less for $\gamma < 1 - \lambda$.

The point set given for $\lambda = 0$ is a random point pattern. It is of course recognized that R is only one of many point sets that could be combined with S_1 or S_2 disturbed points.

Description of Pattern

The disturbance models are described by the underlying lattice P, the density measures λ and γ and the disturbance process ξ. The combination of these parameters produce disturbed and interstitial points and vacant lattice sites in the uniform plane. In a formal sense a model is completely specified by the lattice parameters and the several probability functions. This specification of a model does not, however, describe or summarize in any

useful fashion the point pattern generated by a particular model. But, numerical summary of point pattern M is prerequisite to test of the hypothesis that an observed map pattern is similar to an imperfection pattern.

To measure the level of correspondence between observed and theoretical patterns there is need for (i) measurements on one or more properties of the observed pattern and (ii) theoretical values for the same properties on the pattern defined by the model. In addition, if parameter values for the model are estimated from the observed pattern, the properties for test of similarity between observed and theoretical patterns should be independent of the properties initially used to estimate parameters.

In this report pattern is summarized by two classes of order distance statistics. The methods are described briefly and then their utility as descriptive measures of pattern are indicated.

Point to Point Order Distances

Let i represent any arbitrary point in a point pattern Q. The measured map distance from i to the j nearest point is represented by R_{ij}. J measurements are taken from i and are ordered to satisfy the inequalities

$$R_{i1} < R_{i2} < \cdots < R_{ij} < \cdots < R_{iJ} \tag{10}$$

and the R_{ij} is called the j order distance. For description of a bounded map pattern the j order distance is recorded only if R_{ij} is less than the distance from i to the nearest map boundary. The chance of bias due to the influence of boundaries is reduced by this constraint, but there is loss of information to the pattern description because all distance relations are not utilized.

The R_{ij} measurements reflect the arbitrary map metric. The dimensional constant which eliminates effect of scale is $d^{1/2}$, where d is the density of points in Q. Measurements in Q are reduced to standardized distance by the transformation

$$r_{ij} = d^{1/2} R_{ij}. \tag{11}$$

Standard distances are used in this report to describe all patterns.

Let I denote a collection of points in Q, and $i \in I$. One description of Q uses standard distances from each origin point $i \in I$ to the J nearest points.

Locus to Point Order Distances

A second description of pattern uses distance measurements from coordinate locations to points. Let L define a set of locations in Q and in general a locus $\ell \in L$ is not a point symbol of Q. The measured distance in Q from locus ℓ to the h nearest point is denoted by $R_{\ell h}$. The measurements from ℓ are ordered by distance and put in standard form; in symbols

$$r_{\ell 1} < r_{\ell 2} < \cdots < r_{\ell h} < \cdots < r_{\ell H} \tag{12}$$

$$r_{\ell h} = d^{1/2} R_{\ell h}. \tag{13}$$

The second description of Q uses standard distances from each locus $\ell \in L$ to the H nearest points. The boundary constraint pertains to these distances also.

Sampling Methods

The elements of I may consist of all or a sample of points in Q. For this study a census was taken, largely because of small pattern size.

The loci in L necessarily constitute a sample, and these locations may be designated by random, stratified or uniform sampling methods. The most efficient mesh for plane sampling has been studied by a number of writers, as Zubrzycki (1961) and Dalenius, Hajek, and Zubrzycki (1961), but there are no general conclusions. This study used random sampling, largely because the patterns of interest contain high degree of uniformity in spacing and random sampling is probably less sensitive to this type of spatial bias. However, this topic requires study.

Summary Description of Pattern

A point pattern may be summarized by (i) the lower moments of the j and h order distances or (ii) the frequency distributions of these order distances. The j order point to point distances provide a quantitative summary of the arrangement of points with respect to other points of the pattern, but these distances do not explicitly reflect the arrangement of points with respect to the map space. The complementary h order locus to point distances provide a quantitative summary of the arrangement of points with respect to the loci in L. To the degree the sample mesh of L is a measure of the map space, h order distances also summarize the arrangement of points with respect to the map space. Because these two classes of distances reflect two different aspects of pattern, this type of summary statement captures many of the subtle characteristics composing a point pattern.

Comparison of Map Patterns

The descriptive measures provide a basis for evaluating the degree of similarity between two or more patterns. patterns are called similar if the order distances summarizing each of the patterns have the same statistical parameters. The standardized distances allow direct comparison of any two point patterns, for the distances represented by the variable r (either r_{ij} or $r_{\ell h}$ are normalized to account for differences in scale, unit measurement and density of points. Using either means or frequency distributions of order distances, the hypothesis that two or more sets of measurements belong to the same statistical population may be tested by standard procedures.

Theoretical Order Distances

This paragraph considers the basic derivation of order distances for imperfection models. The derivations are simplified by studying (i) lattices for which $t_1 = t_2$, (ii) nearest neighbor situations only, and (iii) the stochastic variable ξ defined by the normal law.

Two nearest neighbor lattice sites are separated by the distance t $(= t_1 = t_2)$. Let the random variable X denote the distance between two disturbed points associated with any two nearest neighbor lattice sites. It requires only elementary geometry to show that the distance between points (ρ_1, θ_1) and (ρ_2, θ_2) is

$$x = ((\rho_1 \cos\theta_1 - \rho_2 \cos\theta_2 + t)^2 + ((\rho_1 \sin\theta_1 - \rho_2 \sin\theta_2)^2)^{1/2}. \qquad (14)$$

The simplest derivation of order distances is for the complete periodic disturbance model ($\lambda = 1$ and $\gamma = 0$) on the hexagonal lattice. Let m $(= 6)$ denote the number of nearest neighbors to each lattice site. We consider the distances from an arbitrary point i at $(at_1 + bt_2 + \xi_{ab})$. It is assumed that the m nearest points to i are disturbed from nearest neighbor lattice sites only. The x_k is the distance from point i to the k $(= 1, 2, \ldots, m)$ nearest point.

If the disturbance term is identical and independent for each lattice site, the m distances from i may be interpreted as m independent observations in a sample of size m from the population defined by the random variable X. Because the observations are ordered from shortest to longest, x_k is the kth order statistic. It is well known that the distribution function of the kth order statistic is given by

$$\Psi(x_k) = \frac{m!}{(k-1)!(m-k)!} F^{k-1}(\omega) F^{k-1}(\omega)(1 - F(\omega))^{m-k} f(\omega) \tag{15}$$

where $f(\omega) = dF(\omega)$ and the variable X, after making the probability transformation for a specified $f(\rho)$ and $f(\theta)$, is substituted for ω. The z crude moment of the k order statistic for the complete periodic disturbance model is

$$\mu_z{}'(x_k) = \frac{m!}{(k-1)!(m-k)!} F^{k-1}(\omega) \int_0^\infty \omega^z F^{k-1}(\omega)(1 - F(\omega))^{m-k} f(\omega) d\omega. \tag{16}$$

The derivation is far more complex if the lattice is not hexagonal and undoubtedly requires more advanced concepts than provided by elementary probability methods. Moreover, even in this simplified case, numerical evaluation of (16) is not necessarily possible by elementary procedures.

In the statement of disturbance models the normal law was interpreted in polar coordinates by the folded half-normal distribution; that is, the distribution function for location about a lattice site is

$$F(\xi) = F(\rho, \theta) = \int_0^\rho \int_0^\theta f(\rho)\, f(\theta)\, d\rho\, d\theta \tag{17}$$

where

$$f(\rho) = \sqrt{2}\exp(-\rho^2/2\sigma^2)/(\sigma\sqrt{\pi}) \qquad \rho > 0$$
$$f(\theta) = (2\pi)^{-1} \qquad 0 < \theta < 2\pi.$$

It seems appropriate to accept that $f(\xi)$ is identical for each lattice site so that the parameter σ is constant throughout the lattice space. Using (17) to define (14) and substituting the resulting probability transformation into (16) gives an expression for order statistics that, for me, is totally intractable.

Some simplification is gained by interpreting the normal law by the bivariate or circular normal distribution. In this case the distance variable X has a well known form. It may be shown that the distribution function is

$$F(x) = 1/2\exp(-t^2/2\eta^2) \int_0^{(x/\eta)^2} e^{-x/2} I_0(tx^{1/2}/\eta)\, dx \quad x > 0 \tag{18}$$

where $\eta = 2\sigma^2$ and $I_0(\bullet)$ is the modified Bessel function of the first kind of zero order. This expression is recognized as the integral of the non-central χ^2 with two degrees of freedom. In a slightly different form it occurs as a basic distribution function in bombing or coverage problems, Germond (1950). By substituting (18) for $F(\omega)$, (16) gives the z crude moment of order statistics from a non-central χ^2 distribution; however, tables of values have not been published.

It is apparent that even the simplest imperfection model yields equations that are difficult to evaluate. Where $\lambda \neq 1$ and/or $\gamma \neq 0$ the equation systems are immensely more complex and numerical evaluation may be considered, for any practical purpose at this time, impossible. In order to circumvent these mathematical problems the imperfection model has been evaluated by simulation of an equation system for a given set of parameter values.

Analysis of the Pattern of Urban Places in Iowa

The imperfection models were designed to produce types of patterns and distributions studied in the social sciences. Moreover, the particular class of patterns motivating the present formulation are formed by map representations of urban places. As a partial evaluation of the adequacy of the imperfection model to replicate town and city patterns, the distribution of urban places in Iowa, 1950, is studied.

Many parameters of the Iowa distribution are already available in Dacey (1963a). These data provide empirical estimates of parameters for application of the imperfection model to the Iowa pattern. Using estimated parameters, the degree of correspondence of M_2 with the observed pattern of urban places is analyzed. Simulation is used to evaluate the theoretical imperfection model.

Almost Periodic Disturbance Model

The almost periodic disturbance model M_2 is specified by three sets of parameters:

t_1, t_2 and g identify the underlying lattice P,

ξ specifies the disturbance term generating the point set S_2 and

λ and γ are the scale densities for the point sets S_2 and R, respectively.

These three sets of parameters are given numerical values by relating the imperfection concept to structural features of the Iowa map pattern. In this construction, each parameter is described in terms of the corresponding property of the Iowa pattern. Since the theoretical pattern is synthetically fabricated, the definitions and interpretations of parameters are biased toward operational statements.

Lattice Parameters

The M_2 is fabricated as a rectangular map space containing the domain of a square lattice. The domain is of dimensions 12 by 18 and contains 96 points. Thus, the parameters are $t_1 = t_2 = 1$, $g = \pi/2$.

The primitive cells of the square lattice have an abstract correspondence to counties, and in this context lattice points represent the geographic center of counties. This lattice has some resemblance to the Iowa map. In gross form Iowa is roughly a rectangle and most counties in Iowa are approximately square. However, the counties do not form a square grid, largely because of surveying adjustments for the earth's curvature. An alternative, and possibly a closer, approximation to the Iowa structure is the diamond lattice.

The lattice has 96 squares while Iowa has 99 counties. There is no formal advantage to using a lattice of approximately the same dimensions as the study area.

For specification of other parameters the following relations are established between M_2 and the Iowa map:

i. square lattice cells of M_2 are equated with Iowa counties,

ii. lattice points of M_2 are equated with geographic centers of counties,

iii. S_2 and R points are equated with urban places.

Using this dictionary (α) the distribution function for distance from lattice site to S_2 point is estimated from the observed distances from geographic center of counties to nearest urban place and (β) the frequency distribution of points in primitive lattice cells is estimated from the observed frequency distribution of urban places in counties. These two properties are evidently independent of the order distances used to summarize observed and theoretical patterns.

Disturbance Variables

In my earlier study of Iowa it was shown that for interior counties containing an urban place the distance from the geographic center to nearest urban place was closely approximated by the folded half-normal distribution, as defined for $f(\rho)$ in (17), with scale parameter $\sigma = 0.2286$. Observed and calculated frequency distributions are compared in Table 1.

The angular component θ of the disturbance term is taken as a uniform random variable, as defined in (17). No evidence is presented for this assumption, so the uniform variable is entered into the model on the theoretical consideration that a completely chance factor occurs in the disturbance process. However, in examining the location of places with respect to geographic centers I found no evidence of directional bias.

On the basis of these estimates, the vector component ρ and the angular component θ of the disturbance variable ξ are defined for M_2 by the folded, uniform bivariate distribution (17).

Scale Variables

The remaining two parameters of M_2 are the density measures λ and γ. Because M_2 contains only S_2 and R points, the density of all points is $\mu = \lambda + \gamma$. For the Iowa map pattern there are 93 places and 99 counties, so the estimated density of total points in M_2 is $(93/99) = \mu$.

The individual densities λ and γ were estimated from the frequency distribution of urban places among Iowa counties, Table 2. A two parameter probability density function that gives a good fit to the observed frequencies has been stated by Dacey (1963b). By assuming that each disturbed point in S_2 is always located in the primitive cell of its theoretical lattice site and that each random point in R has an equal probability of occurring in each primitive cell, the probability that a cell contains x points is

$$f(x;\lambda,\mu) = (\gamma^{x+1}e^{-\gamma}/x!) + (x\lambda\gamma^{x-1}e^{-\gamma}/x!) \tag{19}$$

where $\gamma = \mu - \lambda$ and $x = 0, 1, \ldots$. The parameter λ was estimated by the method of moments from the distribution of urban places among Iowa counties. Table 2 compares observed and expected frequencies for the parameters $\lambda = 0.74$, $\gamma = 0.20$ and $\mu = 0.94 \cong 93/99$.

Comparison of M_2 and Iowa

A synthetic pattern was constructed from the pattern M_2 for the parameters

$$t_1 = t_2 = 1 \quad g = \pi/2$$

$$\sigma = 0.2286 \quad \lambda = 0.7396 \quad \gamma = 0.1979.$$

These parameters were applied to a space containing 96 lattice sites, so that M_2 contained 71 S_2 points and 19 R points. Tables of random digits and standard normal deviates were used to generate a synthetic M_2. Because of the small pattern size, random digits and normal deviates were tested for randomness.

The M_2 and Iowa patterns were described by (i) distances from origin points to the 10 nearest neighbors and (ii) distances from loci to the 10 nearest points. The boundary constraint was applied so that the number of recorded measurements tends to decrease as the order of neighbor increases.

Order mean distances are listed in Table 3 for point to point measurements and in Table 4 for locus to point measurements. The tabulated data on M_2 give mean distances for the 10 lower order neighbors and the number of recorded measurements for each order. Distances obtained from the Iowa map were standardized by multiplying each observed mean order distance by the square root of the density of urban places. The tabulated data on Iowa give the standardized mean distances and approximate miles for the 10 lower order neighbors. Also tabulated are the absolute and percentage differences between the observed and calculated mean order distances. Many other properties of M_2 and Iowa were collected but are not included in this report.

There are many reasons for not conducting an elaborate analysis for goodness-of-fit of the M_2 data to the Iowa data. Important reasons include the small size of the fabricated M_2 and difficulty in transforming frequency distributions into the normal form. These and similar problems could, largely, be handled in a more careful experimental design. More control was not exercised because I wanted a fast, crude evaluation of an imperfection model to determine whether it possessed any empirical reference, and, hence, merited detailed consideration. A fair test of the imperfection approach to urban systems requires a substantially more sophisticated model than M_2

Though recognizing the 'imperfections' in M_2, it seems sufficiently provocative to justify release of this highly preliminary report. While statistical methods were used to evaluate hypotheses of no difference between M_2 and Iowa (which were not rejected by the available data), reports on levels of significance and other statistical findings do not seem particularly critical at this stage of development.

Evaluation

The synthetic pattern M_2 reproduces with considerable fidelity the Iowa map pattern of urban places. The correspondence between M_2 and Iowa is a statistical rather than a cartographic similarity. This criterion of similarity determines the type of conclusions that can be drawn from the present study.

Both patterns were summarized by sets of distance measurements. These distances represent, however, quite different conceptualizations. The Iowa pattern refers to an observed distribution that exists in the real world, and at a point in time a study area has a single pattern of urban places. In contrast, the synthetic pattern represents a probabilistic model that is an abstract construction. This model does not describe one map pattern. Instead, the model defines a set of theoretical values. It is possible to interpret the model and synthetically construct a pattern that is representative of the model; yet, the model generates only one of an infinity of different patterns that correspond precisely to the statement of the model.

In more formal terms, the reduction of the distribution of urban places to order distances

in a one-to-one mapping but the reduction of the model to a pattern is a one-to-many mapping. So, for the Iowa distribution only one pattern is formally possible (all representations must be conformal) while the mapping of the model is multi-valued. Consequently, while a single map describes the Iowa pattern, there is no cartographic summary of the pattern contained within the theoretical model.

While we reduce a map to a set of numbers we do not return a corresponding set of numbers to the map form. The cost of reducing the Iowa map pattern to a system of equations describing an imperfection model is the loss of the map description of that pattern. Whether this loss is compensated by the substantially greater analytical utility of a mathematical construction is a question that each student must resolve for himself.

In evaluating these questions the role of simulation should be correctly interpreted. Simulation was used only after all parameters of the model were estimated. This is not general in social science investigations of large, complex systems by means of simulation. Often, the model is simulated many times, each run using a different set of parameter values. The model being simulated is then adjudged successful if some set of parameters provides a good fit to the data at hand. This iterative approach is based upon an a priori acceptance of the model. In this application the simulation is used primarily to study properties of a complex model, but it does not provide any independent means of verifying the model itself. Simulation was not used for this purpose; for the imperfection concept simulation serves as the poor man's (mathematically poor, that is) numerical integration of a completely specified probabilistic model which can not be evaluated by analytic methods.

Table 1

Frequency Distributions of Observed and Calculated Standardized Distances, c_1, from Geographic Center of Interior Counties Containing an Urban Place to Nearest Urban Place

Distance c_1/σ	Freq. f_0	Dist. f_c	Error $f_0 - f_c$	$\frac{(f_0-f_c)^2}{f_0}$
0- .243	11	11.72	- .72	0.471
- .486	11	11.04	- .04	0.000
- .729	11	9.82	1.18	0.127
- .972	8	8.23	- .23	0.005
-1.215	6	6.51	- .51	0.237
-1.458	3	4.85	*-1.85*	*0.052*
-1.701	5	3.39	*1.61*	
-1.944	2	2.28	**- .28**	**0.265**
-2.187	2	1.41	**.59**	
-2.430	2	.83	**1.17**	
Over 2.430	0	.92	**- .92**	
Total	61	61		1.157 ($\equiv \chi^2$)
				df=4
				$.90 > Pr(\chi^2 = 1.157) > .75$

Iowa data, f_0 from Dacey (1963a). The standard deviation is $\sigma = 0.2286$. The calculated frequency, f_c, is from the unit half-normal distribution.

Table 2

Comparison of Observed Distribution of Urban Places per County in Iowa, 1950, with Expected Distribution of Points per Primitive cell of M_2

Number of Places x	Frequency Distributions Observed $g(x)$	Expected $E(x)$
0	21	21.1
1	64	64.2
2	13	12.4
3	1	1.2
≥ 4	0	.1

Observed values are from Dacey (1963a). Expected values are computed from (20) with $\lambda = .74$ and $\gamma = .2$.

Table 3

Comparison of j Order Distances for M_2 and Iowa Maps

Order	M_2		Iowa		Error	
j	n_j	$\bar{r}_j$	$d_0^{1/2}\bar{R}_j$	Mi.	$\bar{r}_j - d_0^{1/2}\bar{R}_j$	As % of Iowa
1	65	0.63	0.66	16	-.03	4.7
2	58	0.84	0.84	21	.00	
3	56	0.98	0.99	25	-.01	1.4
4	55	1.12	1.12	28	.00	
5	53	1.24	1.24	31	.00	
6	46	1.35	1.36	34	-.01	1.0
7	44	1.46	1.49	37	-.03	2.1
8	41	1.54	1.60	40	-.06	4.0
9	37	1.65	1.68	42	-.03	2.0
10	36	1.74	1.78	44	-.04	2.0

Iowa data are from Dacey (1963a).

Table 4

Comparison of h Order Distances for M_2 and Iowa Maps

Order	M_2		Iowa		Error	
h	n_j	$\bar{r}_h$	$d_0^{1/2}\bar{R}_h$	Mi.	$\bar{r}_h - d_0^{1/2}\bar{R}_h$	As % of Iowa
1	40	0.42	0.41	10	.01	4.7
2	36	0.72	0.72	18	.00	
3	32	0.97	0.93	23	.04	4.2
4	31	1.07	1.13	28	-.06	4.8
5	29	1.21	1.26	31	-.05	4.0
6	28	1.32	1.39	35	-.07	4.8
7	28	1.43	1.45	36	-.02	1.8
8	27	1.55	1.56	39	-.01	0.8
9	22	1.62	1.65	41	-.03	1.6
10	20	1.71	1.74	43	-.03	1.9

Iowa data are from Dacey (1963a).

October 14, 1963 Philadelphia, Pennsylvania

This original paper by Dacey, when printed in the *Papers* of the Michigan Inter - University Community of Mathematical Geographers, was supplemented with an 'Addendum' reflecting computer programs current at the time by Professor Duane F. Marble and Mr. Marvin Tener, and a second examination of the Iowa data by Dacey (December 13, 1963). A Glossary by Nystuen offered expanded explanations of complicated material for readers uncomfortable with notation. The added materials are not reprinted here.

* The support of the Regional Science Research Institute and of the National Science Foundation is gratefully acknowledged.

** Current address: Department of Geography Northwestern University Evanston, IL

References

Christaller, W. 1933. *Die zentralen Orte in Süddeutschland.* Jena: Fischer.

Dacey, M. F. 1963a. *Iowa: The Classic Plane or Croupier's Table.* Mimeographed.

Dacey, M. F. 1963b. *A Poisson-Type Distribution for Dispersed Population.* Mimeo.

Dacey, M. F. 1963c. "The Status of Pattern Analysis: Identification of Problems in the Statistical Analysis of Spatial Arrangement," paper presented at the Regional Science Association meetings, Chicago, 1963.

Dalenius, T., J. Hajeck, and S. Zubrzycky. 1961. On Plane Sampling and Related Geometrical Problems. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1. Berkeley: University of California, 125-150.

Feller, W. 1957. *An Intorduction to Probability Theory and its Applications, Vol. I.* New York: Wiley. Second edition.

Germond, H. H. 1950. *The Circular Coverage Function.* Santa Monica: RAND, Memorandum 330.

Lösch, A. 1939. *Die Raumliche Ordnung der Wirtschaft.* Jena: Fischer. (Translated by W. H. Woglom and W. F. Stolper as the *Economics of Location*, New Haven: Yale University Press, 1954.)

Seitz, F. 1952. "Imperfections in Nearly Perfect Crystals: A Synthesis," *Imperfections in Nearly Perfect Crystals* (W. Schockley, J. H. Hollomon, R. Maurer, F. Seitz, eds.) New York: Wiley.

Van Bueren, H. G. 1961. *Imperfections in Crystals.* New York: Interscience.

Zubrzycki, S. 1961. "Concerning Plane Sampling." *Second Hungarian Mathematical Cong.* Budapest: Akademiai Kiado.

Readers of *Solstice* might also be interested to note the following additional references to Dacey's work, not noted in the MICMOG publication.

Dacey, M. F. 1960. A note on the derivation of nearest neighbor distances, *Journal of Regional Science*, 2, 81-87.

Dacey, M. F. 1960. The spacing of river towns, *Annals* Association of American Geographers, 50, 59-61.

Dacey, M. F. 1962. Analysis of central place and point pattern by a nearest neighbor method, *Lund Studies in Geography* 24, 55-75.

Dacey, M. F. 1963. Order neighbor statistics for a class of random patterns in multidimensional space, *Annals* Association of American Geographers, 53, 505-515.

Dacey, M. F. 1963. Certain properties of edges on a polygon in a two dimensional aggregate of polygons having randomly distributed nuclei. Mimeo.

Dacey, M. F. 1964. Two-dimensional random point patterns: A review and an interpretation, *Papers*, Regional Science Association, 13, 41-55.

Dacey, M. F. 1964. Modified Poisson probability law for point pattern more regular than random, *Annals* Association of American Geographers, 54, 559-565.

Dacey, M. F. 1965. Order distance in an unhomogeneour random point pattern, *The Canadian Geographer*, 9, 144-153.

Dacey, M. F. 1966. A compound probability law for a pattern more dispersed than random and with areal inhomogeneity, *Economic Geography*, 42, 172-179.

Dacey, M. F. 1966. A county seat model for the areal pattern of an urban system, *Geographical Review*, 56, 527-542.

Dacey, M. F. 1966. A probability model for central place location, *Annals*, Association of American Geographers, 56, 550-568.

Dacey, M. F. 1967. Description of line patterns, *Northwestern Studies in Geography*, 13, 277-287.

Dacey, M. F. 1968. An empirical study of the areal distribution of houses in Puerto Rico, *Transactions*, Institute of British Geographers, 45, 15-30.

Dacey, M. F. 1969. Proportion of reflexive n-th order neighbors in spatial distributions, *Geographical Analysis*, 1, 385-388.

Dacey, M. F. 1969. A hypergeometric family of discrete probability distributions: Properties and applications to location models, *Geographical Analysis*, 1, 283-317.

Dacey, M. F. 1969. Some properties of a cluster point process, *Canadian Geographer*, 13, 128-140.

Dacey, M. F. Similarities in the areal distributions of houses in Japan and Puerto Rico, *Area*, 3, 35-37.

Dacey, M. F. 1973. A central focus cluster process for urban dispersion, *Journal of Regional Science*, 13, 77-90.