

Spatial non-stationarity and the scale of species-environment relationships in the Mojave Desert, California, USA

Jennifer A. Miller¹

¹Department of Geography and the Environment, The University of Texas at Austin, 1 University Station A3100, Austin, Texas, 78712
Telephone: (01-512-232-1587)
Email: Jennifer.miller@austin.utexas.edu

1. Introduction

Species distribution models (SDM) are used to quantify the relationship between some measure of species' importance and environmental factors. They have become a fertile area of research interest at the confluence of spatial ecology and GIScience and have been used to study a wide range of biogeographical phenomena, ranging from invasive species, vector-borne diseases, and biodiversity. One issue that has recently been identified as important is spatial non-stationarity, which addresses how the species-environment relationships vary across space.

Spatial non-stationarity describes modelled relationships that are not constant across space. In other words, the process which underlies the relationship between variables changes with spatial context. This has been addressed in previous work by geographically partitioning the study area before formulating the models (Osborne and Suárez-Seoane 2002). Assunção (2003) outlines approaches for area data for calculating coefficients that vary as smooth functions of geographic locations.

Non-stationarity causes predictor variables to have an impact at different spatial scales, which has important effects on model outcomes. Fitzgerald and Lees (1995) tested the effects of spatial context based on grid cell neighborhoods to map land cover and found that best neighborhood size varied with land cover type and complexity. In studies aimed at predicting biodiversity (and loss thereof), model predictive power was scale-dependent (Rahbek and Graves 2001; Willis and Whittaker 2002).

The relationship between spatial scale and non-stationarity can be examined to identify the spatial scale at which stationarity in the relationship exists, if at all. This research explores the scale of species-environment relationships by using a relatively uncommon (in SDM) statistical method, geographically weighted regression (GWR).

GWR fits potentially different coefficient values for each observation as a function of a spatial kernel weighting scheme, the most important characteristic of which is its bandwidth (the distance within which other areas will have influence; see Fotheringham et al., 2002). The bandwidth can be set at a fixed distance or allowed to shrink and expand ('adaptive kernel') to include a fixed number of observations.

The geographic variation in species-environment relationships and the scale-dependence of SDM predictive ability are emerging issues in Ecology (Rahbek and Graves 2001; Willis and Whittaker 2002; McAlpine et al 2008). Of the few previous GWR applications in species distribution modeling, most have focused on comparing GWR model predictions with global regression model predictions (Kupfer and Farris, 2007) or exploring the spatial variation of model coefficients (Foody 2004; Bickford and Laffan, 2006). These studies typically use either an adaptive kernel or a single optimized bandwidth, and use these results to report the degree of and pattern of spatial non-stationarity. However an underexplored application for which GWR is uniquely suited involves systematically testing a range of fixed bandwidth sizes to determine if there is a spatial scale at which the influence of an environmental variable on species distribution becomes stationary.

This work focuses on several aspects of spatial non-stationarity and scale of species-environment relationships that have not been previously investigated. Although Osborne et al (2007) explored the idea of using geographically weighted regression to determine a scale at which a species-environment relationship becomes stationary, their use of a single species does not allow any comparisons among different types of distributions. The five environmental predictors they used (principal components and distance variables) were also not conducive to examining the scale at which different types of environmental gradients affect species distributions.

2. Data and Methods

The plant data, environmental layers and ~50,000 km² portion of the Mojave Desert used as the study area here (figure 1), are all described in more detail in Miller and Franklin (2006). The plant data consisted of 3819 observations of presence/absence for four species considered to be representative of the different distributions characteristic of the Mojave Desert. *Ephedra nevadensis* (EPNE) is the rarest in the sample (<1%) and occurs in floodplains and washes with well-drained soils; *Larrea tridentata*-*Ambrosia dumosa* (LATRAMDU) is the most widespread alliance in the Mojave Desert (43% of sample) and is common throughout lower elevation areas, particularly alluvial plains and bajadas; *Pinus monophylla* (PIMO) is also rare in the sample (1%), but has strong environmental preferences of cool, moist mountain areas; *Yucca brevifolia* (YUBR), the indicator species of the Mojave Desert (Joshua Tree), occurred in 10% of the sample and is found in a fairly narrow zone at the base of mountains and gentle alluvial fans.

The ten predictor variables represent important environmental gradients comprising climate summaries and extremes and simple and complex topographic variables that describe various soil characteristics and/or modify the effects of the climate factors.

To ensure that any spatial non-stationarity measured here was not a function of model mis-specification, generalized additive models (GAMs) were fit to each of the 40 species-environment pairs and if the resulting shape was significantly nonlinear, a more appropriate transformation of the environmental variables was also tested in the GWR in addition to the linear (L) term.

In addition to a single global regression model for each of the ten species-environment pairs, I used GWR with a range of fixed bandwidths, from ten km to 200 km, with ten km intervals. Model fit and accuracy were of interest here only as a framework for identifying how they are affected by spatial scale and non-stationarity. Akaike information criterion (Hastie et al. 2001) and pseudo- R^2 (Menard 1995) were used to assess model fit.

Jetz et al. (2005) note that R^2 will be inflated at small bandwidths and suggest that a benchmark R^2 be calculated at each bandwidth based on a large number of random predictor variables. The mean R^2 for each bandwidth then represents a null expectation from which the proportional deviation represented by a model with an environmental predictor can be calculated. Here we generated 50 random variables and fit new univariate models to each of the four species at all tested bandwidths (10 km interval). An average was estimated as the null expectation for each species-bandwidth combination, and the change in this value was examined for the species-environment models across all bandwidths.

3. Conclusions

The results suggest that the scale of species-environment relationships varies for both different types of species and different types of environmental variables. Climate and elevation variables generally result in stationary relationships at broader scales, and sometimes do not become stationary at all. Topographic variables operate at finer scales and are often stationary across all scales or become stationary at smaller bandwidths. In addition to being instrumental for examining the effects of scale on spatial non-stationarity for species-environment relationships, GWR also provided important insights on how scale affects a model's explanatory power.

4. Acknowledgements

Acknowledgement of support is gratefully made to the National Science Foundation (grant #0832367) and The University of Texas at Austin College of Liberal Arts. Suggestions by R.Q. Hanham aided in the research design.

5. References and Citations

- ASSUNÇÃO, R. 2003. Space varying coefficient models for small area data. *Environmetrics*, 14, 453-473.
- BICKFORD, S. & LAFFAN, S. 2006. Multi-extent analysis of the relationship between pteridophyte species richness and climate. *Global Ecology & Biogeography*, 15, 588-601.
- FOODY, G. M. 2004. Spatial nonstationarity and scale-dependency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna. *Global Ecology and Biogeography*, 13, 315-320.
- FOTHERINGHAM, A. S., BRUNSDON, C. & CHARLTON, M. 2002. *Geographically Weighted*

- Regression: The analysis of spatially varying relationships*, Chichester, Wiley & Sons.
- HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. 2001. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, New York, Springer-Verlag.
- JETZ, W., RAHBEK, C. & LICHSTEIN, J. W. 2005. Local and global approaches to spatial data analysis in ecology. *Global Ecology and Biogeography*, 14, 97-98.
- KUPFER, J. & FARRIS, C. 2007. Incorporating spatial non-stationarity of regression coefficients into predictive vegetation models. *Landscape Ecology*, 22, 837-852.
- MCALPINE, C., RHODES, R., BOWEN, M., LUNNEY, D., CALLAGHAN, J., MITCHELL, D. & POSSINGHAM, H. 2008. Can multiscale models of species' distribution be generalized from region to region? A case study of the koala. *Journal of Applied Ecology*, 45, 558-567.
- MENARD, S. 1995. *Applied Logistic Regression Analysis*, Iowa City, Sage Publications.
- MILLER, J. A. & FRANKLIN, J. 2006. Explicitly incorporating spatial dependence in predictive vegetation models in the form of explanatory variables: A Mojave Desert case study. *Journal of Geographical Systems*, 8, 411-435.
- OSBORNE, P., FOODY, G. & SUÁREZ-SEOANE, S. 2007. Non-stationarity and local approaches to modelling the distribution of wildlife. *Diversity and Distributions*, 13, 313-323.
- OSBORNE, P. & SUÁREZ-SEOANE, S. 2002. Should data be partitioned spatially before building large scale distribution models? *Ecological Modelling*, 157, 249-259.
- RAHBEK, C. & GRAVES, G. 2001. Multiscale assessment of patterns of avian species richness. *Proceedings of the National Academy of Sciences*, 98, 4534-4539.
- SOBERÓN, J. 2007. Grinnellian and Eltonian niches and geographic distributions of specie. *Ecology Letters*, 10, 1115-1123.
- WILLIS, K. & WHITTAKER, R. 2002. Species diversity- scale matters. *Science*, 295, 1245-1248.