ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Bivand, Roger

Conference Paper

Red Herrings and Club-Convergence: Lessons from Macroecology for Modelling Regional Growth

50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Bivand, Roger (2010) : Red Herrings and Club-Convergence: Lessons from Macroecology for Modelling Regional Growth, 50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/118872

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Red herrings and club-convergence: lessons from macroecology for modelling regional growth*

Roger Bivand
Norwegian School of Economics and Business Administration†

Abstract

Ecologists are often interested in studying the relationship between species and the environment, and between changes in the environment and changes in species occurrence. They consider that the spatial scales of the data generating processes for both the environmental and species variables matter, and that observed spatial dependence may be a "red herring", because the process scales of variables have not been captured appropriately. In studying regional growth, one is often obliged to use administrative entities that may not correspond to the scales of data generating processes. The paper will review the red herring controversy in macroecology, and relate it to the discovery of spatial regimes in connection with the club-convergence hypothesis.

1 Introduction

Spatial data analysis may involve issues of spatial scale, spatial autocorrelation, spatial structural change, and the matching (or mismatching) of units of observation to formative conceptual and empirical spatial processes. Crucially, the models that are estimated in both macroecology and in studies of regional growth raise operationalisation questions in relation to variables and to units of observation. Often, the conceptual causal relationships do not map directly onto observable variables, whether these are measured in the field or are harvested from secondary sources.

*Incomplete early draft, not to be quoted.

†Department of Economics, Norwegian School of Economics and Business Administration, Helleveien 30, N-5045 Bergen, Norway; E-mail: Roger.Bivand@nhh.no

Further, these observations may very well not correspond to the support of the conceptual spatial processes, and as data is often only available in a single format of entities, researchers have little control over this source of misspecification.

Because the ensembles of concepts used to handle variables and units of observation appear to differ across knowledge domains, it is often the case that analysis is carried out in what may be termed relative ontological isolation. Different academic subdisciplines attach varying importance to the conceptual and operational definitions of both variables and observational units. The differences could informally be seen as different prior beliefs and different weightings on thos beliefs about the purpose of the undertaking, assumed to be to try to reach some generalisable statement(s) about some perceived "reality".

In many cases, disciplines differ in the use of primary and secondary data, and in the ways in which any sample design is carried out. In order to access data on response and predictor variables, it is frequently necessary to aggregate, or to change support in one way or another. It may well be that these steps alter the mapping between the observational units in the conceptual model and those actually employed in the chosen operationalisation(s). One reflection of these challenges is the modifiable areal unit problem. The underlying question concerns attempts to make inferences about data generation processes on the basis of patterning in data from cross-sectional and/or temporal observations in entities that may not be well-suited to the process.

It becomes rather problematic if the analysis in hand suggests that spatial regimes, or other forms of non-stationarity in the relationships between the response and predictor variables are immanent, that is intrinsic to some subsets of observations, rather than being an expression of misspecification. The latter is entirely acceptable, because often it is not possible to measure all the variables of interests, so that extracting structure from the residuals may be the only way of improving fitted models. If, on the other hand, spatial heterogeneity is present as a model misspecification, it would seem more appropriate to attempt to introduce alternative model specifications and/or variables or modified functional forms. In this way the misspecification would be alleviated, and hopefully the results of the analysis rendered more general and parsimonious.

This review will attempt to discuss a number of topics in the contexts of macroecology and regional growth. First, we will look at issues of entities (observational units), data collection, and spatial scale. From this we go on to examine proposals for controlling for spatial structure in ecology, and conclude by looking at spatial regimes, non-stationarity and club convergence.

2 Entities, data collection and scale

Wilson (2000, p. 14) distinguishes three dimensions which interact in urban and regional analysis: system articulation, theory, and method. System articulation is in turn made up of three sub-dimensions, entitation, levels of resolution (sectoral, spatial, temporal), and spatial representation. Entitation is stressed by Chapman (1977) in his treatment of human and environmental systems. As Wilson (2000) comments, all too little attention is paid in analysis to careful planning of the main dimensions, with system articulation typically treated in the least satisfactory way. His second and third chapters provide a succinct and enlightening review of why system articulation matters (Wilson, 2000, pp. 6–19).

Interestingly, entitation is also an important concept in ecology: "Entitation — the processes by which entities are recognized and defined, usually by dividing a continuously varying phenomenon into a set of discrete entities. In vegetation ecology entitation refers to the act of segmenting an area of vegetation into homogeneous entities, within which samples (plots) can be placed, or the division of community data (usually plot data) into discrete vegetation classes" (Jennings et al., 2009). Note that the entities are intended to be homogeneous, rather than arbitrary.

Questions of representation and entitation have exercised students of geographical information in important ways over the past two decades. "Early on, the complexities of object definition were confined to the debate over the relative merits of data structures and models. Physical structuring devices were regarded as the primary influences on entitation. Attention to cognitive and philosophical implications of ecoding features using different data models in the early 1990s led to the examination of epistomologies and ontologies" (Schuurman, 2005, p. 27). Discussions in GIScience have not only given attention to the intentionality involved in choosing representations of observational units and variables. They have pointed up the contrasts and conflicts that can arise in the representations given by different groups, which may vary in discipline or in degree of influence in, for example, determining policy outcomes.

Haggett et al. (1977, pp. 260–290) present many aspects of spatial data collection. They draw attention to the fact that in human geography the vast majority of observations are taken from secondary, archival sources, very often of a non-areal nature. They note that: "This dependence on secondary sources has three main consequences: (i) locational analysis is using data which have been collected primarily for non-geographical purposes, and these data are usually oblique in varying degrees to the direct research needs of geographers; ... (iii) data are released in 'bundles' (i.e. for administrative areas) which are inconvenient and anachronistic, and pose extremely acute problems in mapping and interpretation" (Haggett et al., 1977, p. 261–262).

Despite the time that has passed since these concerns were voiced, the availabil-

ity of data better suited to the needs of analysis has not changed. Even access to register data at the individual level, and to other high volume digital sources has not modified the constraint that much of the data is not collected in order to conduct the analysis at hand. Spatial sampling is possible, but most often with regard to primary data. Techniques are discussed by Ripley (1981, pp. 19–27), and the field reviewed by Müller (2007); an extension to areal data is given by Gumprecht et al. (2009).

They comment that: "Besides the nature of the investigated process, the locations in space, where the data are collected, the so-called spatial sampling design, have the most influence on our ability to isolate spatial effects. ... However, it seems a little like negligence that this has never been considered for the very first phase of a spatial study. One explanation for this is that very frequently the sampling design is fixed beforehand. ... Usually these data come at no or little cost ..., and thus there is no need for posing the question: where (to measure the data)?" (Gumprecht et al., 2009, pp. 127–128).

Both the spatial level of resolution and the mode of spatial representation are involved in spatial scale. Scale is intimately connected to the pattern/process matching that is central to analysis, because certain causal effects may be present only at particular scales. If the spatial representation (driven by available data) misses this scale, real causal effects will be obscured. Usually, it is the micro-scale variation of a scale smaller than the observation units (or the distances between them) that is omitted. However, omitted large scale trends may be seen as autocorrelation rather than misspecification (Schabenberger and Gotway, 2005, p. 22).

3 Controlling for spatial structure in ecology

The red herring debate in geographical ecology was ignited by Lennon (2000), who claimed that substantive conclusions about the impact of environmental factors on for example species richness had been undermined by not taking spatial autocorrelation into account. Diniz-Filho et al. (2003) replied, challenging not only the interpretation of the problem in statistical terms, but pointing out that geographical ecology also involves the scale problem, that the influence of environmental factors is moderated by spatial scale.

They followed this up in a study in which the data were subsampled to attempt to isolate the scale problem. But they begin: "It is important to note that we do not present a formal evaluation of this issue using statistical theory ..., our goal is to illustrate heuristically that the often presumed bias due to spatial autocorrelation in OLS regression does not apply to real data sets" (Hawkins et al., 2007, p. 376).

The debate continues with verve in Beale et al. (2007) and Diniz-Filho et al. (2007). This is quite natural, as doubts about the impacts of environmental drivers on species richness raise questions about, for example, the effects of climate change.

How to analyse spatial data is obviously of importance within geographical ecology. However, Diniz-Filho et al. (2007, p. 850) conclude that: "[w]hen multiple assumptions are not being met, as in the case of virtually all geographical analyses, can a result from any single method (whether spatial or non-spatial) be claimed to be better? ... If different spatial methods themselves are unstable and generate conflicting results in real data, it makes no sense to claim that any particular method is always superior to any other".

Multi-author papers have established themselves in ecology as a way of conducting surveys extending beyond the capacity of a smaller group of researchers. Dormann et al. (2007) discuss the practical estimation of spatial autocorrelation, while Bini et al. (2009) examine in depth the actual impact of including or omitting spatial autocorrelation in regression models. They follow up the read herring debate, and conclude: "In sum, we find little coherence or order in the extent and pattern of coefficient shifts in empirical spatially structured data. Given the idiosyncratic patterns of coefficient shifts, we may have little recourse, but to conduct much more complete evaluations of patterns of covariation among variables and report the true levels of uncertainty in our models. This will come as troubling news to some macroecologists, but we feel that, in the end, confusion over understanding the underlying causes of broad-scale patterns in the distribution and abundance of species will be reduced." (Bini et al., 2009, pp. 10–11).

In an interesting intervention, McIntire and Fajardo (2009) suggest the role of space as a surrogate for unobserved processes. They "demonstrate the rapidly emerging use of spatial patterns to help elucidate the underlying ecological processes and formalize this procedure with the title 'space as a surrogate' for unmeasured or unmeasurable processes. This is not just a statistical technique per se, rather, it is a merging of three components that must act together: precise implementation of ecological theory and/or knowledge, a priori inference, and precise application of spatial analytical tools" (McIntire and Fajardo, 2009, p. 47). Warren et al. (2009) have undertaken a survey of selected published studies to try to establish which features of the considered cases lead to difficulties when spatial scale and residual spatial autocorrelation are both present. de Knegt et al. (preprint) have underaken a similar exercise using a virtual dataset.

One promising and flexible approach is based on the eigenfunctions of a spatial weights matrix, as introduced by Griffith (1996). Of course, scale effects that are smaller than the resolution of the entities whose mutual relations are expressed by the weights matrix cannot be captured, constituting micro-scale variation. Spatial filtering using these kinds of Moran eigenvectors has been followed up by Griffith (2003), and further by Tiefelsdorf and Griffith (2007). In addition, this approach is contrasted with a variable-by-variable filtering technique due to Getis in Getis and Griffith (2002). The eigenvectors typically extract patterns of differing scale

from the spatial weights matrix, and maps of the eigenvectors are known as Moran eigenvector maps (MEM).

Borcard and Legendre (2002) appear to have reached a similar eigenvector approach independently, but later work by Legendre et al. (2005), Griffith and Peres-Neto (2006), and Dray et al. (2006) shows how closely the principal coordinates of neighbour matrices (PCNM) approach aligns with spatial filtering using eigenfunctions of a spatial weights matrix. Speculations about the possible use of eigenvectors of spatial weights matrices have occurred in the literature from many authors over many years, for example Bivand (1981, p. 65, in Polish).

These techniques have been followed up in ecology by Legendre et al. (2008), Peres-Neto and Legendre (2010), and Jombart et al. (2009), with applications by Dray et al. (2010), and Qian and Kissling (2010). While overlapping authorships between these papers are clear, the seriousness of the problem of spatial scale being conflated with residual spatial autocorrelation necessarily attracts much attention. In the study of regional growth, studies have been conducted by Fischer and Griffith (2008) using eigenvector spatial filtering in a spatial interaction setting, and by Fischer and Stumpner (2008) using Getis variable-by-variable filtering.

4 Spatial regimes, non-stationarity and club convergence

5 Conclusions

References

Beale CM, Lennon JJ, Elston DA, Brewer MJ, Yearsley JM (2007) Red herrings remain in geographical ecology: a reply to Hawkins et al. (2007). Ecography 30:845–847

Bini LM, Diniz-Filho JAF, Rangel TFLVB, Akre TSB, Albaladejo RG, Albuquerque FS, Aparicio A, Araujo MB, Baselga A, Beck J, Isabel Bellocq M, Boehning-Gaese K, Borges PAV, Castro-Parga I, Chey VK, Chown SL, de Marco P Jr, Dobkin DS, Ferrer-Castan D, Field R, Filloy J, Fleishman E, Gomez JF, Hortal J, Iverson JB, Kerr JT, Kissling WD, Kitching IJ, Leon-Cortes JL, Lobo JM, Montoya D, Morales-Castilla I, Moreno JC, Oberdorff T, Olalla-Tarraga MA, Pausas JG, Qian H, Rahbek C, Rodriguez MA, Rueda M, Ruggiero A, Sackmann P, Sanders NJ, Terribile LC, Vetaas OR, Hawkins BA (2009) Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. ECOGRAPHY 32(2):193–204

Bivand RS (1981) Modelowanie geograficznych układów czasoprzestrzennych. Państwowe Wydawnictwo Naukowe, Poznań

Borcard D, Legendre P (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. ECOLOGICAL MODELLING 153(1-2):51–68

Chapman GT (1977) Human and environmental systems. Academic Press, London

Diniz-Filho JA, Bini LM, Hawkins BA (2003) Spatial autocorrelation and red herrings in geographical ecology. Global Ecology & Biogeography 12:53–64

Diniz-Filho JA, Hawkins BA, Bini LM, De Marco Jr P, Blackburn TM (2007) Are spatial regression methods a panacea or a Pandora's box? a reply to Beale et al. (2007). Ecography 30:848–851

Dormann CF, McPherson JM, Araujo MB, Bivand R, Bolliger J, Carl G, Davies RG, Hirzel A, Jetz W, Kissling WD, Kuehn I, Ohlemueller R, Peres-Neto PR, Reineking B, Schroeder B, Schurr FM, Wilson R (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. ECOGRAPHY 30(5):609–628

Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). ECOLOGICAL MODELLING 196(3-4):483–493

Dray S, Royer-Carenzi M, Calenge C (2010) The exploratory analysis of autocorrelation in animal-movement studies. ECOLOGICAL RESEARCH 25(3):673–681

Fischer MM, Griffith DA (2008) MODELING SPATIAL AUTOCORRELATION IN SPATIAL INTERACTION DATA: AN APPLICATION TO PATENT CITATION DATA IN THE EUROPEAN UNION. JOURNAL OF REGIONAL SCIENCE 48(5):969–989

Fischer MM, Stumpner P (2008) Income distribution dynamics and cross-region convergence in Europe - Spatial filtering and novel stochastic kernel representations. JOURNAL OF GEOGRAPHICAL SYSTEMS 10(2):109–139

Getis A, Griffith D (2002) Comparative spatial filtering in regression analysis. GEOGRAPHICAL ANALYSIS 34(2):130–140

Griffith D (1996) Spatial autocorrelation and eigenfunctions of the geographic weights matrix accompanying geo-referenced data. CANADIAN GEOGRAPHER-GEOGRAPHE CANADIEN 40(4):351–367

Griffith D (2003) Spatial autocorrelation and spatial filtering. Springer, Berlin

Griffith DA, Peres-Neto PR (2006) Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. Ecology 87:2603–2613

Gumprecht D, Müller WG, Rodríguez-Díaz JM (2009) Designs for detecting spatial dependence. Geographical Analysis 41(2):127–143

Haggett P, Cliff AD, Frey A (1977) Locational Analysis in Human Geography. Edward Arnold, London

Hawkins BA, Diniz-Filho JA, Bini LM, De Marco Jr P, Blackburn TM (2007) Red herrings revisited: spatial autocorrelation and parameter estimation in geographical ecology. Ecography 30:375–384

Jennings MD, Faber-Langendoen D, Loucks OL, Peet RK, , Roberts D (2009) Standards for associations and alliances of the U.S. National Vegetation Classification. Ecological Monographs 79:173–199

Jombart T, Dray S, Dufour AB (2009) Finding essential scales of spatial variation in ecological data: a multivariate approach. ECOGRAPHY 32(1):161–168

de Knegt HJ, van Langevelde1 F, Coughenour MB, Skidmore AK, de Boer1 WF, Heitkönig IM, Knox NM, Slotow R, van der Waal1 C, Prins HH (preprint) Spatial autocorrelation and the scaling of species-environment relationships. Ecology

Legendre P, Borcard D, Peres-Neto PR (2005) Analyzing beta diversity: Partitioning the spatial variation of community composition data. Ecol Monogr 75(4):435–450

Legendre P, Borcard D, Peres-Neto PR (2008) Analyzing or explaining beta diversity? comment. Ecology 89(11):3238–3244

Lennon JJ (2000) Red-shifts and red herrings in geographical ecology. Ecography 23:101–113

McIntire EJB, Fajardo A (2009) Beyond description: the active and effective way to infer processes from spatial patterns. Ecology 90(1):46–56

Müller WG (2007) Collecting spatial data. Springer, Berlin

Peres-Neto PR, Legendre P (2010) Estimating and controlling for spatial structure in the study of ecological communities. GLOBAL ECOLOGY AND BIOGEOGRAPHY 19(2):174–184

Qian H, Kissling WD (2010) Spatial scale and cross-taxon congruence of terrestrial vertebrate and vascular plant species richness in China. ECOLOGY 91(4):1172–1183

Ripley BD (1981) Spatial statistics. Wiley, Hoboken, New Jersey

Schabenberger O, Gotway CA (2005) Statistical methods for spatial data analysis. Chapman & Hall/CRC, Boca Raton, Florida

Schuurman N (2005) Social dimensions of object definition in gis. In: Fisher P, Unwin D (eds) Re-presenting GIS, John Wiley, Chichester, pp 27–41

Tiefelsdorf M, Griffith DA (2007) Semiparametric filtering of spatial autocorrelation: The eigenvector approach. Environment and Planning A 39:1193–1221

Warren M, Mcgeoch MA, Chown SL (2009) The detection of spatial structure in populations and communities: An empirical case study. ECOSCIENCE 16(1):95–110

Wilson AG (2000) Complex spatial systems: the modelling foundations of urban and regional analysis. Pearson Education, Harlow