Frame independent spatial analysis †

Waldo R. Tobler

Abstract

The results of an analysis of geographical data should not depend on the spatial coordinates used - the results should be frame independent. This should also apply when areal units are used as the spatial data collection entity. Previous work has shown that some analysis procedures do not yield the same results under alternate areal aggregations, but some of these studies have used measures known to be inappropriate for spatial data, e.g., Pearsonian correlation instead of cross-spectral analysis. And there are some methods of analysis that do seem to yield frame invariant results, especially under alternate partitionings of the geographic space. In other cases it is appropriate to consider aggregations as spatial filters, with response functions that can be estimated *a priori.* There also exist linear spatial models that allow exact calculation of the effects of a spatial aggregation, so that consistent empirical and theoretical results can be obtained at all levels of spatial resolution. It is proposed that all methods of spatial analysis be examined for the invariance of their conclusions under alternative spatial partitionings, and that only those methods be allowed which show such invariance.

From a philosophical point of view it is important that spatial analyses not depend on the units used to identify the geographical location of the objects being studied. In its simplest form this is an assertion that it should not matter whether one is identifying places by rectangular coordinates or by polar coordinates. In this overly simplified example everyone would agree that the names used to identify the locations are irrelevant to the substantive analysis. This same point of view should prevail when areal units are used. But it apparently does not. Openshaw (this volume) quotes Kendall and Yule (1950, p. 313) who warn that one must not lose "... sight of the fact that our results depend on our units". The units in this case are areal units for which agricultural statistics (wheat and potato yields) are assembled. This, I think, is very poor science, and represents a misconception. It is not the areal units that are to blame. The difficulty is that the method of analysis used was inappropriate. This tautology is immediate. If the procedure used gives results which depend on the areal units used, then, *ipso facto,* the procedure must be incorrect, and it should be rejected *a priori.* As an aside, one of the reasons why tensors are used for many calculations in physics is that they give results that are independent of the particular place names chosen. For example, the components of a gradient vector depend on the system of coordinates used, but the gradient itself is a concept independent of these units. We aim for the same type of frame free analysis in geography.

In the particular instance Kendall and Yule were computing correlation coefficients between areal units. That correlations between data sets assembled by areal units are subject to fluctuation has long been known. Openshaw (1984) cites a fifty year old paper by Gehlke and Biehl (1934), in which these authors observed that a correlation coefficient increased when they aggregated the data to larger areal units. Openshaw (1984) goes into some detail here, citing further comments by Kendall and Yule (1950), by Robinson (1950) in his well known study of ecological correlations, and by Blalock (1964).

The fallacy in all of these studies is the assumption that the correlation coefficient is an

appropriate measure of association amongst spatial units. Clearly it is not - the appropriate measure is the spatial cross - coherence function (see Rayner, 1971) and the association between the two variables may be different in different locations - but all of these authors put the blame on the spatial units. This fallacy is compounded when these authors do not recognize that the spatial frequency point of view quickly and easily predicts the types of results that they obtained. For example, Curry (1966) points out that "Administrative units having area dimensions represent a filtering out of wavelengths less than their size". And Casetti (1966) notes that "Aggregating smaller areal units into regions filters out the harmonics whose wavelengths are smaller than the size of the regions", and, if two (or more) "space series have harmonics which are filtered out by a given aggregation, the correlation and regression coefficients of the series before the aggregation will differ from the coefficient obtained after the aggregation." In spite of this clear theoretical understanding Openshaw (1984, p. 13 *et seq.)* feels compelled to perform extensive numerical and computer experiments with empirical data, and does manage to demonstrate that correlation coefficients do indeed perform in the expected unsatisfactory manner. Again the blame is put in the wrong place, on the areal units. Somewhat later Openshaw (1984) demonstrates that similar results hold for regression coefficients, and for a particular spatial interaction model as well as for the simpler measures of association. As noted above this was already anticipated theoretically. Other theoretical insights have not been pursued adequately either. For example Tobler (1969) suggested computation of the spatial response function of an areal partitioning as a method of adjusting for the filtering effects of the partitioning, and Moellering and Tobler (1972) demonstrated how to isolate the most important level(s) of an administrative spatial aggregation.

One of the difficulties, implicit in Kendall and Yule (1950) but explicit in Openshaw (1984), is that many kinds of geographical data inevitably seem to require reporting in some areal units, and that these units are always to some extent arbitrary. The inference is thus that the "modifiable areal unit problem" is unavoidable. This appears to be another fallacy, at least when stated in this naive way. Not all geographical problems are well posed. For example, did Kendall and Yule (1950) really need to use those areal units, or could the problem have been reformulated to be independent of the units? Could they, for example, have gotten the raw data on agricultural fields and used a form of near neighbor analysis (Getis and Boots, 1978), or could they have given the yields in the form of spatially continuous geographical probability density functions (see, e.g., Silverman 1986)?

This last approach is itself not without problems. If one asks for the cancer rate (cancers per 100,000 persons) at a particular latitude and longitude one can get a different answer if the data are computed from national observations, or from state data, or from county data, or from city data, or from census tract data, or from data by city block, or by house. Does this process have a limit? Is there an actual cancer rate at this place? Note the similarity to Richardson's (1926) question "Does the Wind have a Velocity?" We are told that the air, that water, are made up of discrete particles but we also observe that aero- and hydrodynamicists use partial differential equations for the study of these systems, and not quantum mechanics. How can one use calculus in such a situation? The answer, the textbook answer, is in the continium hypothesis. The books (e.g., Batchelor, 1967, pp.4-6) often have a diagram such as that in Figure 1, where the density of a gas (for example) is plotted as a function of the resolution. At some point the density oscillates erratically, and is not a well defined, useful quantity. The student is warned that the analysis procedures, theorems, and techniques that follow in the book do not hold in the vicinity of this region or below. Interestingly I have yet to find a book that is explicit and precise

about where this region occurs. The conclusion that I draw from this is that there may well be problems of areal units, but they are not the ones that have been studied, and have almost nothing to do with correlation and regression.

From this discussion it is clear that the "modifiable areal unit problem" really consists of at least two distinct problems. The first I label the partitioning problem. It can be imagined in this fashion. In some piece of territory there exist discrete (immobile) individuals with attributes. Think of these as dots on a geographical map. This continuous piece of territory is then partitioned into a set of N areal units (put boundaries on the map inside the territory), and some procedure is used to summarize the attributes within each subunit, and to compute a measure of association between the summary attiibutes. Then a *de novo* different partitioning of the territory, again into N areal units, is undertaken, with a comparable summarization of the attributes. To what extent do the associations between the attributes differ for these two partitionings?

There are obviously arbitrarily many ways in which these partitionings can be performed, and the areal subunits can differ in size, shape, and orientation. Perhaps pentominoes (Buttenfield, 1984; Gardner, 1959; Golomb, 1960), which fix the size, can help to study some of the questions via simulations. Does the value of N make a difference? Suppose N is not the same for the two partitionings? The usual situation in practice is that one is given two different sets of data, assembled by two different partitionings and must work with these data, which are all that one can get. Much of the literature suggests that use of data from two such incompatible areal partitionings is only possible by aggregation to some larger unit sizes where the partitionings happen to coincide. Such coincidence often occurs in bureaucratic/political hierarchical spatial partitionings. Here again the conventional notion may not be correct. Pycnophylactic interpolation has recently been suggested (Tobler, 1979) and studied (Rylander, 1986) as a method for converting data from one set of areal units to another, and appears to work quite well. Conversion of data from latitude and longitude to transverse Mercator coordinates does not appear to cause any difficulties. Why should conversion from census tract to school district cause problems? It is helpful to organize geographic conversion problems into a square table, with point coordinates, line coordinates, and areal coordinates along both the side and top of the table. Now fill in the complete table by considering the conversions (and their inverses) between each method of data recording. Areal data are frequently converted to centroids (area → point), lat/lon to UTM (point → point), street addresses to State Plane Coordinates (point → point) or to census tracts (point → area), and so on. Most geographical information systems contain a number of such conversion possibilities. Further experimentation with the types of invariances that can be obtained under such transformations does not appear difficult (see Arbia, this volume, for example).

The second type of "modifiable areal unit problem" I refer to as a true aggregation. Here one starts from data in areal units and, for some reason, groups some of these units together into larger, and consequently fewer, units. This has the effect of coarsening the resolution of the data, where the average resolution is defined as the square root of the size of the territory divided by the number of areal units. The size of the smallest detectable pattern is of course twice that of the resolution. Most analyses are degraded by such a procedure, particularly if it changes the variance in resolution. Here we can point to some positive results, even though correlation and regression may be useless. In the migration model of Dorigo and Tobler (1983) it can be shown that it is possible to calculate exactly all of the model parameters when one combines data from areal units (Figure 2). Aggregation is thus not at all a problem in this model. The only question is why one would want to do it. Although not aggregation invariant the model changes in an

exactly predictable manner. The popular entropy movement model suggested by Wilson (1967) does not have this property and must be recalibrated, with apparently unpredictable results, for every alternate aggregation of the data.

There is another sense in which aggregation is simpler than the partitioning problem. The spatial frequency response point of view allows one to consider the effects of an aggregation to be similar to that of a spatial filter, generally a low-pass filter (see Holloway, 1958; Burr, 1955). The accompanying figures illustrate this dramatically. In each instance the same analysis was performed on data given at different levels of resolution, and in each case the results are as if one had passed a low pass filter over the results of the higher resolution analysis. In this analysis there is no "modifiable areal unit problem". The problem has gone away when we use the correct analysis procedure.

Figure 1 Density as a function of resolution:

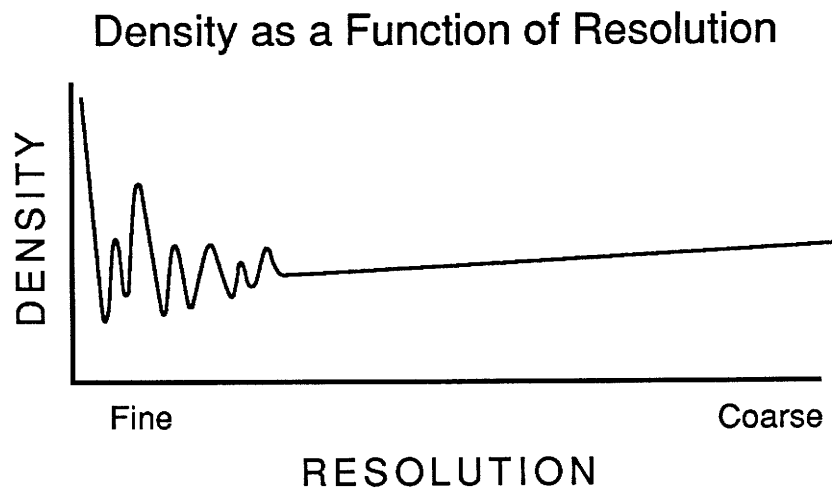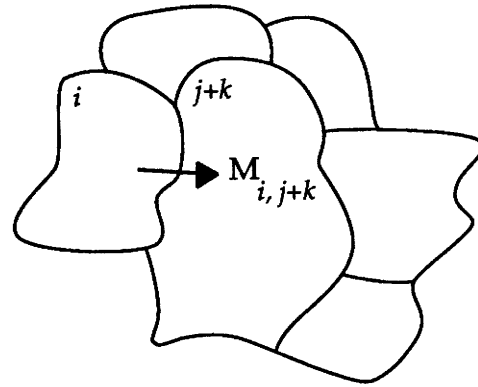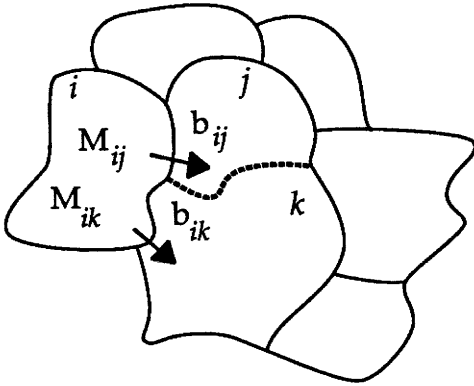Density as a Function of Resolution

Figure 2

The model asserts that migration from i to j is equal to the push from i plus the pull from j times the length of the boundary between i and j. The pull value for the

Computing Migration Before and After Aggregation

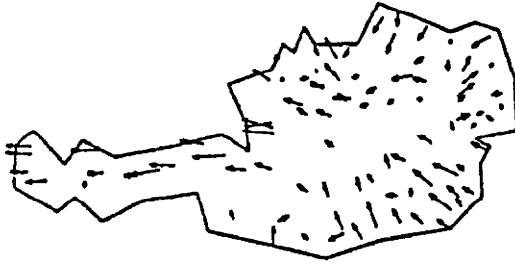$$M_{ij} = (R_i + E_j)\, b_{ij}$$

$$M_{ik} = (R_i + E_k)\, b_{ik}$$

$$M_{ij} + M_{ik} = M_{i,j+k}$$
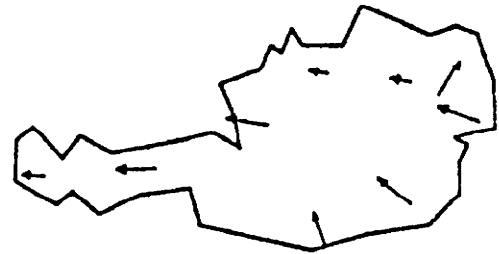
$$= R_i\,(b_{ij} + b_{ik}) + b_{ij}\,E_j + b_{ik}\,E_k$$

combined region i + j is easily computed from the parameters estimated for the disaggregate mode!, as explained in Dorigo and Tobler, 1983.

Figure .3. Reducing resolution acts as a spatial frequency filter:

Migration in Austria

Computed from
98 by 98 Gemeinde data

Computed from
9 by 9 Laender data

Figure 4. Migration patterns computed at three levels of spatial resolution.
3090 Gemeinde, 184 Bezirke, 22 Cantons in Switzerland.
Courtesy of G. Dorigo, University of Zuerich.
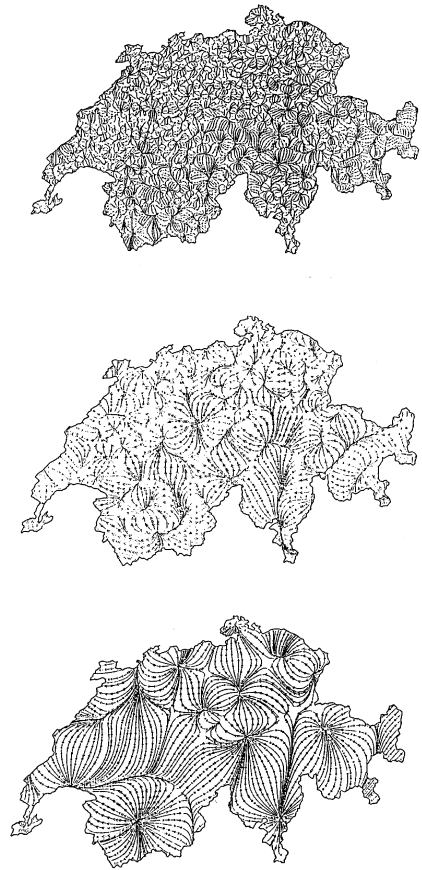**NET** migration computed from the attractivity function.

Figure 5. Migration patterns at three levels of spatial resolution.
3090 Gemeinde, 184 Bezirke, 22 Cantons in Switzerland
Courtesy of G. Dorigo, University of Zuerich
GROSS migration computed from the turnover function.

REFERENCES

Arbia, G., 1988, "Statistical Effects of Spatial Data Transformations", Santa Barbara, these proceedings.

Batchelor, G., 1967, An Introduction to Fluid Dynamics, Cambridge, University Press.

Buttenfield, J., 1984, "Pentominos: A Puzzle Solving Program", Compute!, May; 106-122

Blalock, H., 1964, Causal Inferences in Non Experimental Research, Chapel Hill, U. North Carolina Press.

Burr, E., 1955, "The Sharpening of Observational Data in Two Dimensions", Australian Journal of Physics, 8:30-53

Casetti, E., 1966, "Analysis of Spatial Association by Trigonometric Polynomials", Canadian Geographer, 10:199-204

Curry, L., 1966, "A Note on Spatial Association", Professional Geographer, 18:97-99.

Dorigo, G., & Tobler, W. , 1983, "Push-Pull Migration Laws", Annals, AAG, 73:1—17.

Gardner, M., 1959, Mathematical Puzzles & Diversions, New York, Simon & Schuster; pp. 124-141.

Gehlke, C., and H. Biehl, 1934, "Certain Effects of Grouping upon the Size of the Correlation Coefficient in Census Tract Material", J. Am. Stat. Assn., Supplement, 29:169-170.

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

Golomb, S., 1960, Polyominoes, New York, Scribrier

Holloway, J., 1958, "Smoothing and Filtering of Time Series and Space Fields", Advances in Geophysics, 4:351-389.

Kendall, M., and G. Yule, 1950, An Introduction to the Theory of statistics, London, Griffin.

Moellering, H., & Tobler, W., 1972, "Geographical Variances"' Geographical Analysis IV, 1:34-50

Openshaw, S., 1984, The Modifiable Areal Unit Problem, Catmog # 38, Norwich, Geo Books

Openshaw, S. , 1988, "Learning to Live with Errors in Spatial Databases", Santa Barbara, these proceedings.

Rayner, J., 1971, An Introduction to Spectral Analysis, London, Pion Press.

Robinson, A. , 1950, "Ecological Correlation and the Behaviour of Individuals", Am. Soc. Rev. , 15:351-357.

Richardson, L. , 1926, "Does the Wind have a Velocity", Proc. Roy. Soc., A, 110:709 et seq.

Rylander, G., 1986, "Areal Data Reaggregation", Master's Thesis, Michigan State University, East Lansing, Department of Geography

Silverman, B., 1986, Density Estimation for Statistics and Data Analysis, New York, Chapman & Hall.

Tobler, W. 1969, "Geographical Filters and Their Inverses", Geographical Analysis, 1,3:234-253.

____ 1979, "Smooth Pycnophylactic Interpolation for Geographical Regions", Journal, American Statistical Assn., 74,367:121-127.

Wilson, A., 1967, "A Statistical Theory of Spatial Distribution Models", Transportation Research, 1:253-269.