Functional Regions Reexamined: The Black/White Interaction Model[1]

Valerian T Noronha

Department of Geography, University of Western Ontario
London Ontario, Canada N6A 5C2

1. Introduction

The regional concept has traditionally been one of the cornerstones of geographic thought. The notion of "chorology" was introduced by Ptolemy in the 2nd century AD (Stevenson, 1932); it was developed and elaborated in the 18th and 19th centuries by German scholars, notably Ritter and Passarge (Dickinson, 1969). The evolution of the regional concept accelerated around the turn of this century, as the scope for its application was recognized and exploited. In the United States, the Administration of Woodrow Wilson called upon Isaiah Bowman, then Director of the American Geographical Society, to assist in territorial negotiations at the Peace Conference in Paris (Greenough, 1919). In Europe, several countries were engaged in internal political reorganization; again, geographers were consulted (Gilbert, 1951). As a result of this renewed attention, there were attempts within the discipline to consolidate the concepts surrounding the region (for example, Unstead *et al*, 1937).

Meanwhile, in the middle 20th century, social sciences in general and classification techniques in particular were experiencing a resurrection in the wake of the digital computer. Multivariate statistical tools were increasingly accessible. Classification, once stigmatized by association with subjective description, acquired a new respectability. In the early 1960s, geographers began to recognize the bridges between classification and regionalization. The works of Bunge (1962), Berry (1964) and Grigg (1965) laid the groundwork for the literature on objective regionalization that emerged in the 1970s.

This paper is intended to contribute to the field of objective regionalization, by introducing a model for the conceptualization and delimitation of functional regions. The terminology and classification of regions are examined briefly in the following section, and the literature is reviewed in Section 3. The Black/White Model is then presented, and its implications for regionalization examined in the remainder of the paper.

2. Taxonomies of Regions

Traditionally, regions have been described and differentiated with respect to broad classes of variables. One may thus encounter references to **physiographic** or **climatic** regions, in terms of the physical environment, or human concepts such as **umland** and **rural-urban fringe**. While these notions led to a **descriptive** taxonomy of regions, an alternative classification emerged in the middle 20th century: the **structural** taxonomy differentiated between **formal**, **functional**, **nodal** and **equitable** regions.

---

[1] Based on a PhD thesis written under the supervision of Professors Michael Goodchild, William Warntz and Donald Janelle, and accepted by the University of Western Ontario in August 1985. This version (transcribed 1998 from the original 1985 manuscript) won the Nystrom PhD dissertation award at the meetings of the Association of American Geographers, Minneapolis, 1986. A more comprehensive paper was published by Noronha and Goodchild (1991) "Modeling Interregional Interaction: Implications for Defining Functional Regions" *Annals, Association of American Geographers* **82** (1) 86-102. Contact the author at *noronha@dgrc.ca*

Since regionalization is a classification of space, one first defines **elementary spatial units** or ESUs, within which relevant characteristics are assumed to be constant; these are aggregated into regions. The distinction between formal and other regional types in the structural taxonomy lies in the criteria for aggregating ESUs. The formal region has been defined as "the largest area over which a generalization remains valid" (Minshull, 1967, p 38), directly addresssing the concept of internal homogeneity traditionally associated with the regions covered under the descriptive taxonomy. In a quantitative context, formal regionalization may be achieved by clustering ESUs so as to minimize between-group variance, on one or more variables. The functional region, by contrast, is a phenomenon arising exclusively from human activity, and is best described as a community of interest. Although references to the concept can be found as early as the first half of the 20th century (Platt, 1928), and functionality is accepted as the basis of several popular landuse models, an acceptable definition of the functional region has been elusive. Various authors have proposed descriptions suggesting internal social homogeneity, economic heterogeneity and minimal contact with the rest of the universe. In the quantitative literature, the functional region has often been defined as that aggregation of ESUs which maximizes the ratio of within-region to between-region interaction. There are several operational difficulties with this definition, which will be examined in greater detail in the next section.

The third structural class, the nodal region, was recognized in the 1950s when Whittlesey (1954) introduced the concept of cores and regional dominance in networks. Nystuen and Dacey (1961) translated the concepts into a quantitative methodology for delimiting nodal regions. Finally, the equitable region is a device that has grown out of administrative needs where the primary objective is for regions to contain approximately equal populations — census and electoral districts are two examples. The literature on the problem and appropriate algorithms has been reviewed by Taylor (1979).

3. Delimiting the Functional Region

In the light of the foregoing discussions, it is clear that a methodology for functional regionalization would revolve around analysis of an interaction matrix. Let **I** be a $n \times n$ matrix of interactions between n ESUs. Typically, such a matrix may contain (a) null cells, representing no interaction and (b) missing or null diagonal cells, which arise from the difficulty of measuring interaction with oneself. In the sociology literature, these features have been treated as sampling errors; algorithms exist to estimate entries for these cells (for example, Savage and Deutsch, 1960). There is also debate on whether and how indirect flows should be included in the matrix (Tinkler, 1976).

As mentioned earlier, functional regionalization might be viewed as a search for an aggregation of ESUs that maximizes the ratio of within-region to between-region cell totals. There will usually be differences in size between ESUs; a simple minimization of the above ratio would tie the result of the regionalization to the geocoding system. While there is agreement on the need for standardization of cell entries, the methods of achieving this have been subject to prolonged and inconclusive debate (Goodchild, 1981).

Assume, nevertheless, that the matrix **I** has been suitably pre-processed. At this point there are two divergent approaches in the literature. On the one hand, Garrison and Marble (1963) and others applied a conventional factor analysis to cluster ESUs with others that exhibited similar connective structures. Russett (1967, pp 101, 130) criticized the objective of factor analysis as being inconsistent with the purpose of functional regionalization, which was to aggregate ESUs on the basis of *mutual* relationships, rather than relationships with *all other* ESUs. On the other hand, methods have been proposed to identify clusters of mutual interaction. Brams (1966) and Slater (1976) offered methods based on matrix decomposition; Ng (1969) and O'Sullivan (1977) employed the technique of matrix seriation; Hollingsworth (1971) and Masser and Brown (1975) used clustering procedures to aggregate ESUs while monitoring an appropriate objective function.

One concludes from reviewing the literature that the field of functional regionalization lacks agreement on critical issues at all levels. Methodologies range from Markov chain analysis to matrix decomposition, without a prior conceptualization of the behavioral origins of the functional region. Techniques have been imported liberally from other disciplines, none of which recognize the crucial geographic variable — distance. This often leads to an artificial imposition of contiguity on the derived regions (recall the gravity model). While the detection and delimitation of communities of interest is of concern in a variety of situations (two examples are political mediation and communications planning — see AT&T, 1980) geography lacks the conceptual and methodological resources to provide solutions in these areas.

The key to the understanding of functional regions must lie in a study of human behavior under non-isotropic conditions. The remainder of this paper attempts to model *interaction* across functional regions; the detection of regional boundaries is then achieved by fitting observed interaction matrices to the model.

4. The Black/White Model

The Spatial Interaction Model (SIM) has its antecedents in the Gravity Model, which modeled interaction between two points *i* and *j* as the product of their populations and an inverse function of the distance between them:

$$I_{ij} = K\, P_i\, P_j\, f(d_{ij}) \qquad\qquad \dots (1)$$

where

I_{ij} is the interaction between i and j,

K is a constant of proportionality,

P_i is the population of i and

$f(d_{ij})$ is the distance between i and j.

The principal criticism of the gravity model was that it lacked a theoretical foundation, and that its acceptance in geography was based on empirical fit in the absence of a null hypothesis. Further, the argument for population as the ideal surrogate for Newtonian mass was weak. SIM, on the other hand, specifies the mass terms as *emissivity* of the origin, E_i, and *attraction* of the destination, A_j, each of which may be composites of relevant variables:

$$I_{ij} = E_i\, A_i\, f(d_{ij}) \qquad\qquad \dots (2)$$

The formulation in (2) merely expresses interaction as the product of all forces specific to the origin, those specific to the destination and a function of the distance between them. The probabilistic interaction equation of Huff (1963), Wilson's (1967) entropy maximizing relationship and the Niedercorn and Bechdolt (1972) utility maximizing model can all be interpreted within this generic framework.

The interaction models cited above operate in an abstracted landscape of physiographic and cultural uniformity. Consider the points A, B, etc, in Figure 1.
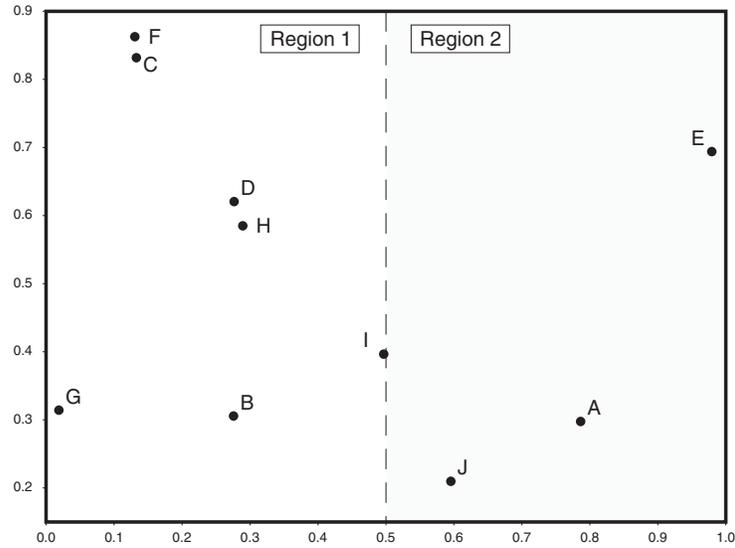
Figure 1. Black/White universe used in examples

Interactions between them, corrected for emissivity and attraction, normally plot as a decreasing function of distance (Figure 2). Now let the two regions indicated in Figure 1 represent two social groups. Their mutual affinity may be measured by a **compatibility** factor ρ ($0.0 \leq \rho \leq 1.0$).
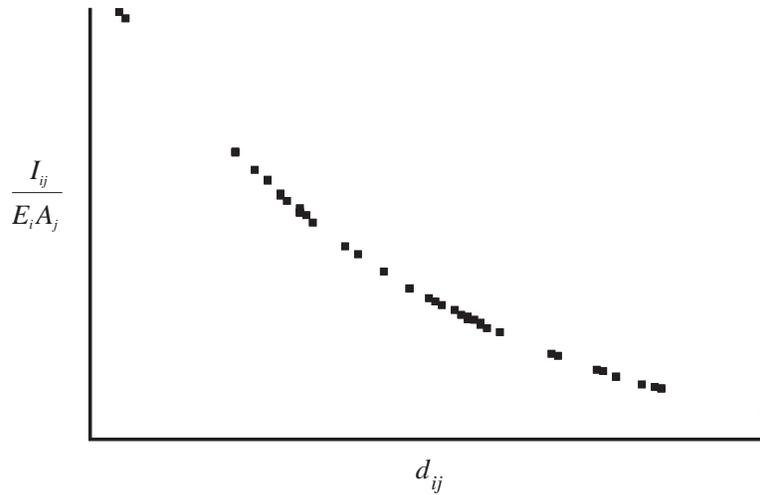

Figure 2. Single-region interaction profile

Except in very unusual cases where ρ could theoretically exceed 1.0, it may be viewed as the probability that individuals would not resist interaction across the regional boundary. A value of $\rho=1.0$ signifies complete disregard for regional distinction; a value of 0.0 indicates extreme sensitivity. The compatibility between any two ESUs i and j in a two-region universe may be expressed as

$$R_{ij} = \rho^{\,\sigma_{ij}} \qquad\qquad \dots (3)$$

where

$\sigma_{ij} =$ 0 if i and j belong to the same region,

1 otherwise   … (4)

Given the Luce (1959) choice axioms, and using the approach of Huff (1963), one may derive the equation:

$I_{ij} = E_i A_j f(d_{ij}) R_{ij}$ … (5)

This is the most general statement of the Black/White Model.

Now consider the pattern of interaction in Figure 1 in the light of the Black/White Model. Flows across the boundary will be reduced by the action of the compatibility factor, and will plot as a second curve (Figure 3). With an appropriate transformation of the axes, the curves appear as two parallel straight lines (Figure 4), their separation determined by ρ.
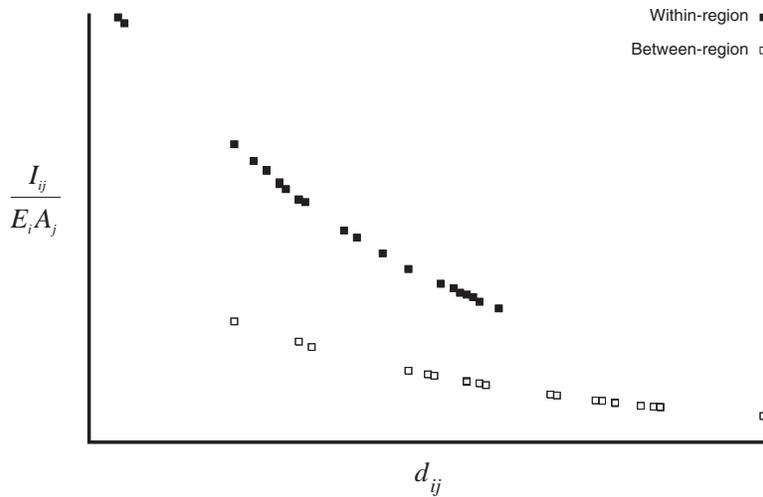

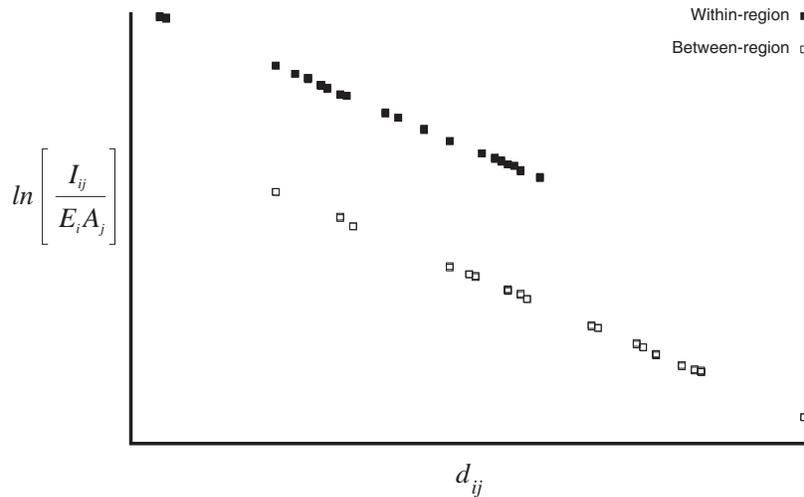Figure 3. Two-region interaction profiles

Figure 4. Linearized two-region interaction profile

Empirical support for this purely theoretical discussion comes from Mackay (1958). In an analysis of long distance calls from Montreal, he discovered three distinct scatters: the uppermost represented calls to points within the province of Quebec; a central scatter was composed of other Canadian destinations, while the lowest contained destinations in the United States (Figure 5).
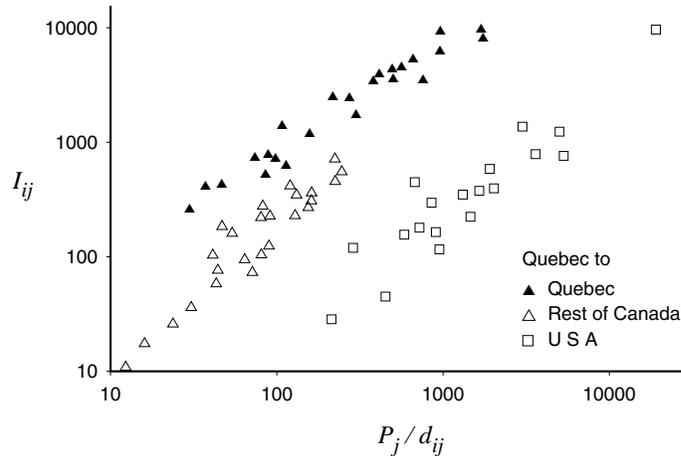

Figure 5. Phone traffic between Canada and USA (after Mackay)

Analysis of Mackay's plot revealed ρ to be 0.2 between Quebec and the rest of Canada, and 0.02 between Quebec and the United States.

The significance of Figures 4 and 5 is clear: since the scatters are differentiated by the regional memberships of the associated ESUs, it follows that in the absence of prior knowledge of regional affiliations of the nodes, an analysis of the dyads contributing to each scatter would reveal the regional memberships.

5. Black/White Regionalization

The above discussion assumes that the emissivities and attractions, **E** and **A**, are known. In typical applications, only **I** and **d** are known, and *f*, the function of distance is usually assumed *a priori* to be power or exponential; other values have to be found by calibrating the Black/White model. Unfortunately, the Black/White equation cannot be calibrated by conventional means, since it may contain more unknowns than knowns. One is forced to rely on estimates, **E'** and **A'**, which are found by calibrating a conventional SIM (that is, assuming no regional differentiation). The deviation of the estimates from the true vectors **E** and **A** will depend on ρ. In general, the error will result in the parallel lines (Figure 4) appearing as **parallel scatters** (Figure 6 shows a simulated example with $\rho=0.4$).
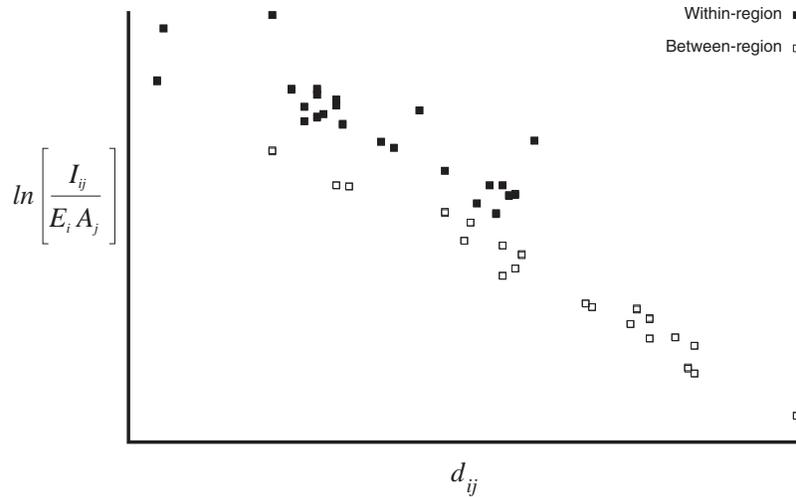
Figure 6. Parallel scatter

Under usual regression assumptions, the maximum likelihood partition of the scatters is associated with the minimum pooled within-scatter sum of squared ordinate residuals generated by all possible regional configurations. In general, there are $2^{n-1}-1$ possible assignments of n ESUs to two regions; these may be examined in sequence, or one may use heuristic methods to isolate a subset of likely configurations.

In tests with a simulated 10×10 data matrix, the method correctly recovered a variety of embedded regional configurations while ρ was varied between 0.01 and 0.99. Next, the tolerance of the method was examined by introducing measured random error into the simulations. The error tolerance, ε, varied with ρ (Figure 7).
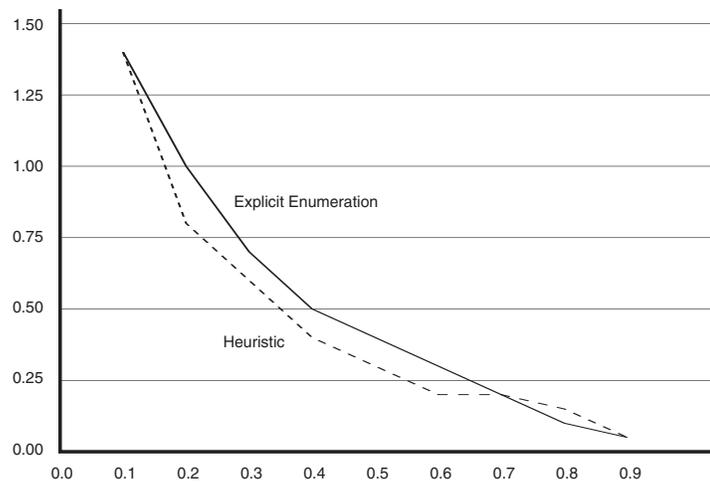

Figure 7. Error tolerance

To give the reader an appreciation for the epsilon scale, the parallel scatter associated with $\rho=0.2$ at the critical epsilon level ($\varepsilon=1.0$) is shown in Figure 8 — equivalent error applied to a straight line produced a scatter with a correlation coefficient of 0.5.
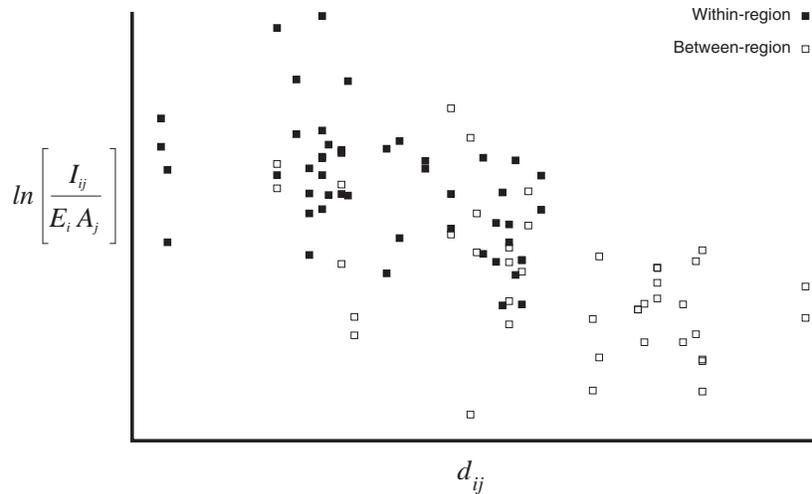

Figure 8. Parallel scatter, =0.2

6. Regions of the United States

The methodology developed in the preceding sections was applied to 49×49 matrices of flows of students between place of residence and place of education in the United States, for 1979 and 1981. The ESUs were the 48 contiguous states and Washington DC.[2] Based on Warntz's (1976) observations on similar data in previous years, it was expected that a partition might emerge between the northern and southern states. A solution by explicit enumeration of all configurations of 49 ESUs would have required over half a million years of computing time; a greedy-add heuristic was developed, which delivered solutions in about 2000 CPU seconds. Figure 9 is representative of the map patterns derived.

The dark shading identifies **polar** ESUs, around which **dominant** companions were aggregated during the execution of the algorithm. The **subdominant** states form the opposite cluster. While a general north-south pattern is clearly discernible, the identification of Nevada, New Jersey, Ohio and Pennsylvania with the south may conflict with traditional notions of state affiliations (interestingly, Washington DC belongs to the northern cluster). An exhaustive discussion of the pattern is beyond the scope of this paper.

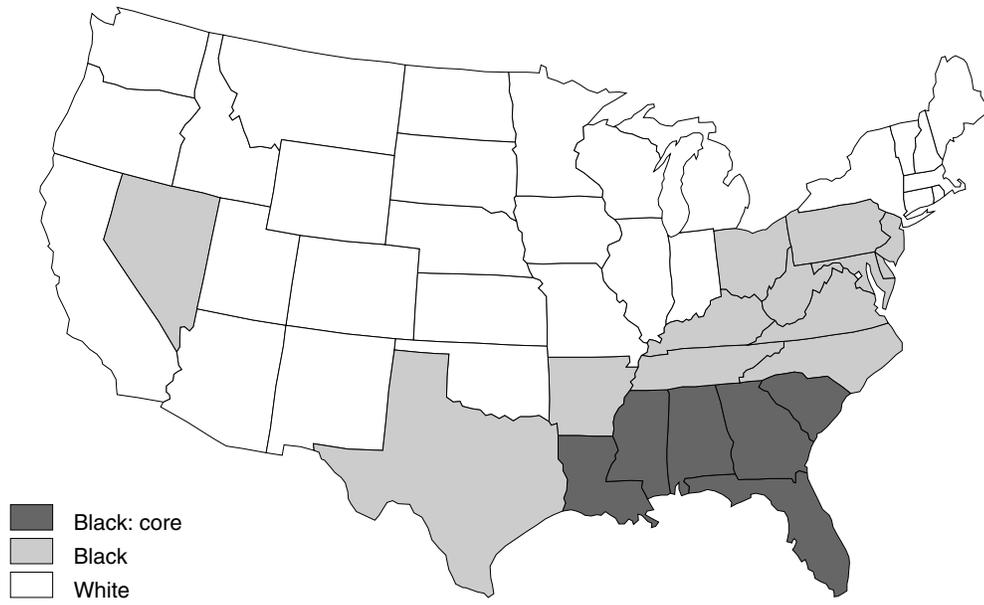[2] Data from National Center for Education Statistics, Washington

Figure 9. Regionalization of the United States

7. Conclusions

This paper has introduced the Black/White model as a conceptualization of functional regions, and as a mechanism for their discovery. This represents several advantages over earlier approaches to functional regionalization. First, the model is derived from choice theory in conjunction with spatial interaction — this theoretical foundation was lacking in much of the earlier literature. Second, since the model addresses the origins of interaction, it permits simulation of matrices; it is thus possible to verify the outcome of a regionalization procedure against a simulated target. Third, the regionalization procedure skirts the controversies over standardization for ESU size, since ESU-specific variations are accommodated by the spatial interaction model. Finally, the method recognizes the role of distance, and has been shown to produce non-contiguous configurations.

There is, however, scope for further research and testing of the procedures. First, the working of the model has been discussed in this paper with respect to *two* regions. Although the extension of the theory to multiple regions is straightforward, the computational logistics involved with testing the performance of the regionalization method are formidable; this will have to await advances in computing technology. Secondly, the amount of error with which the procedures can cope (Figure 8) is so great as to suggest that patterns may be erroneously extorted from random data. There is a need for measures of statistical significance associated with derived patterns.

In conclusion, the Black/White model offers a fresh outlook on a geographical problem that has substantial application potential, and may be expected to provide objective solutions to sensitive social, political and academic problems.

References

AT&T, 1980: *Notes on the Network. Section 2: Numbering Plan and Dialing Procedures* (American Telephone and Telegraph Company)

BERRY B J L, 1964: "Approaches to regional analysis: a synthesis" *Annals, Association of American Geographers* **54** 2-11

BRAMS S J, 1966: "Transaction flows in the international system" *American Political Science Review* **60** 880-98

BUNGE W, 1962: "Theoretical geography" *Lund Studies in Geography, Series C, General and Mathematical Geography, No 1* Gleerup, Lund

DICKINSON R E, 1969: *The Makers of Modern Geography* (Routledge and Kegan Paul, London)

GARRISON W L, MARBLE D F, 1963: "Factor analytic study of the connectivity of a transportation network" *Papers and Proceedings, Regional Science Association* **12** 231-38

GILBERT E W, 1951: "Geography and regionalism" in Taylor T G (Ed) *Geography in the 20th Century* (Methuen, London)

GOODCHILD M F, 1981: "Regionalization and aggregation from interactions and flows" *Association of American Geographers, Annual Meetings, Los Angeles*

GREENOUGH J, 1919: "Geographical record" *Geographical Review* **7** 110-111

GRIGG D B, 1965: "The logic of regional systems" *Annals, Association of American Geographers* **55** 465-91

HOLLINGSWORTH T M, 1971: "Gross migration flows as a basis for regional definition — an experiment with Scottish data" *Proceedings, IUSSP Conference, London, 1969* **4** 2755-65

HUFF D L, 1963: "A probabilistic analysis of shopping center trade areas" *Land Economics* **39** 81-90

LUCE R D, 1959: *Individual Choice Behavior: a Theoretical Analysis* (Wiley, New York)

MACKAY J R, 1958: "The interactance hypothesis and boundaries in Canada : a preliminary study" *Canadian Geographer* **11** 1-8

MASSER I, BROWN P J B, 1975: "Hierarchical aggregation procedures for interaction data" *Environment & Planning A* **7** 509-23

MINSHULL R, 1967: *Regional Geography: Theory and Practice* (Hutchinson, London)

NAGEL S S, 1965: "Simplified bipartisan computer redistricting" *Stanford Law Review* **17** 863-99

NG R C Y, 1969: "Recent internal migration in Thailand" *Annals, Association of American Geographers* **59** 710-30

NIEDERCORN J H, BECHDOLT B V, 1972: "An economic derivation of the gravity law of spatial interaction: a further reply and a reformulation" *Journal of Regional Science* **12** (1) 127-36

NYSTUEN J D, DACEY M F, 1961: "A graph theory interpretation of nodal regions" *Papers and Proceedings, Regional Science Association* **7** 29-42

O'SULLIVAN P, 1977: "The spread of traffic effects and regions of a network" *Environment & Planning A* **9** 917-26

PLATT R S, 1928: "A detail of regional geography: Ellison Bay community as an industrial organism" *Annals, Association of American Geographers* **18** (2) 81-126

RUSSETT B M, 1967: *International Regions and the International System* (Rand McNally, Chicago)

SAVAGE R I, DEUTSCH K W, 1960: "A statistical model of the gross analysis of transaction flows" *Econometrica* **28** 551-72

SLATER P B, 1976: "A hierarchical regionalization of Japanese prefectures" *Regional Studies* **10** 123-32

STEVENSON E L, 1932: *The Geography of Claudius Ptolemy* (New York Public Library, New York)

TAAFFE E J, GAUTHIER H L, 1973: *Geography of Transportation* (Prentice Hall, Englewood Cliffs)

TAYLOR P J, JOHNSTON R J, 1979: *The Geography of Elections* (Croom Helm and Penguin, London)

TINKLER K J, 1976: "On functional regions and indirect flows" *Geographical Analysis* **8** 205-14

UNSTEAD J F, MYRES J L, ROXBY P M, STAMP L D, 1937: "Report on classification of regions of the world" *Geography* **22** 253-82

WARNTZ W, 1976: "The university and the region" *Northeast Regional Science Review* **6** 18-47

WHITTLESEY D, 1954: "The regional concept and the regional method" in James P E, Jones C F (Eds) *American Geography: Inventory andProspect* (Syracuse University Press)

WILSON A G, 1967: "A statistical theory of spatial distribution models" *Transportation Research* **1** 253-69