Spatial Statistics:
Putting the Myths into Perspective

Marie-Josée Fortin and Mark R.T. Dale

[1] Department of Biological Sciences, University of Alberta, Edmonton (Alberta), T6G 2E9. email: mark.dale@ualberta.ca
[2] Department of Ecology and Evolutionary Biology, University of Toronto, Toronto (Ontario), M5S 3G5. email: mjfortin@zoo.utoronto.ca

Abstract

A major aim of including the spatial component in ecological studies is to characterize the nature and intensity of spatial relationships between organisms and their environment. The growing awareness by ecologists of the importance of including spatial structure in ecological studies (for hypothesis development, experimental design, statistical analyses and spatial modeling) is beneficial because it promotes more effective research. There is a drawback however: as more researchers use and include spatial notions and spatial statistics in their analysis some misconceptions about the virtues of spatial statistics that have been carried through the process and years. Here, we review the most common misconceptions about spatial autocorrelation as a list of myths and challenges. We synthesise the problems related to incorporating spatial considerations correctly in the analysis of spatially heterogeneous ecological systems. We conclude in proposing approaches to solutions to those problems.

Background and Relevance

Nowadays, it is increasingly common for ecological studies to acknowledge the importance of the spatial aspects of the systems under study (Dungan et al. 2002, Fortin and Dale 2005, Wagner and Fortin 2005), and to include them as well as possible in the study's design (Legendre et al. 2002) analysis (Legendre et al. 2004) and modeling (Keitt et al. 2002, Lichstein et al. 2002) among others. Spatial effects, however, are many and take various possible interdependent forms including scale effects, spatial autocorrelation, locational and neighbour effects, and the general interaction between spatial pattern and temporal process. The topic is both complex and potentially confusing, and while the increasing awareness of the importance of spatial issues is beneficial, its complex nature has given rise to a certain amount of misunderstanding of the underlying concepts, as well as a number of misconceptions, what could be called "myths", about related issues and their potential solutions. Indeed, spatial statistics are often thought to be the panacea that will solve all the problems of having or not spatially structured data. Although there is help available in the literature and in a number of textbooks (Cressie 1993, Haining 2003, Fortin and Dale 2005), it may require some restating or interpretation, and even with such assistance, challenges will undoubtedly remain. It is our intention to review some of what we see as prevalent myths about spatial effects, with an attempt to correct the related misunderstandings.

Myths, Challenges and Solutions

The most common myth is that there is a *Distance to Independence*, that is if your samples are far enough apart, there is no problem; they are good as independent. This comes from the assumption is that autocorrelation declines with distance (rapidly) to zero and remains approximately zero thereafter, so that distant samples are effectively independent. The problems are, however: (1) the spatial structure is still present in the system investigated regardless of the spaced between the samples; and (2) it is also the *Myth of Insignificance*, that is if autocorrelation at distance x is not significantly different from zero, it can be ignored. Indeed the detection of "significance" is dependent on the numbers of observations at the lag distances investigated; larger distances tend to have fewer observations and therefore are less likely to be detected as significant. Therefore, there may be bias against detecting significant autocorrelation at larger scales, which will depend on the characteristics of the sampling design and the study area. The challenge is the *Lack of independence*, is that autocorrelation is not the only source of lack of independence encountered in spatial studies in ecology: dependence also arises when spatial coefficients are computed at more then one distance lag and because the same data are used again and again in calculations; for example, the data used to calculate autocorrelation at lag 2 are used again in different combinations to calculate autocorrelation at lag 6.

No single solution solves all problems. Solutions are context-, scale-, and system-dependent. As with the choice of spatial methods, some care and judgment will be needed in finding a set of solutions to a multifaceted or complex problem with several possible confounded effects. One note of hope is that while the effects of autocorrelation on univariate tests have currently no obvious solution, there seems to be a general pattern for solutions for bivariate tests. In fact, the more variables that are considered, the less important this problem may become.

Conclusions

In spite of the difficulties of this topic, and despite the myths and misunderstandings, this is an area in which real progress has been made in the last decade, both in a greater general awareness and understanding on the part of ecologists and scientists in related disciplines, and in the technical sophistication of finding solutions to the inherent problems. The continued emphasis in ecology on the importance of the relationship between pattern and process and of the spatial context of ecological systems (configuration, not just composition) ensures that this topic will continue to be a focus in the years to come.

References

Cressie, N. A. C. (1993). *Statistics for Spatial Data*. Second revised edition. New York: John Wiley and Sons, Inc.

Dungan, J. L., Perry, J. N., Dale, M. R. T., Legendre, P., Citron-Pousty, S., Fortin, M.-J., Jakomulska, A., Miriti, M. & Rosenberg, M. S. (2002). A balanced view of scale in spatial statistical analysis. *Ecography*, 25, 626-640.

Fortin, M.-J. & Dale, M. R. T. (2005). *Spatial Analysis: A Guide for Ecologists*. Cambridge University Press, Cambridge.
Haining, R. (2003) *Spatial Data Analysis*: *Theory and Practice*. Cambridge: Cambridge University Press.
Keitt, T. H., Bjornstad, O. N., Dixon, P. M. & Citron-Pousty, S. (2002). Accounting for spatial pattern when modeling organism-environment interactions. *Ecography*, 25, 616-625.
Legendre, P., Dale, M. R. T., Fortin, M.-J., Gurevitch, J., Hohn M. & Myers, D. E. (2002). The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography*, 25, 601-615.
Legendre, P., Dale, M. R. T., Fortin, M.-J., Casgrain, P. & Gurevitch, J. (2004). Effects of spatial structures on the results of field experiments. *Ecology*, 85, 3202-3214.
Lichstein, J. W., Simons, T. R., Shriner, S. A. and Franzreb, K.E. (2002). Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs*, 72, 445-463.
Wagner, H. H. & Fortin, M.-J. (2005). Spatial analysis of landscapes: concepts and statistics. *Ecology*, 86, 1975-1987.