

AREOGRAPHIC REPRESENTATION OF FAUNAL CHARACTERISTICS
THROUGH A 'SECOND-ORDER' RELATIONAL APPROACH

225

Charles H. Smith
Department of Geography
220 Davenport Hall
University of Illinois
Urbana, Illinois 61801

Received April 21, 1983; July 1, 1983

ABSTRACT: Areographic analysis has traditionally depended on primary data consisting of location-specific tallies of presence or absence of given forms. In the present work, an alternate manner of representing distributions is suggested. Regional units are first established, and presence and absence of the forms under consideration in these units is noted. The relation of the biota at any given point location to all others is then established through examination of the former's characteristics of inclusion in the latter. This is accomplished by detailing regional level trends of inclusion and setting up a "second-order" distribution of associations. Two kinds of descriptive information that may be obtained through this approach are discussed, along with an example of a more exploratory kind of analysis grounded in the same mode. Discussion centers on the flexibility of the approach, which may be used to help interpret problems of either historical or ecological nature.

*

*

*

The author has elsewhere (Smith, 1983a, 1983b) suggested a set of guidelines regarding faunal regions classification and used this understanding to prepare a systemization of world mammal faunal regions. In these works, it is argued that efficient and logical hierarchical classification of faunal patterns is not only possible, but desirable in the sense that the resulting information might be used to support the development of models more appropriate to unbiased investigation of the evolution of biogeographic systems. Here, we shall be concerned with a means of further developing this perspective. The distribution of mammalian families is portrayed through a representational mapping procedure that is closely linked to the aforementioned regionalization scheme. The system characteristics of association of particular families with particular subregions are used to generate information that can be used secondarily to enhance description of the faunal characteristics of any given point on the earth's land surface. When a large number of individual locations are so treated, isopleth maps can be prepared from the resulting data which permit new kinds of insight into the causal structure underlying system differentiation.

First and Second-Order Forms of Data Representation

Serious problems confront the investigator attempting to characterize global level distribution patterns. While we often map absolute diversities or simple faunal association measure values pertaining to the members of a particular taxon, these efforts usually reward us with little more than "what is where" factual information. Rapoport (1982) has recently summarized a number of areographic analysis procedures used to deal with such types of data; both the analytical techniques and the data themselves might be referred to as "first-order" representational efforts for their attention to patterns emerging directly out of location-specific "presence/absence" summations. At least two major problems are faced in the interpretation of such systems of representation: 1) the proper framing of analytical studies concerning the causal processes ostensibly giving rise to the observed patterns of distribution; and 2) the impossibility of referring data of this type to any explicit measure of validity: there being no a priori way to establish a relational context across space for any two particular datum values, there can likewise be no standardized meaning attached to differences between them.

*

*

*

*

*

*

*

Evolutionary Theory 6: 225-232 (August, 1983)

The editors thank C.B. Cox and another referee for help in evaluating this paper.

© 1983, The University of Chicago; rights owned by author.

The first problem, involving the simplicity fallacy, is of involved philosophical nature and cannot be dealt with here. In passing, however, it should be noted that naive forms of pattern analysis of first order data representations are risky ventures. It is well known within geographic circles, for example (see discussions by Harvey, 1969; Amedeo and Colledge, 1975; Gale, 1975; and Getis and Boots, 1978) that entirely different causal processes can give rise to identical spatial patterns.

Varying philosophical positions regarding the nature and importance of the second problem alluded to above have created a fundamental disagreement as to what constitutes "proper" biogeographic investigation. Cladists/vicariance biogeographers have been quite vocal in pointing out the fallacy inherent in the studies of "classical" dispersalist biogeographers such as Wallace, Darlington, Mayr, and Simpson of assigning number values to relational measures based on first order distributional data. They argue that the meaning of a similarity value constructed around such numbers cannot be assessed unless it can be shown that identical evolutionary frames of reference obtain for each pair of faunas in question, a nearly impossible condition to meet. The same and related points have been levied at current advocates of the MacArthur and Wilson school of geographical ecology (see, for example, discussion by Cracraft, 1982). So far, there has been little productive response to these criticisms by the latter group. Vicariance biogeographers, meanwhile, have sinned in their own way by advocating pattern analysis procedures that are not strictly scientific for their prior selective rejection of "non-appropriate" patterns within the primary framework of their method. Such selectivity violates the spirit of unbiased pattern analysis and is cause for great concern since it is likely to promote inefficient, fallacious, and narrow-minded courses of investigation.

I should like to suggest here a conceptual/representational tool that might be employed to help circumvent some of these main issues of contention among dispersalist, vicariance, and ecological approaches to biogeographic investigation. I have elsewhere (Smith, 1983a) introduced a world regional mammal faunas classification (see Figure 1) with a property of relevance to the present discussion. This is that subregional divisions have been set such that each specific fauna is as unique as it can be with respect to the sum of all such affinities among all subregions (a variance-maximizing criterion). The multidimensional scaling configuration used to interpret this system of relations is "evenly packed"; i. e., the sum of distances squared among all points in the solution configuration has been minimized. Under these conditions, the mean distance from each point in the configuration to all others is approximately the same for each point. The latter property is particularly useful because it means that each subregion in the system, regardless of its location-specific attributes, is logically/functionally equivalent to each other subregion with respect to its contribution to total system order. (That this is so can be clearly demonstrated through a number of regression analyses involving the use of system level summary statistics to predict the faunal attributes of individual subregions; see Smith, 1983b.) Thus, given the equal weighting of each subregion in the specification of the system, the presence of any particular family in any particular subregion emerges as an equal bit of information in the delineation of world regional faunal patterns.

This fact can be used to characterize the fauna of any particular location on the earth's surface by first giving attention to the total number of subregions that each faunal element found at the given location is present in on a worldwide basis. A distribution of relational values (one for each form specific to the location and ranging in possible magnitude from one to a maximum equal to the number of subregions in the classification) can be tallied from which conventional descriptive statistics may be extracted. Importantly, any such statistic referable to a particular point location is directly comparable to that obtained for any other location, because the basis of comparison is a global frame of reference (i. e. the condition allowing meaningful interpretation of the faunal characteristics at any given location is that each situation be expressed in terms of system level properties). This approach to data representation might be termed "second-order" by analogy to the second-order

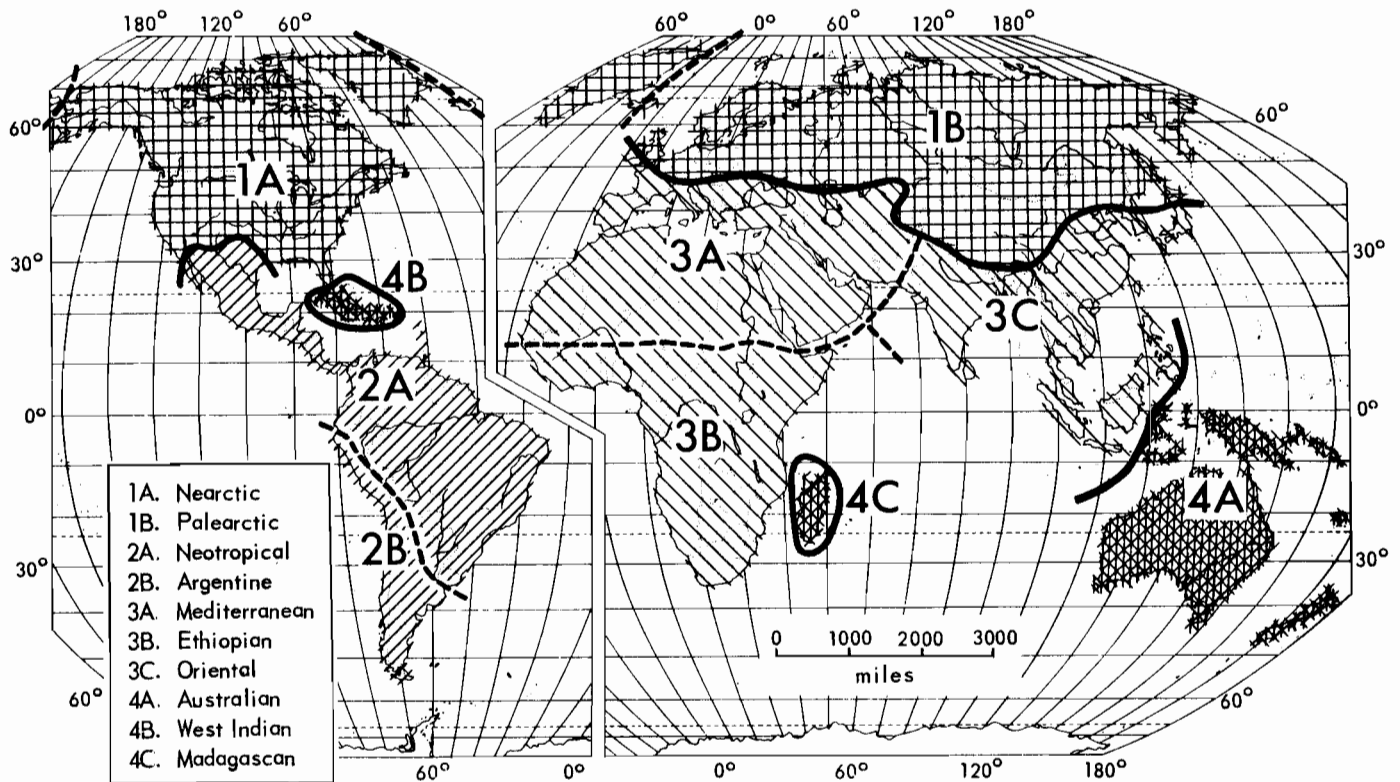


Figure 1. A classification of world regional mammal faunas (after Smith, 1983a). The classification reflects an attempt to order the mammal faunas of the world in such a way as to minimize the sum of squared distances within a multidimensional scaling configuration representing the scaled similarities between all pairs of subregions, a "minimum-entropy" solution. Subregional names are as indicated in the legend; regional divisions are named: 1. Nearctic; 2. Latin American; 3. Afro-Tethyan; 4. Island.

(nearest-neighbor) analysis of point patterns; the major difference lies in the fact that a distribution of relative "presences" is obtained rather than a distribution of distances. In the "second-order" representations of organic distribution described here, therefore, attention is focused not on the attributes of any one, pair of, or entire set of locations, but instead on partial orderings of system-level distributional attributes.

Underlying the maps presented in Figures 2 and 3 are descriptive "second-order" statistics associated with the primary data for 504 point locations. In Figure 2, the statistic mapped is the mean of each second-order distribution, a value that can be referred to as the associated fauna's "mean relative cosmopolitanism". Examination of Figure 2 thus leads to a picture of the variation over space of the mean number of subregions that the mammal families of particular locations are present in. Concerning Figure 2, it can be seen that the values obtained for the mean relative cosmopolitanism statistic range in magnitude from under 3.4 to over 6.6. Highest values are generally associated with suboptimal environments (usually desert or high latitude areas), lowest values with low-latitude conditions of high long-term paleogeographic and/or paleoclimatic stability and currently benign climates. Exceptions occur, however; several peninsular effects are visible, and slight reversals of the general latitudinal trend are commonplace. The peninsular effects are particularly interesting because both positive and negative deviations exist. For example, the mean relative cosmopolitanism values in India and Indochina are lower than those of the areas surrounding them, whereas those in Florida, Korea, and Arabia are higher than those of

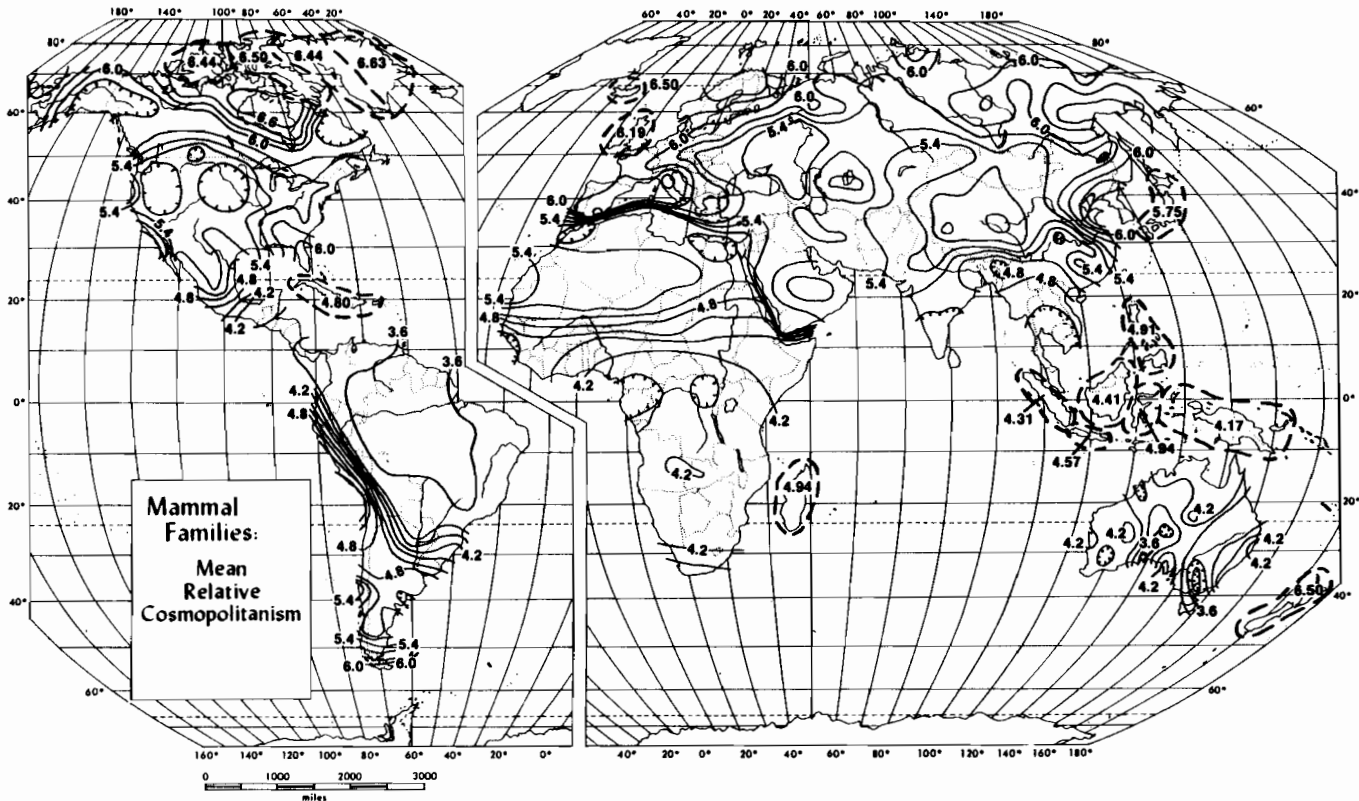


Figure 2. Isopleth map expressing spatial variation in the mean relative cosmopolitanism of world mammal faunas. A value of 5.0 at a given location indicates that the mammal families found there occur, on the average, in five of the subregions shown in Figure 1. Those values found within dashed lines are mean relative cosmopolitanism figures derived for entire islands, parts of islands, or island groups.

surrounding areas. This pattern is likely the result of the differing historical positions of these areas with regard to their relation to diffusing faunas. Subcontinent India, for example, was until recently an "ark" (McKenna, 1973) that eventually provided an important source of new forms for the Asian continent when the two land areas came into contact. Florida, on the other hand, is quite remote from the main sources of mammalian input to North America and emerges as a sink within system affinities.

Figure 3 is a mapping of the population standard deviations associated with the same 504 second-order distributions. This statistic is a bit more remote from intuitive appeal, but nonetheless relays interesting information upon interpretation. The values mapped (which range from under 1.7 to over 3.6) indicate the variation in relative cosmopolitanism of faunal elements inherent at given point locations. Obvious spatial trends are not apparent in Figure 3; there seems to be some correlation, however, between high-latitude and/or suboptimal habitats and faunas whose elements vary less with respect to their degree of mean relative cosmopolitanism. But interesting exceptions to this trend abound; note for example the southern Andes area. This is an area with relatively high variation in the cosmopolitanism of its mammal fauna (as compared to the rest of South America). Such local high variations are likely attributable in general to conditions of recent intermingling of structurally differing faunas (which probably explains the local high in values over the Central American region) or simple juxtaposition of highly restricted and highly cosmopolitan groups (as occurs, for example, on large islands: note the high values for the West Indies, Madagascar, and Australia). The first explanation probably accounts for the southern Andes situation, where an adjustment condition is being occasioned by the relatively

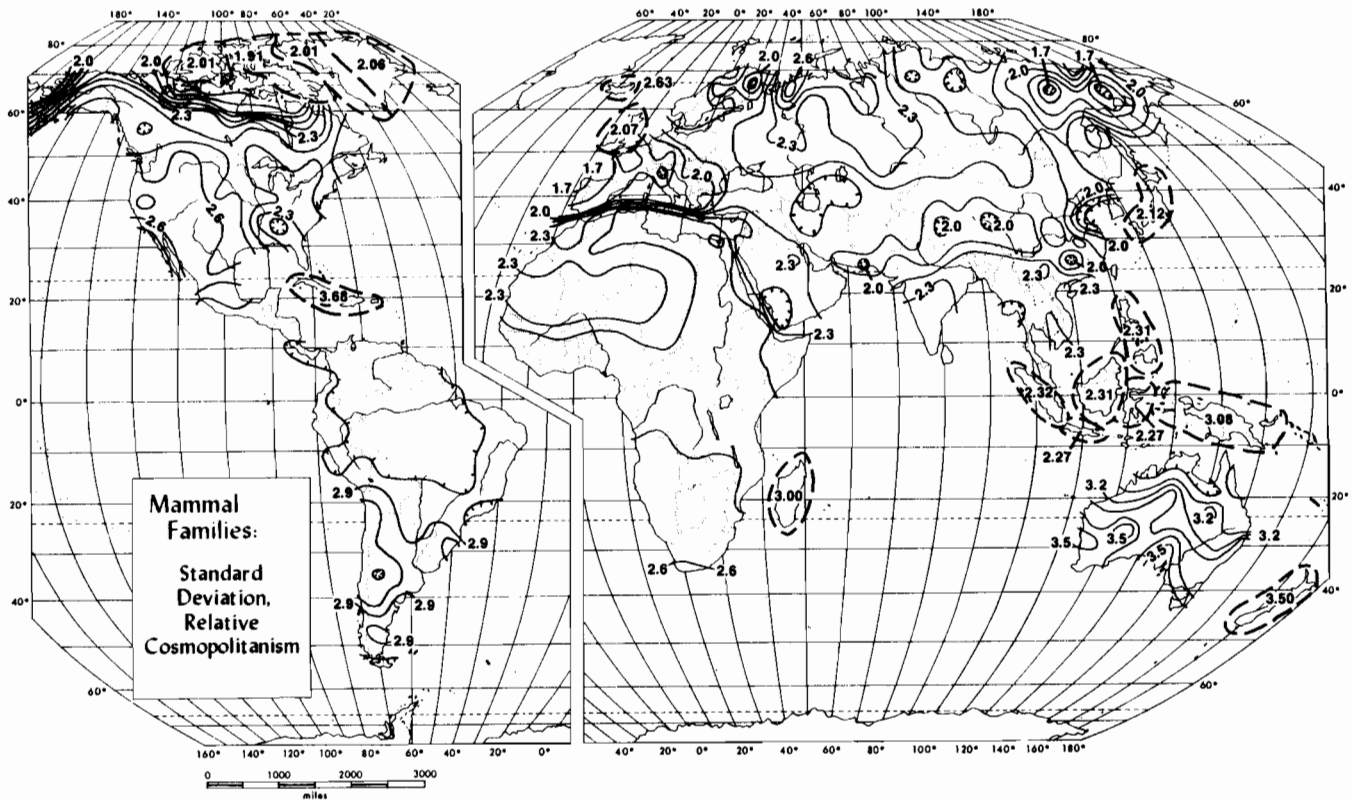


Figure 3. Isopleth map analogous to Figure 2, expressing spatial variation in population standard deviations. See text for further explanation. Those values found within dashed lines are population standard deviations for entire islands, parts of islands, or island groups.

recent introduction of a number of northern cosmopolitan groups.

Other mappings of relevant statistics based on the procedure presented here will be included in a work now in preparation. The information mapped here (and accompanying discussion) is intended primarily to exemplify, and, hopefully, to serve as impetus for further consideration.

Exploratory Analysis using the Second-Order Approach

The author has presented an analysis (Smith, 1983b), also based on the classification system under application here, in which subregional mammalian faunal differences were put into the terms of an entropy maximization approach (Wilson, 1970). Interfaunal similarities were therewith modelled as an inferred system of flows whose relative magnitudes could be explained as a function of: 1) contiguity constraints (producing a distance decay effect on the diffusion of evolutionary innovation over space and time); and 2) random rates of faunal interaction among subregional units (assumed to be the otherwise most probable state of the system). This "no prior assumptions made" approach translates into ten very efficient regression models (each with multiple R^2 of over .98), one each to explain each subregion's affinities with all other subregions. The preferred mode of approach using Wilson's method is first to conceptually account for as much of the variance in the faunal similarities ("flows") matrix as possible through that which can be attributed to conditions of random interaction, and then to proceed to a consideration of the residuals, within which is manifest the distance decay effect (and therefore the indication of contiguity constraints). As long as each areal component in the system is considered potentially both source and sink with regard to its effect on, and influence from, other

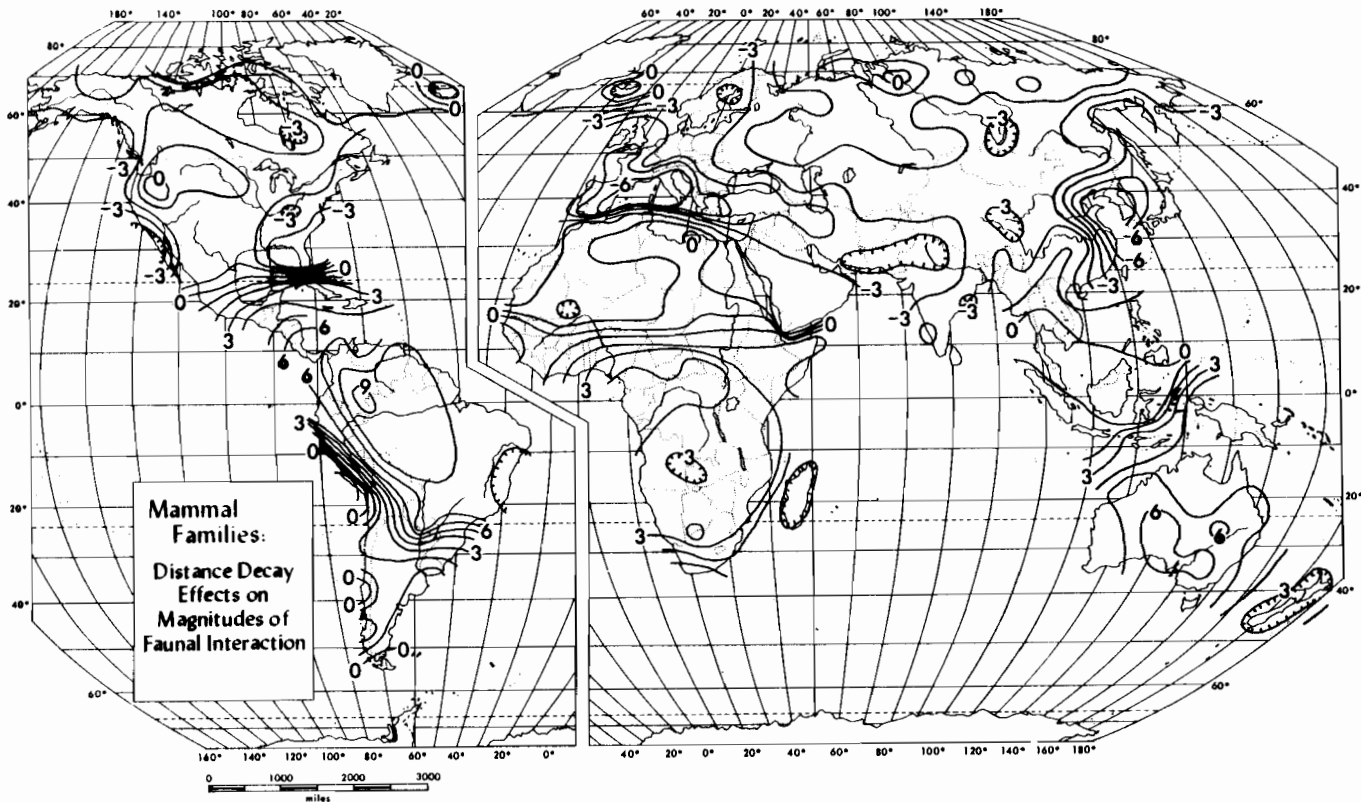


Figure 4. Residual map expressing spatial variation in the degree of influence of all specific subregional mammal faunal groupings on the make-up of the faunas at particular locations. High positive residuals indicate the dominating influence of one (or possibly more) subregional fauna on local faunal make-up; low negative residuals, greater lack of specificity of affinities. See text for further explanation.

faunas, and has relatively equal logical standing as an OGU, the resulting configuration of relationships may be viewed as an effective alternative to or generalization of the MacArthur and Wilson (1963, 1967) island-biogeography accounting framework.

In the subregional-level investigation of contiguity relationships reported in Smith (1983b), the distance-decay surrogate employed (for reasons discussed in that work) was comprised of the same statistics mapped in Figures 2 and 3. In that analysis, therefore, the second-order descriptive statistics were derived after first tallying membership of families on a by-subregion basis. It is also possible, however, to perform basically the same analysis using point locations as the OGU's in the regression model instead of entire subregions. This application of the second-order relational device allows us to obtain a much more detailed and interesting mapping of the distance-decay effect and provides a better starting point for more in-depth investigations. In Figure 4 above are mapped the residuals from a simple linear regression model involving the relative cosmopolitanism characteristics at 400 worldwide point locations. The values predicted in this analysis were the absolute diversities (number of mammal families present) at these locations. The values mapped are the residuals left once the effects of random interaction among locations are removed, and therefore inherently relate a system-standardized distance-decay effect. The mapping of the residuals as themselves indicating the effect is legitimized by the fact that these are almost entirely accounted for in the multiple regression analysis incorporating the contiguity effect surrogates into the model.

A look at Figure 4 suggests that the net domination of all specific subregional faunas on all other location-specific faunas becomes generally less as latitude increases. At least three things almost certainly contribute heavily to this pattern,

namely latitudinal associations with: 1) degree of paleogeographic and paleoclimatic stability over long periods of time; 2) varying degree of past and present physical contiguity of land surfaces; and 3) degree of severity of present ecological conditions. Deviations from the trend can be understood on the basis of the imperfection of the three correlations listed above and more specifically on the possible effects of a number of other influences that will require more study before they can be greatly elaborated upon. One particularly intriguing deviation might be explained as having resulted from human modification of landscape. Mediterranean Europe and coastal China each exhibit strongly negative residual patterns in Figure 4. Given the facts that large-scale human manipulation of these areas is of long standing and that the primary data used to derive the patterns mapped here depict the nature of present mammalian distributions only, it is tempting to believe that these residual patterns (of overestimation) belie, at least in part, human-based long-term selective removal of forms from these areas. The latter are thus geographically "remote" in the sense that they are populated only by those families whose component forms are, relatively speaking, generalists tolerant of such habitat de-stabilization. This is corroborated by the fact that these areas also exhibit high mean relative cosmopolitanism values, as can be seen in Figure 2.

Discussion

The main object of the preceding analysis has been to show how to represent distributional data in a relational manner more conducive to system-level studies. By explicitly equating the faunal characteristics of any given location to a relational state between that fauna and the faunas of the rest of the world, a system of representation is obtained which better conveys the essence of distributional patterns. This rationalist approach to classification (the general philosophy involved extends back at least as far as the time of Spinoza and Leibniz) produces an important dividend. When set within the context of "second-order" analysis, the information inherent in a regional classification recovers that lost in the aggregation procedure that led to its own development. In fact, not only is there no loss of information in the activity, but it appears that we are actually left with a net gain: the values produced are explicitly standardized with respect to the system as a whole and can therefore be directly related to one another across space. It is thus possible to avoid the principle fallacy associated with the study of systems: the idea that the elements of a highly ordered relational grouping can be meaningfully measured in isolation from one another.

Given an internally consistent means of measuring the variation inherent in biogeographical systems, the next step should be the erection of hypotheses concerning the processes resulting in that variation. For example, we might suggest, as I have done elsewhere (Smith, 1983b), that the present characteristics of distribution have devolved as a simple function of the spatial structure of the system, which has constrained the otherwise entirely random process of biological innovation and diffusion into an ordered response: regional faunas. It is useful to think in these kinds of terms because they potentially give us an intelligible interpretation of the relation between evolutionary structure and function, a subject of much interest these days (see related discussion by Gould and Lewontin, 1979). Regional faunas are the essence of evolutionary structure in the zoogeographic sense; their present state may be viewed as one time slice of the continuing enactment of some set of general energy-conserving principles manifest as geographically-varying rates of dispersal, speciation, extinction, etc.

In the more immediate practical sense, the main advantage of the representational approach advocated here is that it permits a reasonably unbiased view of the structure of biogeographical systems; i. e., one which is not initially colored by preconceptions associated with process-linked interpretations of distribution (see related complaints by Eldredge, 1981). Nonetheless, neither does it directly contradict such views. It is as possible, for example, to identify centers of endemism via

this style of pattern analysis as it is to recognize the effects of ecological constraints on organismal diversities. Concerning the former, it should be apparent that the relative cosmopolitanism concept developed here should be highly relevant to the consideration of patterns of endemism. Specifically, the various statistics that can be associated with this (such as the mean and standard deviation discussed earlier) could prove useful in providing an objective format within which interrelated instances of vicariance/dispersal can be sorted out. These same statistics could also be used to test ideas on the effect of ecological/spatial controls on distribution; two examples were given earlier but many more can be imagined. In short, the method permits flexibility, and hopefully this flexibility might serve to help bring some now very inflexible schools of biogeographic thought closer together.

Acknowledgments

The preparation of the manuscript was aided by the helpful comments of Susan L. Coon and Larry A. Hauser of the University of Illinois. Cartographers Chris Dunbar and Alan Hobscheid assisted in the preparation of the maps.

References

1. Amedeo, D., & Golledge, R., 1975. An introduction to scientific reasoning in geography.
2. Cracraft, J., 1982. Historical analysis versus ecological determinism: are there really two methods of biogeography? (pre-publication draft)
3. Eldredge, N., 1981. Discussion (of "The riddle of dispersal: dispersal theories and how they affect vicariance biogeography", by M. D. F. Udvardy). in G. Nelson & D. E. Rosen, eds., *Vicariance biogeography; a critique* (Columbia University Press, New York): 34-38.
4. Gale, S., 1975. Simplicity, again, isn't that simple. *Geographical Analysis* 7 (4): 451-455.
5. Getis, A., & Boots, B., 1978. *Models of spatial processes*. Cambridge University Press, Cambridge.
6. Gould, S. J., & Lewontin, R. C., 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. Royal Soc. London Ser. b*: 581-598.
7. Harvey, D. W., 1969. *Explanation in geography*. Edward Arnold, London.
8. MacArthur, R. H., & Wilson, E. O., 1963. An equilibrium theory of insular zoogeography. *Evolution* 17: 373-387.
9. MacArthur, R. H., & Wilson, E. O., 1967. *The theory of island biogeography*. Princeton University Press, Princeton, New Jersey.
10. McKenna, M. C., 1973. Sweepstakes, filters, corridors, Noah's arks, and beached Viking funeral ships in palaeogeography. in D. H. Tarling & S. K. Runcorn, eds., *Implications of continental drift to the earth sciences* (2 vols., Academic Press, New York): 295-308.
11. Rapoport, E. H., 1982. *Areography*. Pergamon Press, Oxford, England.
12. Smith, C. H., 1983a (in press). A system of world mammal faunal regions. Part one. Logical and statistical derivation of the regions. *J. Biogeogr.* 10.
13. Smith, C. H., 1983b (in press). A system of world mammal faunal regions. Part two. The distance decay effect upon interregional affinities. *J. Biogeogr.* 10.
14. Wilson, A. G., 1970. *Entropy in urban and regional planning*. Pion Limited, London.