5

Archaeological applications of the 'Percolation Method' for data analysis and pattern recognition

A. De Guio*

G. Secco†

5.1 Introduction

'Percolation' is a highly formalized and expanding field of data analysis within the larger context of numerical taxonomy and automatic pattern recognition (Tremolliers 1984). Tremolliers has already presented several different versions of the original procedure ('Percolation Normale', 'Percolation Généralisée' and ' Percolation Equilibrée), each of which demonstrates a trend towards increasing complexity (Tremolliers 1979, Tremolliers 1981, Tremolliers 1982, Tremolliers 1984).

It is not our intention to reiterate the entire set of supporting assumptions, definitions, relevant technical terminology and the long series of formulas involved in this process, but rather to emphasize the key concepts and main procedures. In other words the level of detail used to illustrate the Percolation perspective is the minimum required to create a generic frame of reference, intended as a starting point for understanding our subsequent research design. This ultimately enters into a very specific archaeological field of enquiry and therefore retains very little of the original methodology.

Initially Percolation relies on the basic assumption that there is a direct connection between the 'natural' notion of 'groups' and the identification of 'uni-modal zones' within multi-dimensional distributions. Thus, given a set of values (or simply 'points', as defined in the conventional working approximation given by the author (Tremolliers 1984)), Normal Percolation identifies the underlying structural organisation of such distributions by dividing these points into three different classes (taxa):

* Istituto di Archeologia,
Università di Padova,
P.zza Capitaniato n. 7, 35100 Padova
Italy

† Dipartimento di Geografia,
Università di Padova,
Via del Santo n. 26,
35123 Padova
Italy

1. 'Group points' (G), or 'modal points', which constitute the subset of points x attributable to a single mode within the data;
2. 'Frontiers points' (F), or the subset of points x attributable to two or more modes;
3. 'Isolated points' (I), or those points not attributable to any mode.

The partitioning algorithm relies on a density function ('fonction de densité'), related to a certain 'perception threshold' ('seuil de perception'). If we imagine the simpler case of a bi-dimensional universe I (a set of points, such as those displayed by the starting 'map' of unweighted points of our case study in Fig. 5.1), we can easily reckon for each point i a score

$$P_i = \Sigma_{j=1}^{k} W_j$$

where j=element of I and W = a weight always having a value of 1 ('unweighted density'), and located at a distance $D_{ij} \leq T$, where T='Perception Threshold' (in other words the radius of a circular area for density calculation centered on i). If we now imagine such a density function as the third dimension built upon our starting bi-dimensional landscape, we can easily approach the key concept of Percolation: the existence of a multi-modal distribution with the possibility to attribute each point to one or more 'modes' on the basis of an intuitive clustering/partitioning algorithm, close to our capabilities of basic perceptual pattern recognition. A single point should be ascribed to a group because there exists a related chain of points of decreasing density scores, starting from the mode which characterizes the group itself. The procedure is very simple:

1. rearrange the total number of P_i scores (=n) in decreasing order (P_1 = highest; P_n =lowest);
2. the first highest P_i (= P_1) characterizes the mode of the first group G (G_1);
3. identify the second highest P_i (= P_2). If the distance between the related point (group) 1 and the trial point 2 is less than or equal to T ($D_{12} \leq T$), then point (that is element) 2 is ascribed to G_1, otherwise it begins to form a new group G_2;
4. repeat step 2 for each decreasing P_i up to P_n and attribute the points to old groups or form new groups, according to the same rule in relation to the nearest neighbour of an already formed group G.

The simple path is, of course, complicated by the possible existence of 'Frontier Points' (F), single, multi-modal points at the boundary of two or more unimodal groups G, that is at the same distance from them.

It seems useful, in addition, to note as 'Isolated Points' (I) those points with no other points within a distance T (the density radius) from themselves.

Generalized Percolation offers the further capability of discovering a new (although rare in real cases) class of points as an extension of the previous class: the 'Multi-Modal Points' (M), that is those subsets of homogeneous (equal-density) points, which act as a bridge('no-man's land') between two contiguous-connected groups G.

Other relevant areas in the course of development are subsequently grouped under the rubric of the new 'Equilibrated Percolation'. These attempt to combine the proper

Percolation approach (the stepwise individuation of modal groups) with other dynamic-allocation algorithms, such as those based on centroid or nearest neighbour clustering procedures, involving possible rearrangement and the dissection of previously aggregated uni-modal groups.

Without offering further details, it is important to summarize here the distinguishing features common to the entire concept of Percolation building. Initially we must recognize its principal focus as that of 'pattern recognition', intended to detect 'natural' grouping on the basis of multi-modal distribution of density values. Tremolliers (Tremolliers 1984) views such an analytical device as amounting to a sort of automatisation of visual perception, open to future development within the realm of computer-vision (Artificial Intelligence).

The mechanics of the clustering and partitioning algorithm (at least in the simpler three-dimensional perspective stressed above), can be easily recoded with little difficulty as a pictorial simulated landscape. We have just to imagine a completely submerged morphological panorama (x,y coordinates and the density function as z), which 'discovers' new groups and/or bonds to old-ones (on the basis of density and threshold distance values), while the water level is lowered (or 'percolates') from the highest 'peaks' to the bottom. The apparent simplicity and lack of sophistication of the model do not affect its high potential level of analytical performance in a wide spectrum of field-areas. It could indeed be directly applied on to any type of multi-dimensional data within the archaeological domain, from numerical taxonomy of material culture to inter/intra- site spatial analysis. In the latter case, a promising field, we are employing this approach to recognize contiguous (often blurred) space activities within excavated and ploughed surface contexts. We are using different 'strata' of Percolation analysis, each restricted to a specific type of artefact or ecofact involved with different degrees of actual discriminant spatial and functional capabilities.

It is not, however, our intention to thoroughly investigate the potentially infinite analytical horizon open to the Percolation method. Ultimately, we are dealing with with a new and general class of cluster analysis, able to partition a given universe into discrete, operational taxa, the only major underlying theoretical assumption being a previous recognition of density as a workable and meaningful device for structuring the universe itself.

Our Interest, indeed, restricted to an extremely limited domain, which we were actively examining when we initially encountered the Percolation perspective. This is the so called 'Landscape of Power' (henceforth 'L. of P.').

G. Secco

5.2 Percolation and 'landscape of power'

As a matter of convention we would like to include under the mnemonic heading of 'L. of P.' (Renfrew 1984, pp. 23–77) an emergent research area located at the intersection of political anthropology, geography and archaeology (cf. e. g. Alden 1979, De Guio 1985a, De Guio *et al.* 1986, Friedman & Rowlands 1982, Hodder 1979, Hodder & Orton 1976, Johnson 1981, Renfrew 1982, Renfrew & Cherry 1986, Renfrew & Level 1979, Steponaitis 1981, Voorips 1981). This involves the integration of formal mathematical models for analysis, pattern recognition and the simulation of ancient political organisations in space, using archaeological data as a starting point. It is not the case, in this instance, to reiterate the entire corpus of theory building and the relevant literature associated with it. It suffices to recall Renfrew and Level's (Renfrew & Level 1979) development of a proper axiomatic foundation of the field, leaving to a later discussion some critical reservations and additions. The building blocks of such a theoretical construct can be reduced to six principal axioms:

1. the human social group defined as the habitual association of persons within a territory;
2. the segmentary-modular nature of human social organisation of space;
3. the tendency of basic social groups to affiliate into larger groups;
4. the hierarchical status of human society and its derived stratified spatial organisation;
5. the possibility of identifying the effective polity, the highest order social unit, on the basis of scale and distribution of central places;
6. the possibility of finding uniformities in artefact distribution, due to special interactions between polities, which are not documenting societies or peoples.

In addition Renfrew and Level have presented four detailed assumptions:

1. the expected spatial continuity of polities, without intervening parcels of land;
2. the likelihood that each parcel of land (with the possible exception of 'no-man' s land') will fall under the 'jurisdiction' of a single autonomous authority, or its deputies within a hierarchical resolution of power.
3. the highest probability of identifying polity 'capitals' as the largest settlements in their territories;
4. the expected positive correlation between the 'size' of the the autonomous 'capital' centre and the territorial area under its jurisdiction (Renfrew & Level 1979, pp. 145–146).

Such a frame of reference is no more than a provisional setting, in anticipation of a fully integrated and detailed archaeological theory of L. of P. It is sufficient, however, to assert the identity of L. of P. as an autonomous field of investigation. The axioms and assumptions are, in addition, strictly related to the specific simulation model presented by Renfrew and Level ('XTent' model). This pertains to that subset of formal models within the L. of P. domain, which we could label properly as 'spatially based'. This fully exploit a large amount of possibly relevant sources of spatial information: such as 'size', inter-site distance and distributional patterns. Among these we could mention a handful, drawn from an already wide repertoire based, for instance, on 'Central Place' theory', 'Thiessen Polygons', 'Gravity theory ', 'Information theory' and even 'Catchment Area Analysis'(cf. e.g. Alden 1979, Renfrew & Level 1979, Hodder & Orton 1976, Steponaitis 1981). The model presented here (shortened to 'Percol' model) and some others implemented in recent years ('Tect' and 'Top' models: De Guio 1985a, De Guio *et al.* 1986) fall within this type of approach and share a wide spectrum of declared or inherent conventions and approximations. Among these are included the search for simpler 'dominance' relationships or the more ambitious design of recognizing different nested arrangements of power in space ('modules' with critical within group/between groups rates of similarity).

In our estimation the unique perspective of Percolation appears to address two rather neglected sources of spatial information within the L. of P. literature, density as well as its multi-modal distribution, displayed by relevant entities (normally sites or meaningful parts of them, such as monuments, or upper order modular inter-site units). Within this context, the first trial hypothesis assumes that the spatial arrangement of density values within a L. of P. should demonstrate some constraints and regularities. In particular we should expect some degree of homomorphism between modal zones (G groups) and possible patterns of modular resolution of power.

The partial failure of such expectations compelled us to re-orient the research design towards new and rather original approaches that sometimes retained very little of the original Percolation method.

A. De Guio

5.3 The 'Percol' model: new functions

What remains of the original Percolation method is essentially the recognition of different modal groups (G), given a certain function and a certain 'distance threshold' (T) related to our 'points'. On the contrary, the relevant additions and modifications involve a number of key areas.

Initially different functions are required instead of the simple, unweighted density function. In effect a direct-unmodified application of the original Percolation method, based on such a function, produced what immediately appeared to be meaningless results in terms of our L. of P. case study analysis. This is largely due to the fact that a L. of P. is far more then a 'perceptual' partition into modal groups. We subsequently experimented with nine different trial functions, borrowed in part from a well established geographical-analytical literature (drawn primarily from the fields of gravity and population dynamics), as well as from newly introduced formulas, directed towards the specific goal of a L. of P. simulation. Here we present four of them, of which only the final one ('Dominance') is referable to the 'original' subset.

1. 'Weighted Density' (WD), which is still a density function, but, at least a weighted one. In effect for each 'point' i (such as a site), it summarizes a related weight W_j (for instance a size value of the settled area or the estimated population of site j):

$$WD_i = \Sigma_j^m W_j$$

   where m is the number of sites within an assumed T 'distance threshold' from the site i;

2. 'Demographic energy' (DE):

$$DE_i = K * P_i * \Sigma_{j=1}^n P_j / D_{ij}$$

   where K is a constant, P_i and P_j are the population of sites i and j, D_{ij} is a 'distance' value between them, and n is the total number of sites;

3. 'Potential of Population' (PP):

$$PP_i = K * \Sigma_{j=1}^{n} P_j / D_{ij}$$

The last two formulas are borrowed from the 'social physics' analytical literature related to the study of population dynamics (cf. Zanetto 1979). In particular we rely upon a number of results attained by Steward and Warntz (Steward & Warntz 1958), who integrated previous theoretical issues into a coherent body, by demonstrating how some important demographic indexes not only related one another, but also how they related isomorphically to Newton's law of gravitation, and even more closely to Lagrange's subsequent enhancements of this. In particular (Warntz 1964, p. 174; we use here his original notation), the interrelatedness of three key measurements can be demonstrated as follows.

The 'Demographic Force' (F) of attraction between two population groups (P_1 and P_2) at a distance r from one another is: $F = K * P_1 P_2 / r^2$, where K is a demographic gravitational constant left to future determination, and often kept to a value of 1 in the majority of real case-studies (cf. for instance Reilly's 'Index of Attraction' and break-event point determination: Reilly 1929). Their 'Demographic Energy' (E) is: $E = K * P_1 P_2 / r$ (cf. for instance Zipf's 'Index of Interactance': Zipf 1946). The 'Potentials of Population' (V) are respectively: $V_1 = K * P_2 / r$ and $V_2 = K * P_1 / r$ (as regards the problem of 'self potential' Warntz suggests the use a r value equal to one-half the radius of the population group area assumed to be circular);

The interrelation then becomes:

$$2E = P_1 * V_1 + P_2 * V_2$$

The use of the two formulas ('Demographic Energy' and 'Potential of Population') for L. of P. simulation purposes involves a key assumption, (which we could better consider a working hypothesis to be tested in real case-studies) of an underlying homomorphism between two distinct classes of spatial-functional behaviour: 'power' on the one hand and demography on the other. If we consider the dominance (and especially the hierarchically-modular dominance) process as ultimately reducible to a tendentially 'minimum effort' (optimizing) phenomenon of energy and information transfer, the ranked scores of the two demographic indexes, measuring the inter-site attraction and potential accessibility in terms of optimal population transfer, should be expected to reflect at least to some degree the hierarchical resolution of power in space. In order to avoid any naïve expectations in this regard, we have to consider not only an abstract functional (and basically synchronic) perspective, but also the potentially divergent (especially in terms of relative 'inertia') diachronic behaviour of the two spheres of interaction. This could amount to a build-up through time of two progressively distinct 'morphogenetic landscapes' (L. of P.. and demographic landscape), displaying an increasing loosening of an assumed common homomorphic matrix;

4. 'Dominance' (DM). The formula represents a substantial revision of some of our previous formal approaches to the L. of P. simulation (De Guio 1985a, De Guio *et al.* 1986), with which it shares a number of axioms and assumptions, in

particular 'dominance' as a function of some 'size' value and of some inter-site 'distance' value:

$$DM_i = K * W_i * \Sigma_{j=1}^{n} 1/D_{ij} * (W_i - W_j)/(W_i + W_j)$$

where is W = the 'size' value and D_{ij} = the inter-site 'distance', K = a constant ('coefficient of power': we imagine that it should assume higher values for higher levels of integration of the political dimension. It is, however, left to future determination and maintained here at a trial value of 1 as in the demographic formulas) and n is the total number of points (settlements). Each partial score can thus be positive, negative or null, suggesting a working analogy (and no more than that), respectively with possible statuses of positive and negative dominance (and even possible degrees of these, related to the absolute values of the score) or no-dominance (independence): basically, the shorter the distance and the higher the positive difference in size, the higher the resulting positive 'dominance'. The W_i component acts as a weighting (and therefore ranking) factor throughout the summation procedure.

The above four formulas can be split usefully into two relevant categories, 'local' (WD) and 'global' (the others). The first class in fact produces local values from a local constraining basin of information (or neighborhood). In other words it can assume different values according to the variation of the T parameter, which is used in this case both for the density size unit and for the clustering algorithm. On the contrary the second class produces a single set of values for the function involved. Every element (or more directly, 'site', in our L. of P. frame of reference) of the universe assumes the whole universe itself (that is all the n sites under examination, not just those of an assumed neighborhood), as the channelling information basin. The different scores, given such an 'equal opportunity' scheme, result from size and location.

A. De Guio, G. Secco

5.4 The 'Percol' model: new algorithm and computer program

Another key difference in comparison to the original Percolation model pertains to the clustering algorithm. Instead of using just one T value for cluster formation (which involves the logical justification of our choice criterion), we assume all its possible discriminant values, that is only those values able to partition our universe into different arrangements. Therefore, considering the T range of variation, we will employ discrete values corresponding to increasing values of the D_{ij} (inter-site distance) range of variation. This range extends from a minimum (corresponding to the minimum D_{ij}), up to a maximum (in this instance normally far below the maximal D_{ij}), corresponding to a value at which the clustering algorithm ends to form just one covering group G. The supporting algorithm thus become a stepwise process, producing at each step a differential partition (or 'Percolation landscape') of the universe.

In the case of a local function such as our WD, the values of T are, at each step, the same used for calculating the function itself and for forming the clusters.

Moving into the L. of P. problem area and, therefore, into our starting set of axioms and assumptions, we decided to rely upon two of the group categories used by the various Percolation approaches: G groups (or 'modal groups') and I points (or 'isolated points'). The recognition of Frontier Points (F) and Multi-modal Points (M, or 'bridge-points'), as meaningful categories in terms of L. of P. (these, as potential indicators, respectively, of inter/intra-polity frontiers and no-man's land situations), although attractive at first glance, can be quite hazardous. As an initial approach, we prefer to consider dominance as a basically dichotomous entity, so that every 'site' must be attributed in each case to a polity or to a module within it. This is accomplished by relying, in the case of matching scores not only upon the nearest neighbour, but also upon the second or third, and so on. This continues up to the first appearance of a differential score for a discriminant attribution to one of the 'rival' groups (in the extreme case—never encountered—of a persistent matching, the decision is taken on the basis of the summation of the relative sizes of the same groups).

The computer program which implements our Percol model was written by G. Secco on a Data General MV 6000 minicomputer located by the Department of Geography of the University of Padua, using the FORTRAN 77 language. It requires a data input file where every record is made up of a case identifier variable (sequence number or character variable) for the case, x, y coordinates, and a variable number of weighting variables (such as size value for sites in different chronological phases, with a 0 value denoting a site not settled in the relevant phase). For everyone of the four above-mentioned functions in each step of the cluster formation process, there is an output displaying the T value, total number of groups formed, and a matrix of case values (case identifier, x, y coordinates, weight, score) subdivided by groups (with a group identifier and a group counter), with each group ordered according to decreasing score values. At the end there is a summary with a sequence of paired values given for each step, which represent the total number of groups formed and their relative T value. The C.P.U. time required, using a data-base of 90 cases (the highest figure run by us thus far) and a single function is an average of 4 seconds for each step. The overall cluster formation path, in the case of the 'global' functions, which produce an hierarchical partitioning process, is represented by a dendrogram graph (Figs. 5.7, 5.8) with a vertical axis corresponding to increasing (from top to bottom) values of the T parameter.

G. Secco

5.5 A case-study

The Percol model has been tested in the environs of Uruk, in Southern Mesopotamia, between the Early Dynastic II–III and the Old Babylonian Periods (cf. the composite map in Fig. 5.1). The data-base is drawn from Adams' principal study of this area (Adams 1981), derived from Adams and Nissen's survey campaigns. This 'heartland of cities' is a primary focus for the study of urban and state formation processes, and has been investigated in one of the most impressive long term survey projects carried out to date (cf. Ammerman 1981). Consequently, it presents an extremely relevant set of field data for our L. of P. approach.

Our principal aim, however, is not an attempt to offer a substantial contribution to the socio-political history of this region. Instead, we hope to provide a new body of formal constructs against which Mesopotamian scholars might compare a wide spectrum of existing information drawn from a variety of field and textual data.

Among the six succeeding Periods, we have chosen to emphasize the third one, the Uruk–Isin–Larsa, which presents the greatest number of sites with a wide variation in size, and is therefore likely to be one of the most promising as regards the spatial resolution of power. This selection was further influenced by our previous analytical work here, concerning demographic trend analyses and, in particular, the testing of other types of L. of P. simulations (De Guio 1985a).

In order to fit Adams' data to the prerequisites of our model we have made the following decisions:

1. the 'size' of a settlement (hectares of settled area: cf. Figs. 5.2–5.4) is assumed to correspond to the central values of a seven-rank classification system offered by Adams (from 'trace', to which we have given the minimum conventional value 1, up to 200+ ha: cf. Adams 1981, Tab. 14);
2. the inter-site 'distance' is assumed to be linear and is measured on Adams' principal reference map;
3. the population estimates are made on the basis of Adams' assumption of 100 persons per ha (Adams 1981, p. 69);
4. the K constant of our functions is kept to a value of 1.

The relevant results of the Percol approach are illustrated here by a small range of statistical and iconic models (cf. Figs. 5.5–5.14). Initially we must examine the overall stepwise process of cluster formation. The above-mentioned dendrogram scheme (cluster against T values on the vertical axis—cf. Figs. 5.7–5.8) allows us to recognize the general rate and trend of such a formation process. We can further denote the detailed 'path' followed by a single site (which constitutes an independent group G in its 'leaf' status at the top of the dendrogram and is ultimately fused into a final all-encompassing group at the 'root' of the tree), or group of sites, wherever such a group is formed, and is then absorbed into a major cluster situated along the T dimension.

Every group contains an internal ranking corresponding to the scores of the function, and a 'leader' (the most scored site), which represents a mode in the overall multi-modal 'Percolation panorama'. The dendrogram model is used exclusively for 'global' functions, as these produce a single set of scores, and therefore forms a hierarchical partitioning configuration quite suitable to the dendrogram model itself, as well as to a 'modular' L. of P. simulation approach.

The branching process of group formation appears to suggest some kind of analogy with a 'recognition path' of a L. of P. We should, however, climb such a tree with caution, avoiding any naïve pretense of finding an exact, one to one correspondence between the clustering 'nodes' and the 'nodes' of the power resolution in space. The taxonomy of power is a function of the level of political organisation pertaining to the specific socio-cultural group being considered. The notion of L. of P., as a functional and cognitive category, involving at least two hierarchical levels (dominating/dominated), is not a naturally inherent condition of human organisation and perception of social space, and constitutes instead a key processual and evolutionary achievement. Once established a given L. of P. is subject to more or less periodical cycles of integration and disintegration. The degree of hierarchy, therefore (within a maximum range, extending from multi-national, 'imperial capitals', to 'national', 'regional', and 'district ' capitals down to local centers), is expected to be highly metastable through time and possibly variable in space as well, even within the same polity. In this regard we should consider from a functional point of view, that a L. of P. (even if we are able to to isolate a perfect synchronic panorama within its evolutionary trajectory, thus avoiding any diachronic palimpsest), could still show some spatial differentiation into modules displaying distinct, 'local' levels of power resolution. These ought constitute different sub-regional or district modules, some of them highly articulated into a number of nested levels of hierarchy, and some others displaying no inner ranking at all.

With such horizons of expectation we should consider our clustering schemes as no more than an empirical toolkit, with a wide 'degree of freedom' left to interpretation, and intended to be tested by other independent sources of information. The clustering properties (hierarchical partitioning) produced by the 'global' functions involve an isomorphic spatial counterpart. In effect the area encompassed by the sites of a given group (that is that one bounded by a closed line connecting all the peripheral sites) will be a spatial sub-set of the area pertaining to every major, 'parent' group agglutinating it at a 'succeeding' step of the overall cluster formation process. There are no (set and spatial) lateral intersections and rearrangements, as is often the case with the 'local' Weighted Density function. This line of reasoning involves a working analogy with the general, expected trend of a modular arrangement of power in space. Such an analogy is, however, quite limited in scope. In fact, our 'modules' are merely conceived as conventional operational units, and no effort has been made to more closely emulate a territorially continuous model of L. of P. by splitting every intermediate piece of land between contiguous groups, thereby attributing a trial 'territorial jurisdiction' to each of them.

Our three 'global' functions therefore offer the possibility to usefully plot, at a certain degree of complexity (number of groups), the various stepwise 'Percolation panoramas' as a composite map (cf. Figs. 5.13–5.14). In this instance the embedded modules are always contained spatially within the parent ones (with a number of possible 'generations' according to the degree of power integration) and, for immediate recognition, they are represented by darker shading. Such an 'overlay model' can also be used for the 'Percolation panoramas' given by 'local' formulas, but only to a limited extend, in as much the hierarchical partitioning rule (or simply 'modular rule') is not broken (cf. Fig. 5.12). In addition the dendrogram model is again an optimum reference scheme for measuring the 'rate of survival' of groups and group leaders along the T axis. This is reckoned by the T range within which a group remains the same or a group leader retains its status within the same group or in the upper order parent groups. Such group and leader paths could represent a statistical measurement of their relative 'stability' throughout the variation of the T parameter. The longer the paths, the more reliable the group and the leader, along with their possible L. of P. counterparts.

Such a 'survival' analogy suggested to us the importance of exploring a proper 'Survival Analysis' (henceforth S.A.) approach (cf. De Guio 1985b, De Guio 1986), in order to carry out a formal general comparison of performances among the four functions. In this regard we should consider the 'life span' of a single site or group of sites as the total, absolute span of the T range, within which it remains 'independent' (that is, not absorbed into another group). At each step within our Percol model (starting from the lowest T value), one or more group are absorbed into a major one and 'die'. S.A. offers a rather sophisticated statistical repertoire (cf. De Guio 1985b, De Guio

1986) for measuring such a 'pattern of death' (that is, the group formation process) as well as comparing different survival performances offered by the four different functions (the function variable thus become the 'stratifying variable').

The simpler Cumulative Survival analysis (Fig. 5.9) presents a cumulative, superimposed plot of the group formation rate (vertical axis= percentage; horizontal= T values, subdivided into 1 km wide entry intervals), in relation to increasing values of T. We perceive a very marked trend common to the four functions: an initial steep slope with a high rate of group formation within a small range of the lower tier of T, up to a critical area (around a T value of 10 km) from which the slope is increasingly flattened with a progressively smaller rate of group formation at comparatively wider intervals. A closer look at such survival profiles of the functions points out to a number of interesting features. The graphs of DE and PP coincide perfectly, with the only small differences, irrelevant to our entry interval span, concerning the 6-groups step (T=13.7 km for DE compared to to 13.3 km for PP) and the 33-groups step (with T=5.2 compared to T=5 respectively). Therefore, for the sake of simplicity, we can assume a single profile as common to both the DE an PP functions, and compare it to the others (WD, DM). (For the same reason we employ the single dendrogram graph for DE, the small differences in comparison to the PP graph being neither significant nor perceptible on our plotting scale: cf. Fig. 5.7).

The comparative 'survival performances' are formally measurable on the basis of a U score (cf. De Guio 1985b, De Guio 1986), which is computed by comparing the survival time of each member of a 'stratum' (one type of function) with those recorded for each member of the other strata (different functions). The results for the four strata are as follows: −17.011 for WD; +1.4667 for DE; +1.4111 for PP; +14.133 for DM. Although such a range of variation is not estimated as statistically significant by the Lee-Desu statistics (cf. De Guio 1985b, De Guio 1986; score=1.012 involving a probability value of O.3145), at least it provides us with a comparative ranking we can use as a rough index of the 'aggregation rate'. In effect, the lower the survival profile, the higher the 'aggregation rate', since an average shorter T range values is required for 'dying', that is for the group formation process. Thus WD has the greatest capacity of aggregation and therefore is the most 'hierarchical' function in this respect. Another indicator giving somewhat different results is the highest T value of the function, which represents the inter-site distance at which all the clusters are absorbed into a single final, all-encompassing group, and the formation process thus terminates: 28 km for WD, 41.3 for both DE and PP, 39.7 for DM. WD still remains the most hierarchical function, but in this instance DM is more hierarchical than DE and PP.

Two other useful S.A. graphs, relative to the the Density and the Hazard functions (Figs. 5.10–5.11), offer a more detailed insight into the group formation profiles. They indicate respectively the overall probability of 'dying' (forming groups) for each T interval (1 km) and the relative risk of 'dying' for each interval in comparison to the immediately preceeding one. Thus, both such indicators denote the most critical (absolute and relative) T values in the overall group formation process. While the Density graph (Fig. 5.10) simply adds minor details to the marked trend already identified by the cumulative survival graph, the Hazard one (Fig. 5.11) seems to present some fresh information. The WD profile is again clearly distinguished from the others, while the DM profile, although much closer to the DE-PP ones, shows localised, interesting deviations from these latter, especially for the highest risk value at the isolated 24–25 km interval. The traits of the hazard curve with a 0 value (no group formation) also permit an immediate perception (and a measurement) of the stability of the overall cluster configuration attained at a given step. On the contrary the critical points (peak values) demonstrate key steps in the aggregation process. Among these the most interesting are those located in the middle-upper T range (from the elbow area displayed by the Cumulative Survival graph onwards), where the higher level and more stable groups (the most likely and reliable in terms of possible L. of P. counterparts) begin to form. In particular they are the 9–10 km and 12–13 km intervals for DW,16–19 km and 24–25 km for DM, 13–14 km, 17–19 km, 32–33 km for DE-PP, along with the Top T values for the final group formation stressed above.

At this juncture, it is important to comment briefly upon our sets of 'Percolation panoramas', keeping in mind that our main gaol is simply to offer a number of set-configurations which scholars can use as a frame of reference against which to compare their (textual and field) data. We limit our comparative analysis to the upper levels of group hierarchy, in the hope of finding L. of P. counterparts in terms om major modules (such as 'region', 'sub-region', 'district', and so on) of an expected three to five rank hierarchy which we should reasonably expect to encounter within a study area as this one (cf. Adams 1981, De Guio 1985a). The minor modules are likely to reflect some kind of interaction sphere (e.g. social, economic, demographic) not necessarily meaningful in terms of a proper political setting. They could represent perhaps expected lines of preferential fracture along which an assumed cumulative process of power integration process is likely to functionally dissect through time.

As far as the first sets of WD 'Percolation panoramas', we can only represent with our overlay model the 3-group and 7-group steps (Fig. 5.12). The algorithm does not, in fact, aggregate the 2, 4, 5 group configurations, since, at the relevant T values, more than one group is added to the old ones. In addition the first new group configuration following the 7-group set, begins to break the hierarchical partitioning rule by producing lateral rearrangements among the clusters. Basically we could recognize a rather 'stable' (28–12km T range) 3 group pattern (Northern, Central, Southern). The group leaders (we individuate every group from its leader; in case of matching rank we mention all the co-leaders) are n. 432 (30 ha), n. 428 (2 ha), n. 430 (2 ha), n. 431 (7 ha), n. 429 (2 ha), all with a score of 429 for the Southern group, n. 175 (30 ha; 276 score) for the Northern one, and n. 131 (30 ha; 225 score) for the Central one. For the 7-group partition we merely stress how it subdivides into five sub- units the Northern cluster: 1) n. 172 (2 ha), Zabalam (100 ha), Umma (100 ha) each with a score of 202; 2) n. 49 (7 ha) with 160; 3) n. 11 (2ha) with 137; 4) n. 190 (15 ha) with 63; 5)n. 235 (2 ha) and n. 213 (30 ha) both with 32. It thus becomes apparent throughout the stepwise procedure that there is no correspondence between the score ranking and the size ranking, both within the entire set of sites as well as within each group. Although we have already classified the WD function as the most hierarchical, since it aggregates the total system within a shorter T range, it appears not to 'recognize', via the scoring procedure, any L. of P. ranking. This should display at least a critical amount of congruence with the 'size' sets of values (henceforth we shall address such an optimum property as 'ranking capability'). The WD score is basically a function of location, and the group leaders represent nothing more group 'barycenters'.

In summary we must attribute to the WD approach the limited capability, itself quite close to the original percolation model, of recognizing 'modal groups' of weighted entities. In terms of L. of P. analysis such a property should produce better results to the extent that the case-study pattern approaches some ideal spatial configurations, such as those postulated by the Central Place theory, which we should consider rather exceptional in our real world experience. Again, the lack of 'ranking capability' and the frequent exceptions to the 'modular rule', cause us to rely more upon the other set of 'global' functions.

In addition, these last functions have a great deal in common. Beyond an absolute concordance with the 'modular rule', they show an average high 'ranking capability'. The best results are given by DE with a perfect matching between size and score ranks, followed by PD and DM, with a limited number of no-matching values restricted to the lower tier of size range (1 to 7 ha).

Moving now to a quick comparison of the cluster configurations produced by the three functions, we must first recognize the highest degree of similarity between the two demographic functions. They are perfectly identical as far as the content of all the groups at each step of the formation process is concerned, and differ to a very limited extent only in the score ranking within some groups, and, by a minimum fractional value, in the T values at which the 6 group and 34 group configurations are produced (cf. above). This permits us to argue in terms of DE-PP phenomena and to rely upon just one 'overlay model' plot and one dendrogram plot. Such an impressive similarity suggests a rather interesting consideration: the weighting (and ranking!) factor Wi, which distinguishes the two formulas from one another, offers DE a better 'ranking capability'; however, the 'simpler' Potential of Population function recognizes the same cluster configuration as that produced by DE. This reasserts our expectation concerning a reasonable amount of homomorphism between L. of P. and a 'demographic landscape'. If we now compare the DM ('Dominance', that is, an ad hoc, 'power recognition' oriented function) versus DE-PP clustering configurations (cf. Figs. 5.5–5.8, 5.13, 5.14), we again discover a substantial amount of similarity (compared to the very poor concordance with the previous WD function). This appears to reassure us regarding the latter assumption. Some very interesting divergences, however, are visible (in order to hasten our comparison procedure we resort to a tripartite formula so as to individuate a given group: formula abbreviation/ group leader identifier (number or name)/ group formation step number; the = symbol signifies the identity of content between groups; the ≈ symbol points out a considerable similarity, that is a wide intersection between the two sets). Both approaches recognize an initial (starting from the cluster root) 2 group configuration (DM/ Larsa/ 2 ≈ DE-PP/ Larsa / 2; DM/ 4/ 2 ≈ DE-PP/ UMMA/ 2). The difference in content is quite insignificant, involving just two very small and peripheral sites (n.70/ 2 ha, and n. 78/ 1 ha). Such a pattern recognition appears rather odd and somewhat surprising, but it is nonetheless a very suggestive one. The bi-polar 'subregional' configuration perfectly adheres to two distinct channel/irrigation systems as shown by Adams (Adams 1981, Fig. 31). It is easy to argue that such an impressive, energy spending infra-structure had to involve and/or imply a properly politically-founded management system (as something akin to 'power infrastructures').

The major stability suggested by the widest 'life spans' of such groups can be further evaluated as a rough indicator of their higher reliability in terms of possible modular L. of P. counterparts. It is worth noting here the existence a very interesting feature: the highest 'survival' value is not that of the overall leader Larsa (200 ha), but that of Umma

(100 ha). In fact in the following Old Babylonian Period (Fig. 5.4) Umma becomes the leader, while the Larsa module is reduced and disintegrated, and Larsa itself (the exact opposite of Umma!) halves its size. The new and more nucleated settlement system of the Old Babylonian Period, which presents a marked shift in the focus of dominance to the North (and gives rise to a new bi-polarity between the Northern and Southern sub-regions) seems also to solve some previous anomalies, for instance, the close proximity of the two equal-size higher order (100 ha) sites of Zalabam and Umma itself. A similar phenomenon seems to take place between the Akkadian (more nucleated) and the Uruk–Isin–Larsa Periods (Figs. 5.2–5.3), with the disappearance of the 100 ha site n. 168 located even closer to the equal-size Zabalam, within a context of a more diffused settlement pattern. We could evaluate such processes in the once fashionable terms of cybernetics (cf. Clarke 1978, p. 88–91), arguing for 'contradictive variety' entering an information system and causing 'equivocation'. But we should also consider the L. of P. morphogenetic path within a tighter processual perspective. In particular we must take into account the time-span of each Period, which surely involved a great deal of evolutionary and spatially discriminating trends, artificially compressed into an unlikely, operational time-slice.

A. De Guio

5.6 Some cautionary reservations

It is necessary to point at the major areas of reservations as regards our Percol model. The majority of these are common to an entire class of L. of P. approaches, and are therefore of a more general interest (cf. De Guio 1985a).

As to the the reliability of Mesopotamian survey work in general, we refer the reader to Oates' (Oates 1977) thorough review of this material. Still, one of the most constructive sources of criticism, especially concerning his own data, is Adams himself (Adams 1981). On our part, in order to manipulate this data-base to operationalize the Percol approach, we were forced to implement further simplifications and approximations. Our reservations largely concern:

1. the size measurements which do not represent precise estimates, but instead the central values of arbitrary broad size classes (based on the minimum rectangular area encompassing the sites) as offered by Adams who stresses their non-optimum level of accuracy (Adams 1981, pp. 131, 170);
2. the low chronological resolution of data. The chronological grid is so loose that each presumable L. of P. does not constitute a reliable synchronic map of power. In reality it represents a thick palimpsest (and probably an 'overestimated map of power'), encompassing an unknown quantity of evolutionary processes (perhaps highly discontinuous or cyclical, and sometimes 'catastrophic', as seen in the frequent 'pulsatory' alternation of nucleated and diffused settlement systems (cf. Renfrew & Poston 1979)). This is likely to be relevant to our Ur III–Isin–Larsa Period, which spans across some 300 years, and involves three successive and distinct political settings, each with its own capital (Adams 1981, p. 143).;
3. the rather artificial limits of the area under examination and the 'boundary problems' involved with these.

Figure 5.1: From Early Dynastic II–III to Middle Babylonian Period: composite map (cf. Adams 1981, enclosed map).

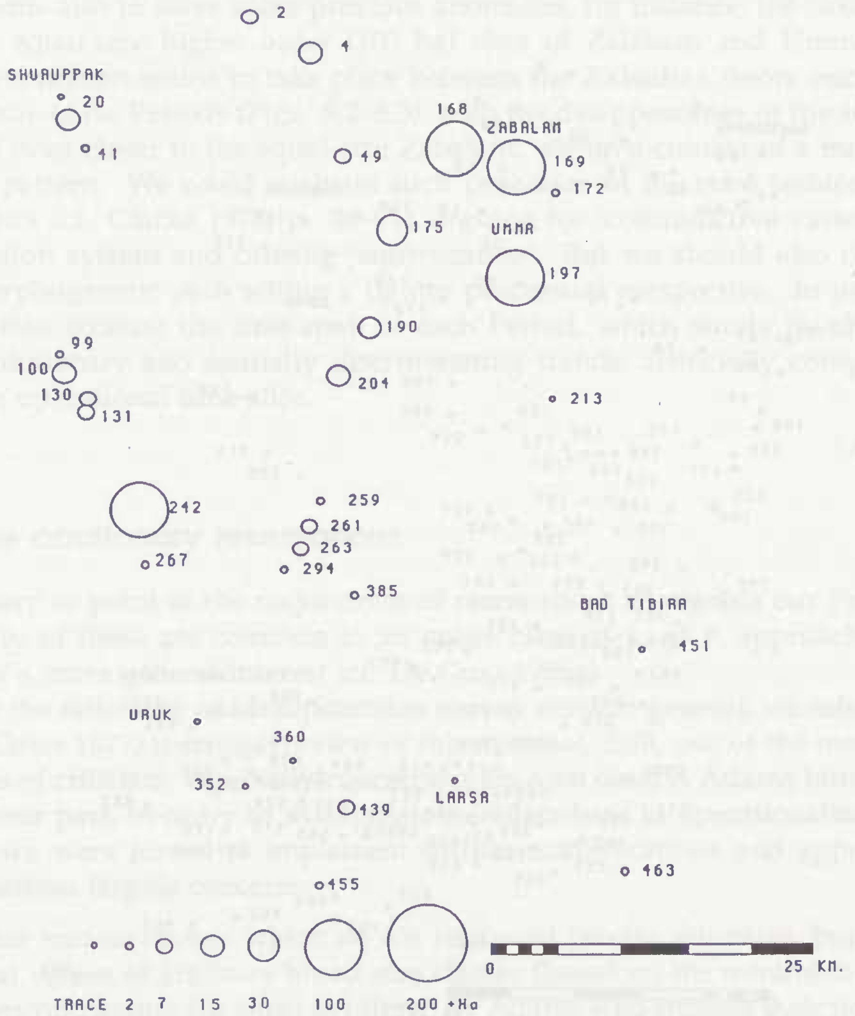

Figure 5.2: Akkadian Period: distributional map (cf. Adams 1981, Tab. 14, Fig. 30).

Figure 5.3: Ur III–Isin–Larsa Period: distributional map (cf. Adams 1981, Tab. 14, Fig. 31).

Figure 5.4: Old Babylonian Period: distributional map (cf. Adams 1981, Tab. 14, Fig. 33).

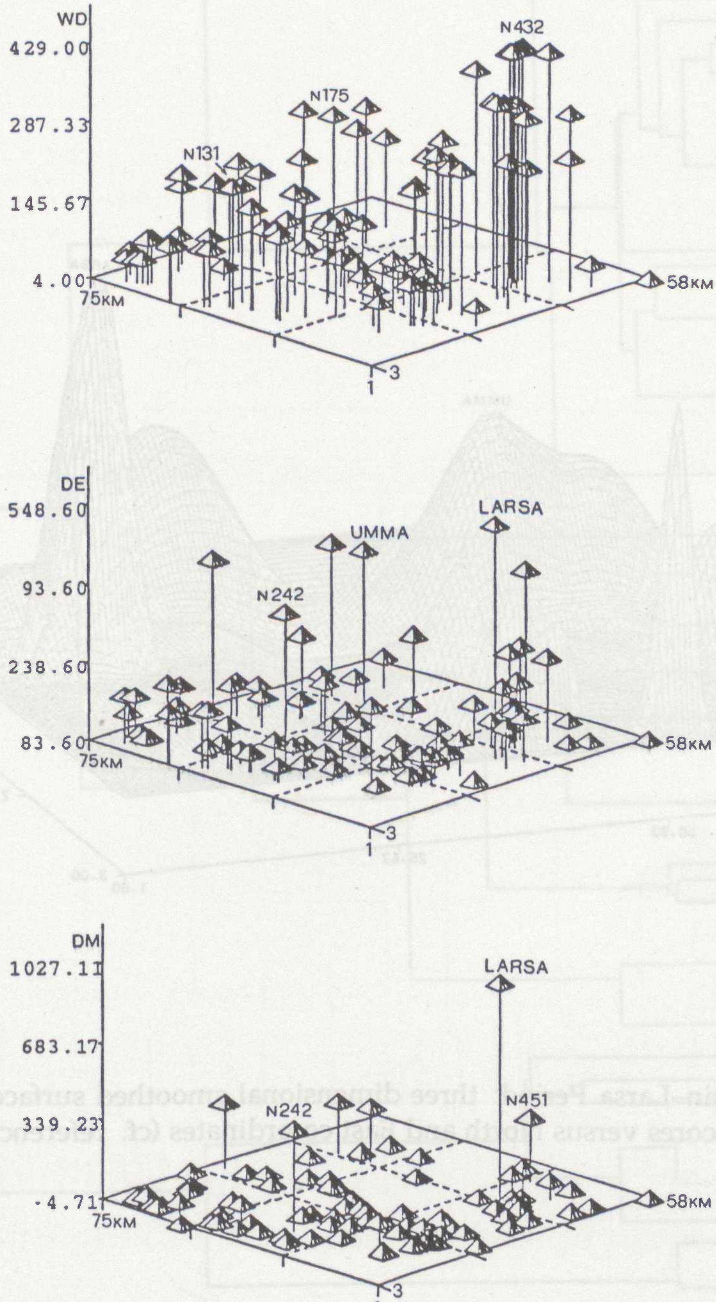

Figure 5.5: Ur III–Isin–Larsa Period: three dimensional scatter plots of Weighted Density (WD: the values are those of step 3), Demographic Energy (DE) and Dominance (DM) scores versus North and East co-ordinates (cf. reference map in Fig. 5.3).

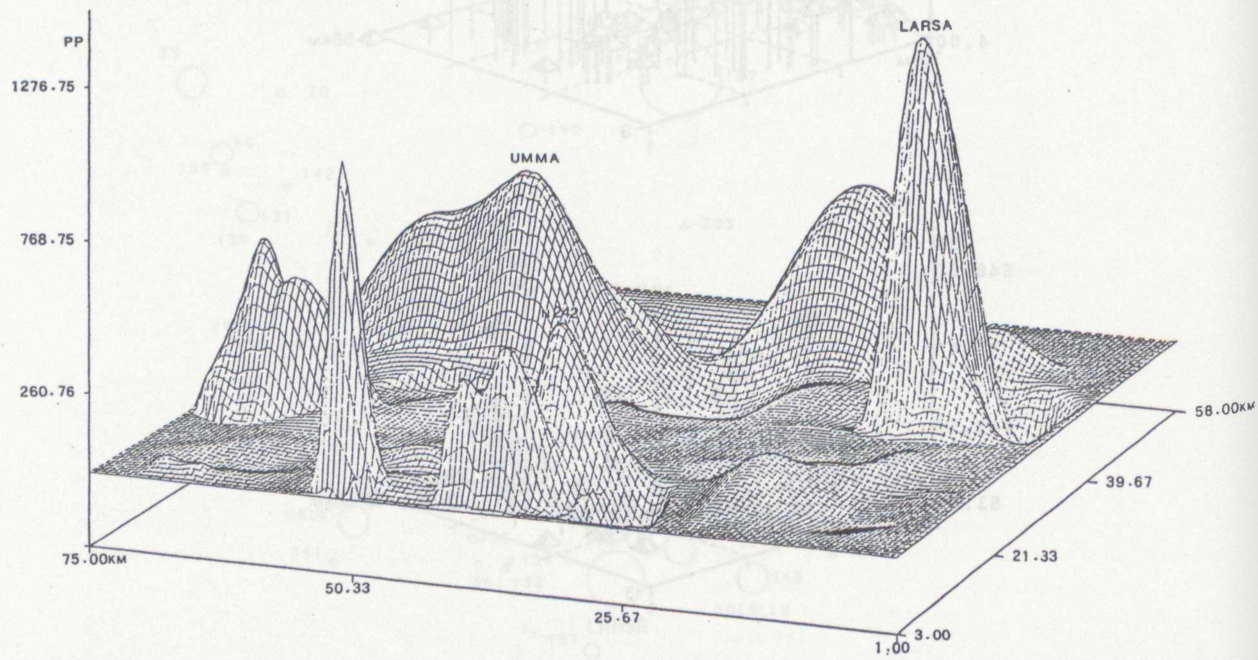

Figure 5.6: Ur III–Isin–Larsa Period: three dimensional smoothed surface of Potential of Population (PP) scores versus North and East co-ordinates (cf. reference map in Fig. 5.3).

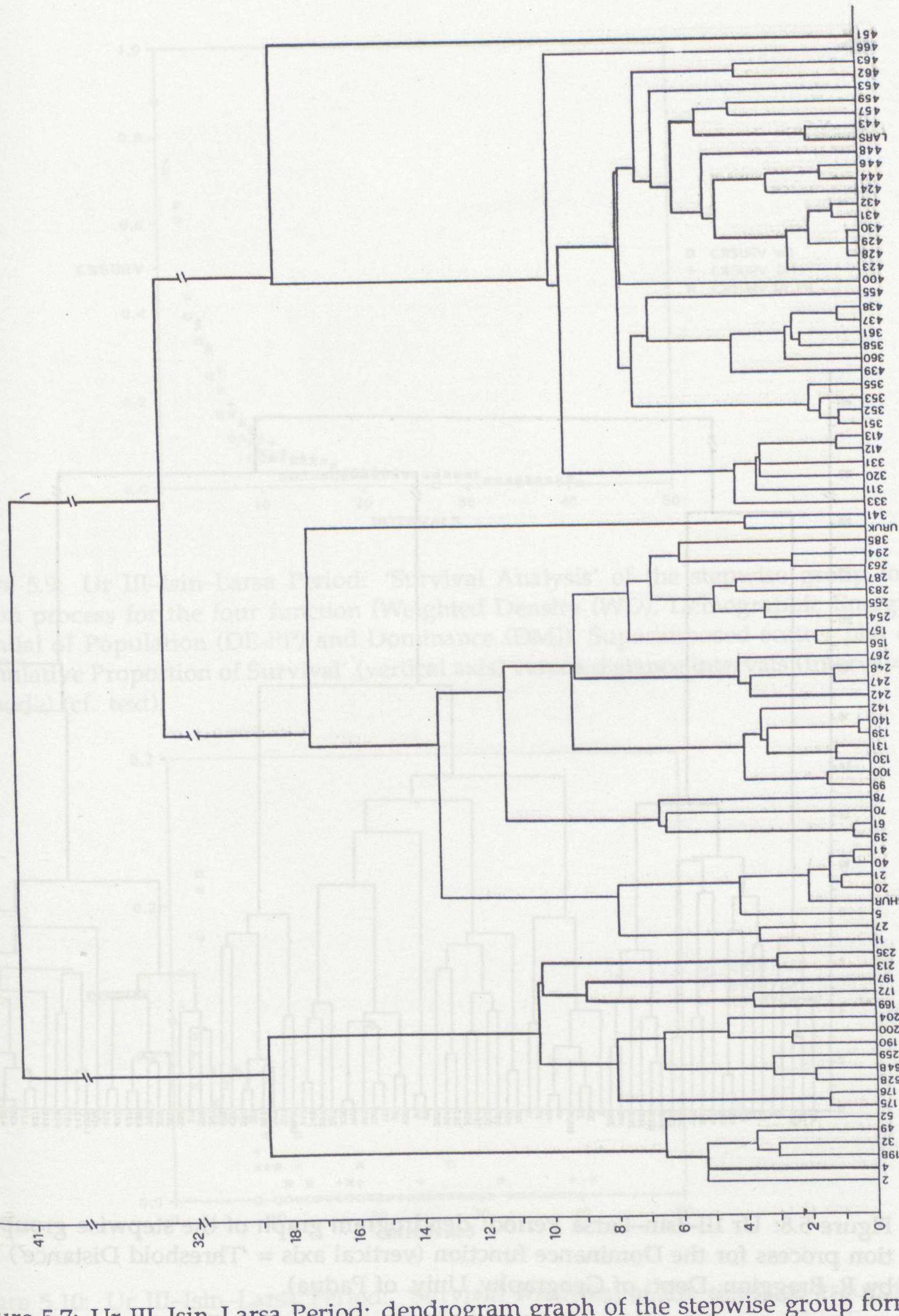

Figure 5.7: Ur III–Isin–Larsa Period: dendrogram graph of the stepwise group formation process for the Demographic Energy function (vertical axis = 'Threshold Distance') (drawn by R. Braggion, Dept. of Geography, Univ. of Padua).

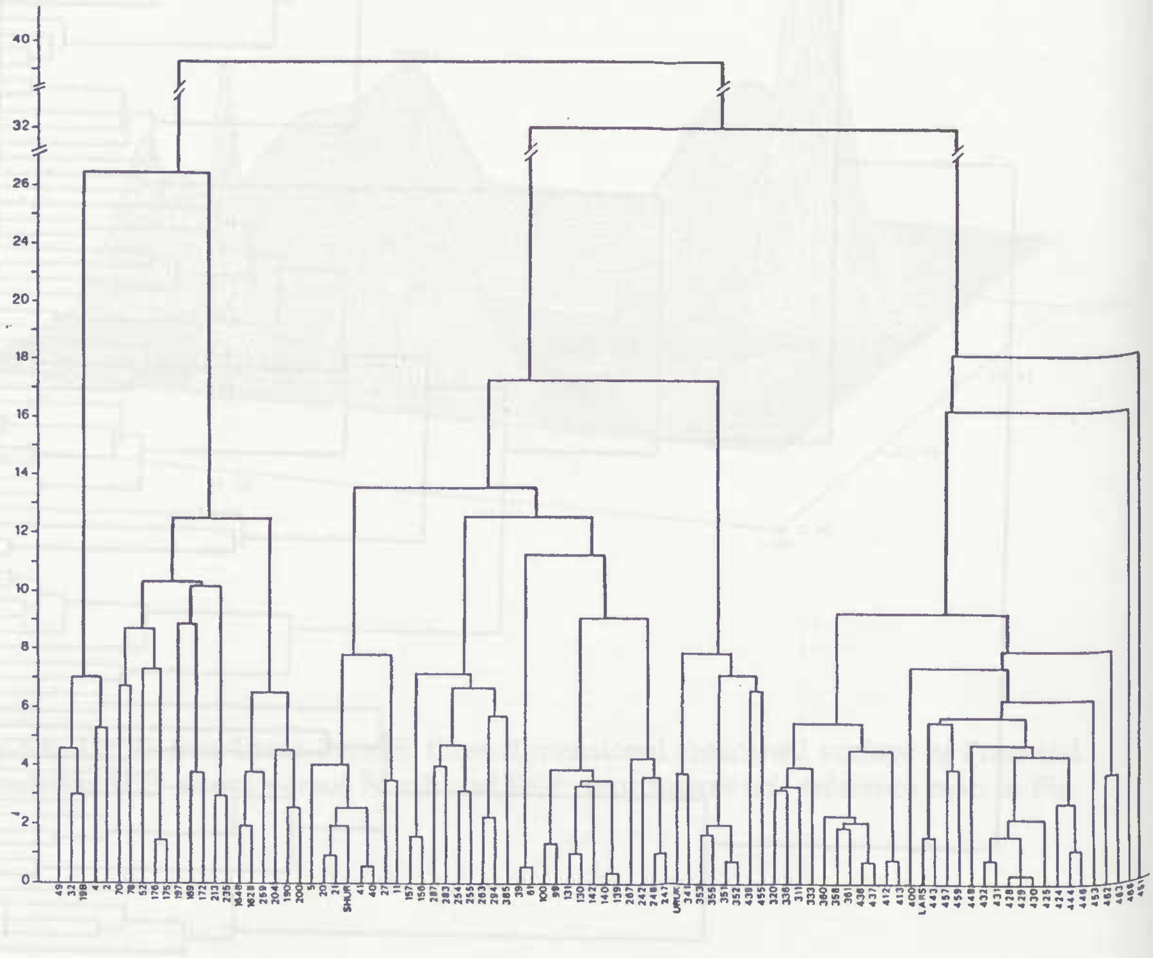

Figure 5.8: Ur III–Isin–Larsa Period: dendrogram graph of the stepwise group formation process for the Dominance function (vertical axis = 'Threshold Distance') (drawn by R. Braggion, Dept. of Geography, Univ. of Padua).

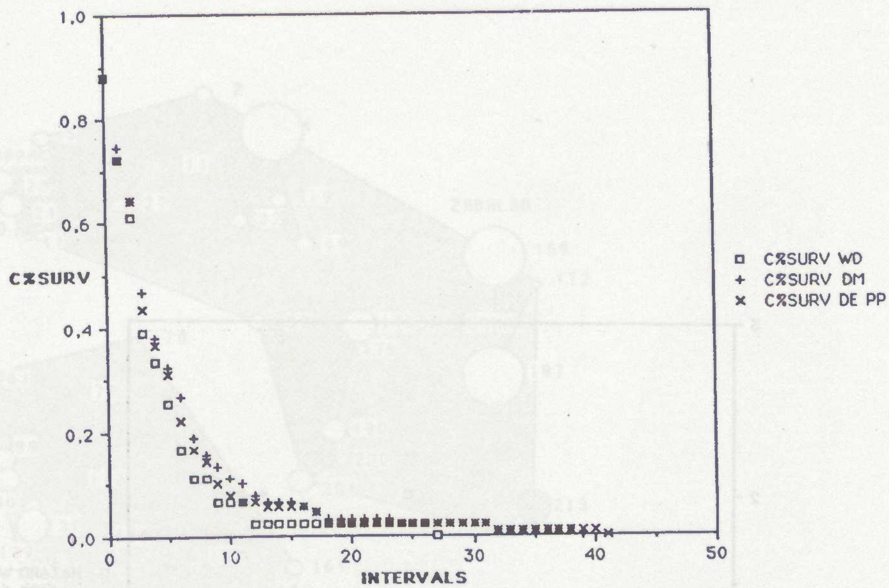

Figure 5.9: Ur III–Isin–Larsa Period: 'Survival Analysis' of the stepwise group formation process for the four function (Weighted Density (WD), Demographic Energy-Potential of Population (DE-PP) and Dominance (DM)). Superimposed scatter plot of 'Cumulative Proportion of Survival' (vertical axis) versus distance intervals (Interval=1 km wide) (cf. text).

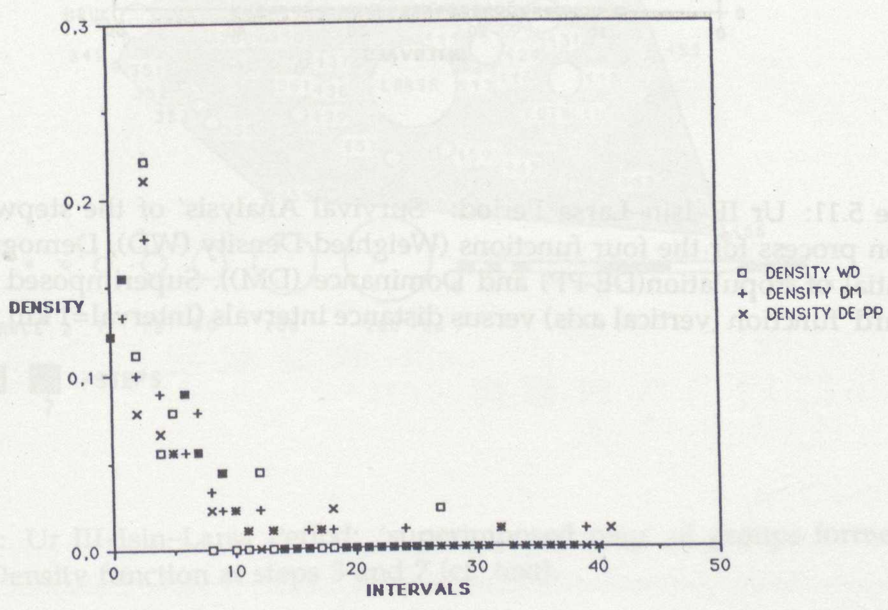

Figure 5.10: Ur III–Isin–Larsa Period: 'Survival Analysis' of the stepwise group formation process for the four functions (Weighted Density (WD), Demographic Energy-Potential of Population (DE-PP) and Dominance (DM)). Superimposed scatter plot of 'Density' function (vertical axis) versus distance intervals (Interval=1 km wide) (cf. text).

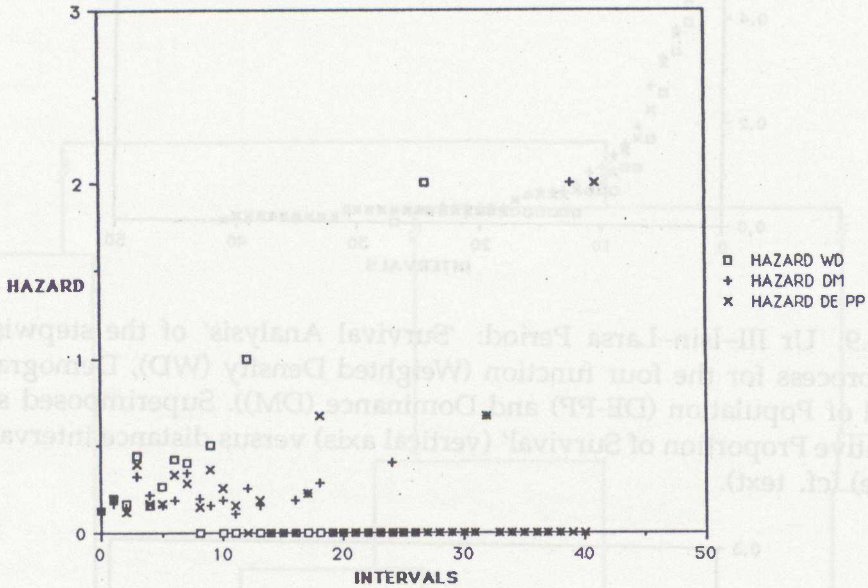

Figure 5.11: Ur III–Isin–Larsa Period: 'Survival Analysis' of the stepwise group formation process for the four functions (Weighted Density (WD), Demographic Energy-Potential of Population(DE-PP) and Dominance (DM)). Superimposed scatter plot of 'Hazard' function (vertical axis) versus distance intervals (Interval=1 km wide)(cf. text).

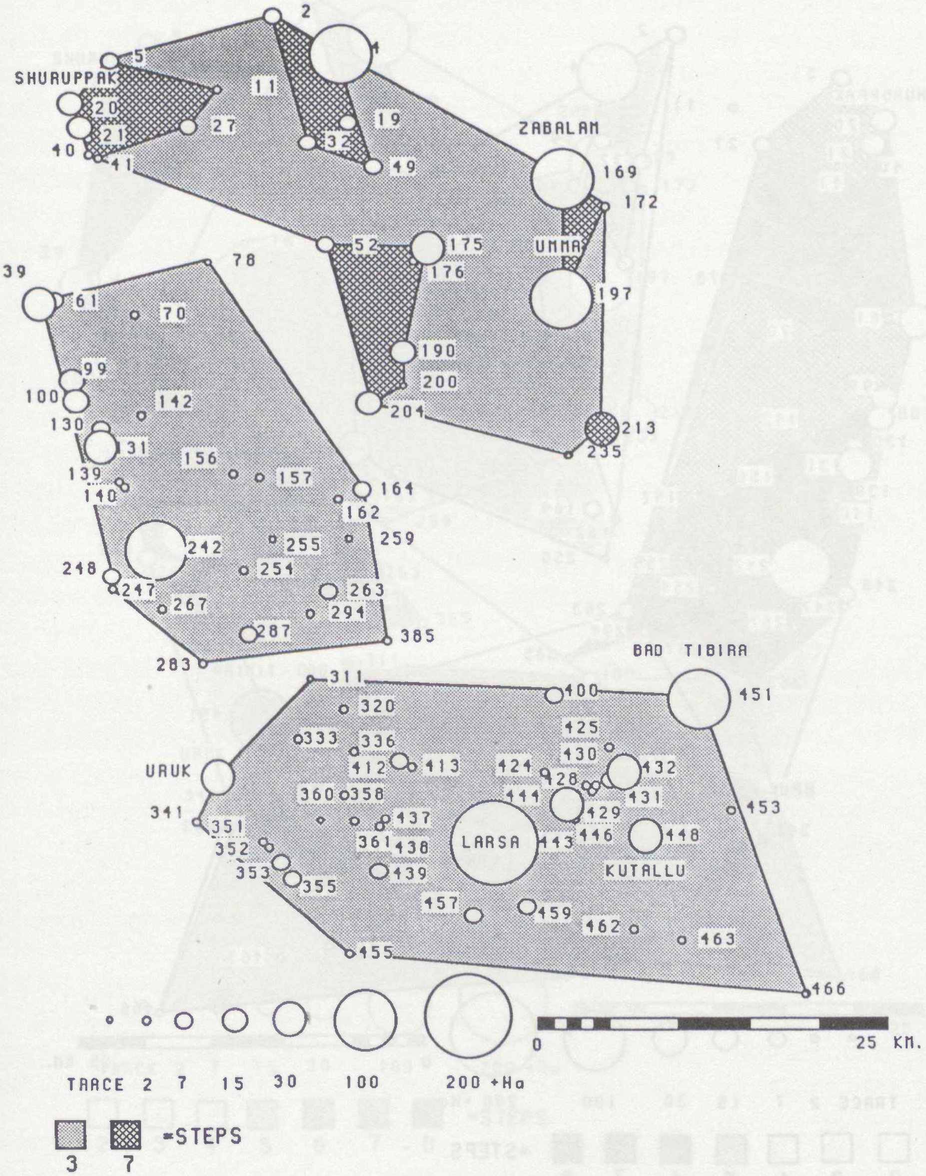

Figure 5.12: Ur III–Isin–Larsa Period: 'superimposed map' of groups formed by the Weighted Density function at steps 3 and 7 (cf. text).

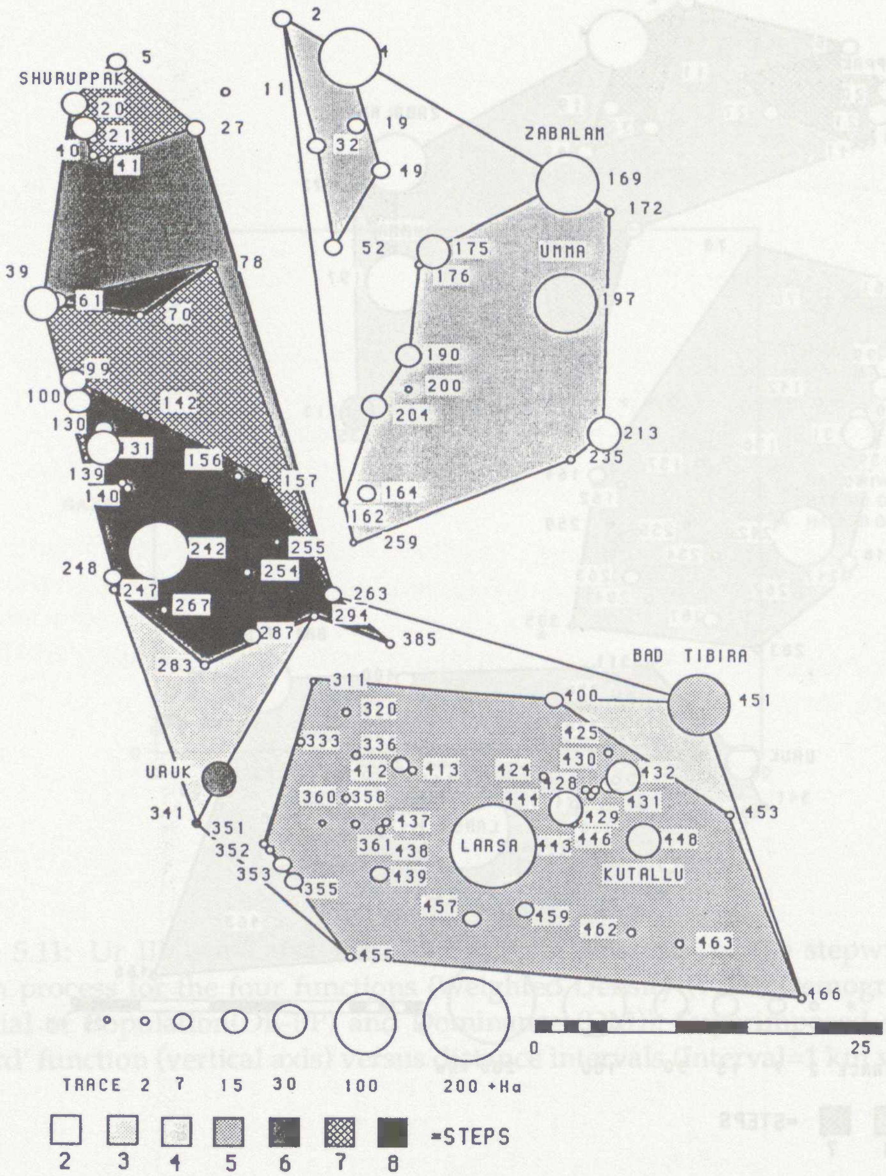

Figure 5.13: Ur III—Isin–Larsa Period: 'superimposed map' of groups formed by the Demographic Energy- Potential of Population functions at steps 2 to 8 (cf. text).

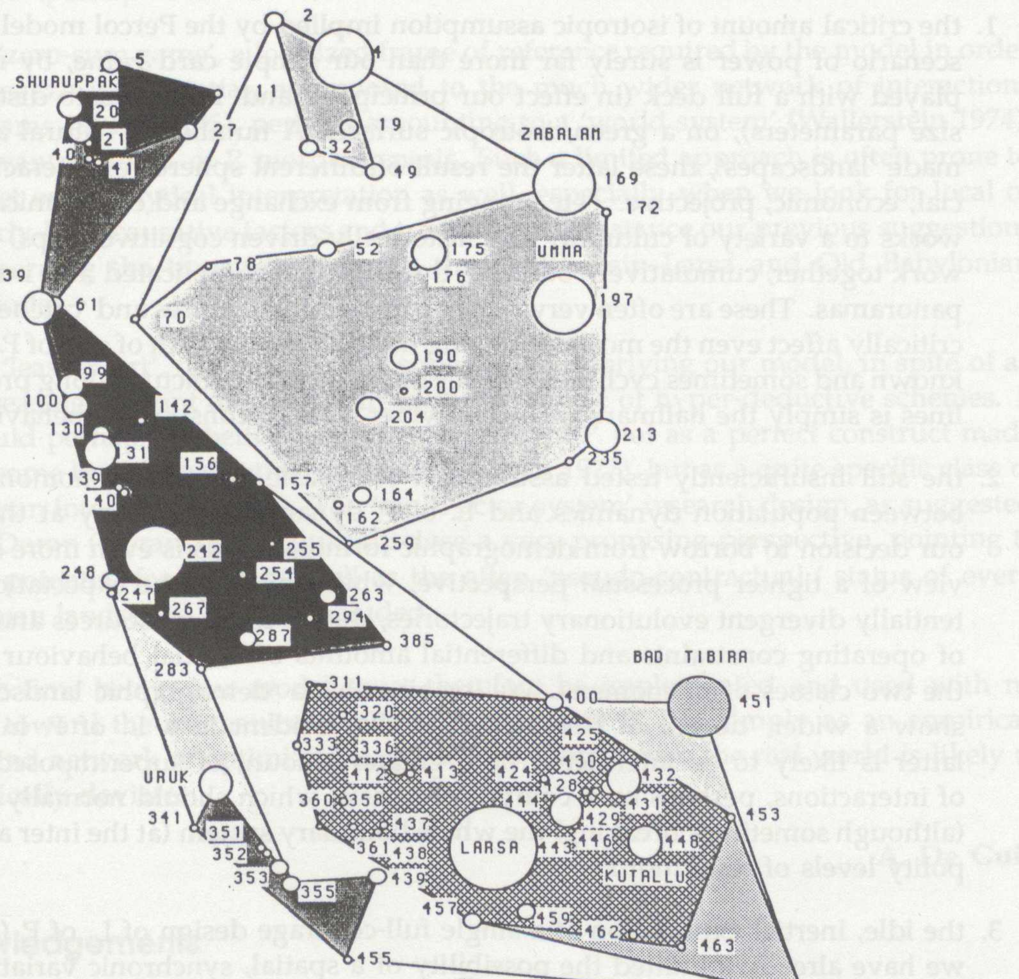

Figure 5.14: Ur III–Isin–Larsa Period: 'superimposed map' of groups formed by the Dominance functions at steps 2 to 8 (cf. text).

Other relevant areas of weakness pertain to our theoretical approach towards the L. of P. area:

1. the critical amount of isotropic assumption implied by the Percol model. A likely scenario of power is surely far more than our simple card game, by no means played with a full deck (in effect our principal 'hand' is limited to distance and size parameters), on a green, isotropic surface. A number of natural and man-made 'landscapes', these latter the result of different spheres of interactions (social, economic, projective... etc., ranging from exchange and communication networks to a variety of culturally and historically driven cognitive maps) normally work together, cumulatively building through time complicated and intersecting panoramas. These are often very rich in constraining features and 'catches', which critically affect even the most optimizing morphogenetic path of a L. of P. The well known and sometimes cyclical tendency of a territory to fracture along preferential lines is simply the hallmark of such a wider class of constrained behaviour;

2. the still insufficiently tested assumption of a basic underlying homomorphism between population dynamics and L. of P. dynamics, which lay at the root of our decision to borrow from demographic formulas. This is even more critical in view of a tighter processual perspective, which involves the expectation of potentially divergent evolutionary trajectories, due to different sources and degrees of operating constraints and differential amounts of inertial behaviour between the two classes of phenomena. We imagine that a 'demographic landscape' will show a wider 'degree of freedom' than that evident in a L. of P. In fact the latter is likely to be intersected by a greater amount of superimposed spheres of interactions, perhaps on a compulsory basis, which should normally reinforce (although sometimes breaking) the whole boundary system (at the inter and intra-polity levels of resolution);

3. the idle, inertial expectation of a single full-coverage design of L. of P. (although we have already admitted the possibility of a spatial, synchronic variation in its modular resolution), against the suggestions of more complicated and intersecting designs, such as those supposed for the Middle Babylonian Period by Adams himself (Adams 1981, pp. 150–153): a basic social and political dichotomy of the L. of P. between a city- state system and a rural system, the latter principally tribal and kin-based in nature;

4. the difficulty in locating a reliable amount of external sources of information in order to test any given L. of P. simulation. The 'iconography of power' whatever form it may assume, from boundary stones to spatially opposed active symbols (cf. Hodder 1982), is very limited in relevant content, even for this case-study area. The considerable amount of textual sources, which initially seemed so promising, was not of particular assistance. Even our expectations concerning the recognition of a cognitive taxonomy of the local, ancient L. of P. (see, for instance, the arguments regarding terms such as 'uru-sag', 'uru', alu', 'kaprum', 'e-duru', 'maskanu': Edzard 1964, Hallo 1971, Leemans 1975), seem to authorize little more than a rough dichotomy between a non-urban settlement category (in a trial range between 'trace' and 4 ha sites, probably encompassing a wide

functional spectrum of something akin to villages, manors, processing facilities, temporary camps and so on), as opposed to a equally generic and inclusive city status (perhaps over 10 ha);

5. the 'zero-sum game', a localized frame of reference required by the model in order to realize its potential, as opposed to the much wider network of interactions (Adams 1981, p. 135), perhaps amounting to a 'world system' (Wallerstein 1974), relevant to our L. of P. morphogenesis. Such a limited approach is often prone to affect our analytical interpretation as well, especially when we look for local or nearly-local causative factors and trends (cf. for instance our previous suggestions concerning the transition between the Ur III—Isin–Larsa and Old Babylonian Periods);
6. the 'least effort', optimizing philosophy still underlying our model, in spite of an increasing amount of skepticism towards the use of hyper-deductive schemes. It would perhaps be better to approach the L. of P. not as a perfect construct made by some kind of 'analytical engine' (cf. Klejin 1977), but as a quite specific class of human interaction sphere. The 'multi-actor system' research design, as suggested by Doran (Doran 1982), could disclose a very promising perspective, pointing to the potential for error, as well as the often 'pseudo-contractual ' status of every human landscape, L. of P. included.

For the time being, our model must therefore be implemented and used with no pretense toward the mechanical replication of a L. of P., but simply as an empirical, background network of optimizing expectations, from which the real world is likely to systematically deviate.

A. De Guio

Acknowledgements

We are very grateful to Dr. C. N. Sandberg (Dept. of Anthopology, Harvard University) for his very impressive effort in improving clarity and content of the text and to Mr. S. Rahtz for the final editing.

References

ADAMS, R. M. 1981. *Heatland of Cities. Surveys of ancient settlement and land use of central foodplain of the Euphrates.* The University of Chicago Press, Chicago.

ALDEN, J. R. 1979. "A reconstruction of Toltec period political units in the Valley of Mexico". *in* Renfrew, C. & Cooke, K. L., (eds.), *Transformations. Mathematical approaches to culture change,* pp. 169–200. Academic Press, New York.

AMMERMAN, A. J. 1981. "Survey and archaeological research", *Annual Review of Archaeology,* 10: 63–88.

CLARKE, D. L. 1978. *Analytical archaeology.* Methuen & Co, London.

DE GUIO, A. 1985a. "Analytical tools for simulating morphogenetic processes of 'landscapes of power"'. *in Data Management and Mathematical Methods in Archaeology—Denver, 29th April–1st May 1985.* U.I.S.P.P.

DE GUIO, A. 1985b. "Archaeological applications of survival analysis,". *in* Voorips, A. & Loving, S. H., (eds.), *To pattern the past,* pp. 361–381. Council of Europe, Strasbourg.

DE GUIO, A. 1986. "'Analisi della Sopravvivenza': dalle scienze biomediche all'archeologia", *Aquileia Nostra,* LVII: 101–128.

DE GUIO, A., S. P. EVANS, & A. RUTA SERAFINI 1986. "Marginalità territoriale ed evoluzione del ≪paesaggio≫ del potere: un caso di studio nel Veneto", *Quaderni di Archeologia del Veneto,* II: 160–172.

DORAN, J. E. 1982. "A computational model of sociocultural systems and their dynamics" *in* Renfrew, C., Rowlands, M. J., & Segraves, B. A., (eds.), *Theory and explanation in archaeology,,* pp. 375–388. Academic Press, New York.

EDZARD, D. O. 1964. "Mari und Aramaeer ?", *Zeitschrift für Assyriologie,* 56: 142–149.

FRIEDMAN, J. & M. J. ROWLANDS, (eds.) 1982. *The evolution of social systems.* Duckworth, London.

HALLO, W. W. 1971. "Antediluvian cities", *Journal of Cuneiform Studies,* 23: 57–67.

HODDER, I. 1979. "Simulating the growth of hierarchies". *in* Renfrew, C. & Cooke, K. L., (eds.), *Transformations. Mathematical approaches to culture change,* pp. 117–144. Academic Press, New York.

HODDER, I. 1982. *The Present Past.* B. T. Batsford, London.

HODDER, I. & C. ORTON 1976. *Spatial analyis in archaeology.* Cambridge University Press, Cambridge.

JOHNSON, G. A. 1981. "Monitoring complex system integration and boundary phenomena with settlement size data". *in* van der Leeuw, S. E., (ed.), *Archaeological approaches to the study of complexity,* pp. 144–188. Universiteit van Amsterdam, Amsterdam.

KLEJIN, L. S. 1977. "A panorama of theoretical archaeology", *Current Anthropology,* XVIII: 1–43.

LEEMANS, W. F. 1975. "The role of landlease in Mesopotamia in the early second millennium B.C", *Journal of the Economic and Social History of the Orient,* 18: 134–145.

OATES, J. 1977. "Archaeology and geography in Mesopotamia". *in* Bintliff, J., (ed.), *Mycenean geography,* pp. 101–106. British Association for Mycenean Studies, Cambridge.

REILLY, W. J. 1929. "Methods for the study of retail relationships", *University of Texas Bulletin,* 2944.

RENFREW, C. 1982. "Space, time and polity". *in* Friedman, J. & Rowlands, M. J., (eds.), *The evolution of social systems*, pp. 89–112. Duckworth, London.

RENFREW, C. 1984. *Approaches to social archaeology*. Edinburgh University Press, Edinburgh.

RENFREW, C. & J. F. CHERRY, (eds.) 1986. *Peer polity interaction and socio–political change*. Cambridge University Press, Cambridge.

RENFREW, C. & E. V. LEVEL 1979. "Exploring dominance: predicting polities from centres". *in* Renfrew, C. & Cooke, K. L., (eds.), *Transformations. Mathematical approaches to culture change*, pp. 145–167. Academic Press, New York.

RENFREW, C. & T. POSTON 1979. "Discontinuities in the endogenous change of settlement pattern". *in* Renfrew, C. & Cooke, K. L., (eds.), *Transformations. Mathematical approaches to culture change*, pp. 437–461. Academic Press, New York.

STEPONAITIS, V. 1981. "Settlement hierarchies and political complexity in non market societies: the formative period in the Valley of Mexico", *American Anthropologist*, 83: 320–363.

STEWARD, J. Q. & W. WARNTZ 1958. "Macrogeography and social science", *The Geographical Review*, 48: 167–184.

TREMOLLIERS, R. 1979. "The Percolation Method for an efficient grouping of data", *Pattern Recognition*, 11: 225–269.

TREMOLLIERS, R. 1981. *Introduction aux fonctions de densité d'inertie*. Etudes et Documents 234. Institute d'Administration des Enterprises, Aix en Provence.

TREMOLLIERS, R. 1982. *Qualitative data clustering by the Percolation Method*. Etudes et Documents 248. Institute d'Administration des Enterprises, Aix en Provence.

TREMOLLIERS, R. 1984. *La nouvelle Percolation Généralisée pour l'analyse des données et la reconnaissance des formes*. Etudes et Documents 306. Institute d'Administration des Enterprises, Aix en Provence.

VOORIPS, A. 1981. "To taylor the inflected tail: reflections on rank–size relationships". *in* van der Leeuw, S. E., (ed.), *Archaeological approaches to the study of complexity*, pp. 189–196. Universiteit van Amsterdam, Amsterdam.

WALLERSTEIN, I. 1974. *The modern world system: capitalist agriculture and the origins of the European world-economy in the sixteen century*. Academic Press, New York.

WARNTZ, W. 1964. "A new map of the surface of population potentials for the United States, 1960", *The Geographical Review*, 54: 170–184.

ZANETTO, G. 1979. "Il potenziale: da modello a strumento", *Rivista Geografica Italiana*, 86: 298–320.

ZIPF, G. K. 1946. "The P_1P_2/D hypothesis: on the intercity movement of persons", *American Sociological Review*, 11: 667–686.