Review & synthesis

Ecography 32: 193–204, 2009
doi: 10.1111/j.1600-0587.2009.05717.x

Subject Editor: Jens-Christian Svenning. Accepted 8 January 2009

Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression

L. Mauricio Bini, J. Alexandre F. Diniz-Filho, Thiago F. L. V. B. Rangel, Thomas S. B. Akre, Rafael G. Albaladejo, Fabio S. Albuquerque, Abelardo Aparicio, Miguel B. Araújo, Andrés Baselga, Jan Beck, M. Isabel Bellocq, Katrin Böhning-Gaese, Paulo A. V. Borges, Isabel Castro-Parga, Vun Khen Chey, Steven L. Chown, Paulo de Marco, Jr, David S. Dobkin, Dolores Ferrer-Castán, Richard Field, Julieta Filloy, Erica Fleishman, Jose F. Gómez, Joaquín Hortal, John B. Iverson, Jeremy T. Kerr, W. Daniel Kissling, Ian J. Kitching, Jorge L. León-Cortés, Jorge M. Lobo, Daniel Montoya, Ignacio Morales-Castilla, Juan C. Moreno, Thierry Oberdorff, Miguel Á. Olalla-Tárraga, Juli G. Pausas, Hong Qian, Carsten Rahbek, Miguel Á. Rodríguez, Marta Rueda, Adriana Ruggiero, Paula Sackmann, Nathan J. Sanders, Levi Carina Terribile, Ole R. Vetaas and Bradford A. Hawkins

L. M. Bini, J. A. F. Diniz-Filho, P. de Marco, Jr and L. C. Terribile, Depto de Biologia Geral, ICB, Univ. Federal de Goiás, CP 131, 74.001–970 Goiânia, GO, Brazil. – T. F. L. V. B. Rangel, Dept of Ecology and Evolutionary Biology, Univ. of Connecticut, Storrs, CT 06269, USA. – T. S. B. Akre, Dept of Biological and Environmental Sciences, Longwood Univ., Farmville, VA 23909, USA. – R. G. Albaladejo and A. Aparicio, Depto de Biología Vegetal y Ecología, Univ. de Sevilla, c/Prof. García Gonzalez no 2, ES-41012 Sevilla, Spain. – F. S. Albuquerque, D. Montoya, I. Morales-Castilla, M. Á. Olalla-Tarraga, M. Á. Rodríguez and M. Rueda, Depto de Ecología, Univ. de Alcalá, ES-28871 Alcalá de Henares, Madrid, Spain. – M. B. Araújo, A. Baselga, J. F. Gómez and J. M. Lobo, Depto de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), ES-28006 Madrid, Spain. – J. Beck, Dept of Environmental Sciences, Inst. of Biogeography, Univ. of Basel, St.Johanns-Vorstadt 10, CH-4056 Basel, Switzerland. – M. I. Bellocq and J. Filloy, Depto de Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Univ. de Buenos Aires, and CONICET, Ciudad Universitaria Pab. 2 (1428) Buenos Aires, Argentina. – K. Böhning-Gaese and W. D. Kissling, Inst. für Zoologie, Johannes Gutenberg-Univ. Mainz, Becherweg 13, DE-55099 Mainz, Germany. – P. A. V. Borges, Depto de Ciências Agrárias, Univ. dos Açores, CITA A (Azorean Biodiversity Group), Terra Chã, PT-9700–851 Angra do Heroísmo, Terceira, Açores, Portugal. – I. Castro-Parga, Depto de Ecología, C/Darwin 2, Univ. Autónoma de Madrid, ES-28049 Madrid, Spain. – V. K. Chey, Entomology Section, Forest Research Centre of Sabah, Sepilok, P.O. Box 1407, 90715 Sandakan, Malaysia. – S. L. Chown, DST-NRF Centre of Excellence for Invasion Biology, Stellenbosch Univ., Private Bag XI, Matieland 7602, South Africa. – D. S. Dobkin, High Desert Ecological Research Inst., 15 S.W. Colorado Ave., Suite 300, Bend, OR 97702, USA. – D. Ferrer-Castán, Área de Ecología, Facultad de Biología, Univ. de Salamanca, ES-37007 Salamanca, Spain. – R. Field, School of Geography, Univ. of Nottingham, Nottingham NG7 2RD, UK. – E. Fleishman, National Center for Ecological Analysis and Synthesis, 735 State St, Suite 300, Santa Barbara, CA 93101, USA. – J. Hortal, NERC Centre for Population Biology, Imperial College at Silwood Park, Ascot SL5 7PY, UK. – J. B. Iverson, Dept of Biology, Earlham College, Richmond, IN 47374, USA. – J. T. Kerr, Dept of Biology, Univ. of Ottawa, Ottawa, ON, Canada KIN 6N5. – I. J. Kitching, Dept of Entomology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK. – J. L. León-Cortés, Depto de Ecología y Sistemática Terrestre, El Colegio de la Frontera Sur, Carr. Panamericana y Av. Periférico Sur s/n, San Cristóbal de las Casas, Chiapas 29290, México. – J. C. Moreno, Depto de Biología, C/ Darwin 2, Univ. Autonoma de Madrid, ES-28049 Madrid, Spain. – T. Oberdorff, IRD, DMPA, Museum National d'Histoire Naturelle, 43 Rue Cuvier, FR-75005 Paris, France. – J. G. Pausas, Centro de Investigación sobre Desertificación (CIDE, CSIC), Apartado Oficial, ES-46470 Albal, Valencia, Spain. – H. Qian, Research and Collections Center, Illinois State Museum, 1011 East Ash Street, Springfield, IL 62703, USA. – C. Rahbek, Center for Macroecology, Dept of Biology, Univ. of Copenhagen, DK-2100 Copenhagen, Denmark. – A. Ruggiero and P. Sackmann, Laboratorio Ecotono, Centro Regional Universitario Bariloche, Univ. Nacional del Comahue – INIBIOMA-CONICET, Quintral 1250 (8400) Bariloche, Rio Negro, Argentina. – N. J. Sanders, Dept of Ecology and Evolutionary Biology, Univ. of Tennessee, Knoxville, TN 37996, USA. – O. R. Vetaas, UNIFOB Global, Univ. of Bergen, NO-5015 Bergen, Norway. – B. A. Hawkins (bhawkins@uci.edu), Dept of Ecology and Evolutionary Biology, Univ. of California, Irvine, CA 92697, USA.

A major focus of geographical ecology and macroecology is to understand the causes of spatially structured ecological patterns. However, achieving this understanding can be complicated when using multiple regression, because the relative importance of explanatory variables, as measured by regression coefficients, can shift depending on whether spatially explicit or non-spatial modeling is used. However, the extent to which coefficients may shift and why shifts occur are unclear. Here, we analyze the relationship between environmental predictors and the geographical distribution of species richness, body size, range size and abundance in 97 multi-factorial data sets. Our goal was to compare standardized partial regression coefficients of non-spatial ordinary least squares regressions (i.e. models fitted using ordinary least squares without taking autocorrelation into account; "OLS models" hereafter) and eight spatial methods to evaluate the frequency of coefficient shifts and identify characteristics of data that might predict when shifts are likely. We generated three metrics of coefficient shifts and eight characteristics of the data sets as predictors of shifts. Typical of ecological data, spatial autocorrelation in the residuals of OLS models was found in most data sets. The spatial models varied in the extent to which they minimized residual spatial autocorrelation. Patterns of coefficient shifts also varied among methods and datasets, although the magnitudes of shifts tended to be small in all cases. We were unable to identify strong predictors of shifts, including the levels of autocorrelation in either explanatory variables or model residuals. Thus, changes in coefficients between spatial and non-spatial methods depend on the method used and are largely idiosyncratic, making it difficult to predict when or why shifts occur. We conclude that the ecological importance of regression coefficients cannot be evaluated with confidence irrespective of whether spatially explicit modelling is used or not. Researchers may have little choice but to be more explicit about the uncertainty of models and more cautious in their interpretation.

Ecologists widely agree that type I error probabilities are not reliable in the presence of spatial autocorrelation (SAC), which has serious implications for hypothesis testing in geographical ecology (Legendre 1993, Fortin and Payette 2002, Legendre et al. 2002, Diniz-Filho et al. 2003, 2008, Fortin and Dale 2005). However, there is little consensus about how to deal with SAC when analyzing spatial ecological data. The dominant practice is to use multiple regression and rank the standardized partial regression coefficients (hereafter standardized coefficients; Sokal and Rohlf 1995) or associated t-values of coefficients of explanatory variables (Hawkins et al. 2003, Tognelli and Kelt 2004, Kissling et al. 2008) under the assumption that higher coefficients represent stronger "effects" on species richness or any other spatially structured variable (e.g. body size, range size or abundance). The idea is to identify the set of most probable ecological and evolutionary drivers of spatial patterns based on the relative magnitudes of their standardized coefficients. This methodological point of view is also related to recent changes in the philosophy of data analysis, with the gradual replacement of classical null-hypothesis significance testing of models with alternative statistical approaches where the goal is to rank competing models in terms of information content or conditional likelihood (Burnham and Anderson 2002, Hobbs and Hilborn 2006, Hoeting et al. 2006, Araújo and New 2007, Diniz-Filho et al. 2008).

Despite a rapidly developing consensus that null-hypothesis significance testing is inappropriate for most broad-scale ecological data, which sidesteps the problem of SAC and type I error, a second serious issue still divides ecologists: does the presence of SAC in either response or explanatory variables lead to bias in the ecological interpretation of results due to shifts in the rank of standardized coefficients in multiple regression models ("coefficient shifts")?

Lennon (2000) argued that environmental predictors with the highest levels of SAC (i.e. variables with the strongest broad-scale spatial structure, referred to as red-shifted variables) have a biased importance in ordinary least squares (OLS) multiple regression models because their magnitudes are inflated and their errors are underestimated, which he called the red shift problem. He concluded that the perceived importance of broad-scale environmental predictors as drivers of species richness has been overstated in the literature as a consequence of SAC in the data. Based on a spatially-explicit (GLS) regression model for a simulated data set, he also argued that spatial modelling corrects for this problem by giving more weight to variables with less spatial structure (blue-shifted variables).

However, it is important to stress that the statistical problems associated with SAC occur mainly when autocorrelation remains in model residuals rather than in the variables themselves. The underlying ecological reason is the difference between broad-scale spatial structures (e.g. spatial dependence) and spatial autocorrelation (sensu Legendre 1993; see also Kissling and Carl 2008). Environmental and biological processes generate all spatial patterns in nature, whereas in macroecological data spatial structure at short distances may also include elements of pseudoreplication, usually due to the presence of false positives in adjacent cells/points or counting mobile individuals more than once (autocorrelation in a strict sense). Related to this, Diniz-Filho et al. (2003) showed that Lennon's (2000) red shifts appeared even when no SAC was observed in model residuals (since broad scale trends generating autocorrelation in the response variable were already taken into account by fitting environmental data). These authors interpreted the changes in the relative importance of the variables when using spatial methods as a consequence of differences in the spatial scales at which explanatory variables influence species richness (i.e. hierarchical patterns of spatial dependence), rather than an effect of SAC per se.

There is broad agreement that the t-values and their significance levels are affected by SAC since errors of regression coefficients, which are the denominators of t-values, are too narrow due to inflated degrees of freedom in autocorrelated data (i.e. data provide less information than would be obtained from data with uncorrelated regression residuals). But since OLS is unbiased (Cressie 1993, Schabenberg and Gotway 2005), SAC should not affect the underlying structure of the relationships between response and explanatory variables measured by the raw (conventional) or standardized regression coefficients (Bini et al. 2000, Hawkins et al. 2007, Araújo et al. 2008, Kissling et al. 2008). However, additional recent claims about the interpretation of spatial and non-spatial regression coefficients have appeared since Lennon (2000), explicitly arguing that the presence of SAC alters the relative importance of regression coefficients and thus can change the interpretation of broad-scale macroecological analyses (Dormann 2007, Kühn 2007).

To date, it remains unclear if standardized coefficients are sensitive to SAC, and the behavior of regression coefficients in spatially structured data is not well understood. However, coefficient shifts can be large enough to change interpretation of the ecological and evolutionary mechanisms presumed to be driving the observed patterns in some cases (Kühn 2007). Uncertainty about why results vary has also been fueled by the large number of spatial methods that have been devised to model SAC (see below). Another complication is that previous studies have been based on simulated data or single data sets (Dormann 2007), making it difficult to evaluate the general applicability of the results. Thus, we believe that the results provided by simulated data should be complemented with results provided by comprehensive, real-world datasets.

We assembled 97 spatial ecological data sets of highly variable spatial extents and grain sizes that contain a number of ecological response variables to address three questions: 1) how frequent are shifts in the relative importance of the explanatory variables accounting for the variance in the response variable when one moves from a non-spatial to a spatial model, 2) how great are differences in standardized coefficients estimated with non-spatial and spatial regression techniques, and 3) can we predict when coefficient shifts are likely based on the structure of a data set? Previous comparative studies have used simulated data (Dormann et al. 2007) or compilations of results from the literature in which data sets were analyzed with different spatial techniques and assumed different spatial structures (Dormann 2007). Here we included as many data sets as we were able to compile and compared eight spatial regression techniques. Our primary goals were to quantify the frequency of coefficient shifts in real ecological data and to generate a predictive framework for when workers might expect a lack of congruence between a non-spatial and spatial approach. Although we value the power of simulated data sets to test hypotheses related to the performance of statistical methods, ultimately we need to know how methods behave when confronted with real data.

Methods

The data sets

The data sets come from a variety of taxonomic groups, geographic regions, spatial extents, and sample sizes (Supplementary material Table S1). Response variables comprise species richness (74 cases), mean body size (14 cases), mean range size (5 cases), and abundance (4 cases). The number of explanatory variables available for the regression models range from 3 to 20. Sample sizes vary from 22 to 2403, being smallest for island archipelagos (4 cases) and largest for subcontinental or continental regions (46 cases). Plant data sets span non-vascular, vascular, native and non-native species. Animal data include terrestrial and aquatic taxa, invertebrates and vertebrates, and seasonal and year-round residents. We analyzed 97 data sets that were spatially explicit (i.e. in which latitude and longitude of each sampling unit is known) and included a minimum of three predictor variables. Although many additional data sets exist in the literature, our collection was compiled without taxonomic or geographic bias, and we believe they are representative of the data analyzed by macroecologists and geographical ecologists.

The spatial methods

To compare the effects of non-spatial and eight spatial regression models on the estimation of standardized coefficients and to identify characteristics of data that generate coefficient shifts, we used an automated analysis procedure using the statistical library developed for SAM (Rangel et al. 2006). For each data set we initially fitted non-spatial ordinary least squares (OLS hereafter) with all available explanatory variables. We then fitted the data using eight spatial methods: lagged predictor (LagPred), lagged response (LagResp), simultaneous autoregressive (SAR), conditional autoregressive (CAR), moving average (MA), and three variants of eigenvector-based spatial filters (or spatial eigenvector mapping, SEVM) models (see Haining 1990, 2002, Cressie 1993, Schabenberg and Gotway 2005, Rangel et al. 2006, Dormann et al. 2007 for general descriptions of the model types). Lagged-response and lagged-predictor are also forms of simultaneous autoregression, although we retained the term SAR only for the SAR method which incorporates the autocorrelation in the residual covariance (the SAR-error method – see Kissling and Carl 2008 for comparison). Although eigenvector mapping is a unique method that expresses "space" as a set of eigenvectors extracted from a symmetrical matrix expressing spatial relationships among spatial units, we refer to "variants" of SEVM in terms of using different criteria to select eigenvectors to be added to the model (Borcard and Legendre 2002, Borcard et al. 2004, Diniz-Filho and Bini 2005, Griffith and Peres-Neto 2006), and this can have a substantial impact on the regression coefficients associated with environmental variables. The first variant included as spatial predictors all eigenvectors for which Moran's I >0.1 (SEVM-v1). The second included only eigenvectors that were significantly correlated ($p<0.05$) with the response variables (SEVM-v2). In the third variant, we selected all eigenvectors that were significantly correlated with the residuals of OLS models that resulted from regressing each response variable against the environmental predictors (SEVM-v3). The SEVM-v3 is roughly comparable to the criterion adopted by Griffith and Peres-Neto (2006) of selecting eigenvectors that minimize Moran's I in regression residuals.

We did not use generalized least squares (GLS or Kriging regression) because of the need to fit semi-variograms in SAM iteratively, and because all analyses were based on a connectivity matrix (see below) rather than on a matrix of continuous distances.

A Gabriel connection was used to describe the spatial relationship between spatial units (Legendre and Legendre 1998). Gabriel networks approximate the rook scheme when the data are in a regular grid, whereas in irregularly spaced spatial data Gabriel networks can form a more complex link between neighboring spatial units. Using these short-distance connections is preferable (i.e. more conservative; Griffith 1996) than using inverse-decaying distances, since in most empirical data sets residual spatial autocorrelation tends to be stronger at smaller distances classes.

Metrics of coefficient shifts

We created three metrics of the coefficient shifts that result from moving from OLS to spatial modelling, following the reasoning of Dormann (2007). All were calculated for each spatial method independently. First, the relative spatial autocorrelation effect (rSACe) was calculated as described by Dormann (2007): $rSACe = (b_{OLS} - b_{SM})/\text{Maximum}(b_{OLS}, b_{SM})$, where b_{OLS} and b_{SM} are the standardized coefficients estimated by OLS or a spatial method (SM). The rSACe measures the relative change in the coefficients of each explanatory variable, independently of their magnitudes. Second, we calculated a simple difference (diff) between standardized coefficients estimated by OLS and each spatial method ($b_{OLS} - b_{SM}$). In this measure, shifts in variables with relatively high importance (as measured by standardized coefficients) are given greater weight, since their absolute difference will be larger. The third metric was the Spearman correlation between the ranks of the standardized coefficients of the explanatory variables in the OLS model for a given data set and the ranks of the variables as given by the standardized coefficients estimated by each spatial method. A perfect positive correlation indicates that a spatial method did not influence the relative importance (ranks) of the explanatory variables even if the method changed the values of the coefficients. The first two metrics (rSACe and diff) were calculated for each predictor variable and then averaged across each data set (so that metrics for coefficient shifts could be paired with other characteristics of the data set, see below), whereas the Spearman correlations are a data set-level metric and therefore could be directly paired with the characteristics of the data sets.

Correlates of coefficient shifts

We examined eight possible explanations for coefficient shifts:

1) n – the number of observations. A preliminary examination of the data sets found that the number of data points was positively associated with levels of spatial autocorrelation in the response variables (Pearson's correlation between n and Moran's I coefficients $= 0.52$, $p < 0.05$). There was a similar correlation for the predictors (mean correlation within and across all data sets $= 0.62$, $p < 0.05$).

2) m – the number of explanatory variables. Data sets with more explanatory variables are likely to contain more collinearity due to both the relationships among the explanatory variables themselves and the shared variance explained by those variables and spatial coordinates (spatially structured environmental variation sensu Borcard et al. 1992). Higher levels of collinearity increase instability in the estimation of partial regression coefficients (Sokal and Rohlf 1995).

3) VIF – variance inflation factor. The effect of multicollinearity on parameter estimation in OLS is widely known (Graham 2003) and can be assessed by the mean VIF of the explanatory variables. The sensitivity of spatial methods to multicollinearity is not necessarily similar to that found in non-spatial regression (i.e. this sensitivity can vary among spatial methods as well). This explanatory variable was thus used to evaluate the extent to which the changes in the standardized coefficients were influenced by the level of multicollinearity independently of the number of explanatory variables [see variable (2)].

4) Moran's I_{ave} – the average level of spatial autocorrelation in the explanatory variables. According to Lennon (2000), "Correlation between an autocorrelated response variable and each of a set of explanatory variables is strongly biased in favour of those explanatory variables that are highly autocorrelated – the expected magnitude of the correlation coefficient increases with autocorrelation even if the spatial patterns are completely independent" (p. 101, Abstract). Although Lennon (2000) did not use the term coefficient shifts explicitly, following his reasoning, relative spatial autocorrelation effects (rSACe), absolute differences in standardized coefficients (diff), and changes in the relative importance of the predictors (Spearman) would also be higher in the data sets with higher mean levels of autocorrelation (i.e. stronger spatial structure) in the explanatory variables. In all cases, Moran's Is were calculated using the same spatial structure used for modelling, i.e. the Gabriel network among sampling units.

5) R^2 – adjusted coefficient of determination of OLS. Models explaining a low fraction of the total sum of squares can differ more in relation to spatial models and, thus, coefficient shifts may arise more frequently.

6) Residual Moran's I – spatial autocorrelation in the residuals of OLS models. As pointed out by Cressie (1993), spatial structure in model residuals may indicate model misspecification because one (or more) important explanatory variable is missing. This instability may also be linked to Lennon's (2000) red shift because of the spatial component in the explanatory variables. When these unmodelled explanatory variables, which are spatially structured, are added to the model, the level of spatial autocorrelation in the residuals decreases. Conversely, when spatial autocorrelation is controlled by a spatial regression method, the relative importance of explanatory variables without spatial structure may increase.

7) g1 – skewness of the residuals of the OLS models. All of the methods we use assume that residuals are normally distributed. Spatial and non-spatial models may have different levels of robustness against violations in this assumption.

8) F – stationarity, as measured by the F-statistic derived from an ANOVA in which the OLS model is compared against a geographically weighted regression model, calculated using the residual sum of squares for the standard OLS model and that for the GWR model (Fotheringham et al. 2002, pp. 91–92). A spatial data set is considered stationary if the relationship between the variables is constant throughout the entire spatial extent of the data. This is a basic assumption of spatial modelling, although it is rarely

tested by ecologists. Even so, GWR has recently been used to investigate patterns of non-stationarity in macroecological data by examining spatial variation in the regression coefficients (Foody 2004, Wang and Tenhunen 2005, Bickford and Laffan 2006, Cassemiro et al. 2007). Although GWR cannot be directly compared to the other spatial methods because there is a set of regression-estimated parameters for each spatial unit, the F-statistic can be used to measure the gain in fitting a GWR instead of an OLS model and thus to estimate the amount of non-stationarity in the data (Fotheringham et al. 2002). Violation of the stationarity assumption could underlie coefficient shifts because if a predictor is strongly non-stationary, it will generate a larger effect when spatial methods are used (see the results for topography in Hawkins' et al. 2007 analyses). High F-values could also capture effects of broad-scale non-linearity in modelling. For instance, fitting a linear model to a quadratic response would create structure in residuals that could also be expressed as non-stationarity. Although our automatic procedure cannot disentangle these two "sources" of non-stationarity, the main issue is that this statistic can be used to test if these effects can create coefficient shifts.

Analysis

First, we compared the standardized coefficients obtained from all nine methods (OLS and the eight spatial methods) and classified each data set and each spatial method according to its coefficient shifts: 1) no coefficient shifts (the ranks of the explanatory variables were identical), 2) primary shifts (the most important explanatory variable of OLS was not the same as in the spatial model), and 3) secondary shifts (the most important explanatory variable was the same in OLS and the spatial model, but rank changes occurred in other coefficients). Thus, we established the frequency at which either primary or secondary shifts occurred for each spatial method. This is roughly similar to calculating correlations between coefficients, but it provides a qualitative description of the problems caused by coefficient shifts. Then, for each spatial method, a quantitative evaluation of coefficient shifts was performed based on the data sets' characteristics (the eight potential correlates of coefficient shifts). Each metric of coefficient shift for each method was regressed against the eight variables listed in the previous section. Subsequently, multi-model inference based on model averaging was used to estimate the relative importance of each predictor (Burnham and Anderson 2002).

Because the effects of spatial models on coefficient shifts may be correlated among the methods, we also considered it potentially informative to evaluate all methods simultaneously. Canonical correlation analysis (CCorA) was used to investigate the relationships between the metrics of coefficient shifts (rSACe, diff and Spearman) and the eight correlates of coefficient shifts. In these analyses, each data set corresponds to one observation, and one matrix comprises one metric of coefficient shifts for each of the eight spatial model types with a corresponding matrix comprising the eight correlates of coefficient shifts. We performed an independent CCorA for each metric of coefficient shift.

Results

Residual autocorrelation, frequency of shifts, and variation in metrics of coefficient shifts

High, but variable levels of spatial autocorrelation (SAC) were found in the residuals of OLS models across the data sets (Fig. 1). This indicates that, if one assumes that residual SAC is the root of the problem, the potential for coefficient shifts was present in most cases. LagResp, LagPred, SAR and MA were most effective on average for eliminating or minimizing residual SAC. The remaining four methods were less effective in eliminating residual SAC in the majority of the data sets using our standardized method. In other words, under our standardized analytical protocol not all methods controlled for SAC to the same extent, and no spatial method controlled for SAC in all data sets.

Irrespective of the degree to which the spatial methods controlled for SAC, the frequency of coefficient shifts generated by the various methods was moderate to high (Table 1), ranging from 64% of the data sets when using the third variant of spatial eigenvector mapping (SEVM-v3) to 94% in the first variant of spatial eigenvector mapping (SEVM-v1). In general, about half of the coefficient shifts were primary shifts. That is, a clear discrepancy between coefficients generated using OLS and spatial regression occurred between one- and two-thirds of the time, depending on the type of spatial model selected. Thus, all spatial methods generated at least different coefficient shifts

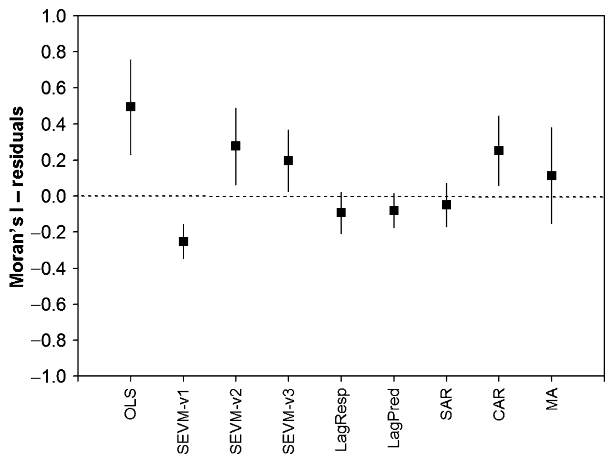

Figure 1. Moran's I of the residuals obtained with different models correcting for spatial autocorrelation across the 97 data sets (using a Gabriel network in all cases). Moran's I of the residuals estimated with ordinary least squares (OLS) is also shown for comparison. Methods are: spatial eigenvector mapping variants 1–3 (SEVM-v1, SEVM-v2, SEVM-v3), lagged response (LagResp), lagged predictors (LagPred), simultaneous autoregressive (SAR), conditional autoregressive (CAR) and moving average (MA).

Table 1. Summary of the sensitivity of standardized regression coefficients to spatial regression methods for 97 data sets, ranked in terms of the number of primary shifts across data sets. "No shifts" represent cases where ranks of coefficients are identical to those obtained using OLS, "secondary shifts" represent cases were the highest coefficients are identical for OLS and each spatial method but other coefficients change ranks, and "primary shift" represent cases where highest coefficients differ for OLS and each spatial method. The spatial methods tested are lagged predictor (LagPred), lagged response (LagResp), simultaneous autoregressive (SAR), conditional autoregressive (CAR), moving average (MA) and three variants of eigenvector mapping (SEVM-v1, SEVM-v2 and SEVM-v3).

Spatial method	No shifts	Secondary shifts	Primary shifts
SEVM-v3	35	35	27
CAR	28	41	28
MA	24	43	30
SEVM-v2	23	39	35
SAR	12	47	38
LagPred	8	31	58
SEVM-v1	6	39	52
LagResp	7	28	62

when compared to OLS, although the probability of a shift occurring, and its severity, varied among methods.

Based on the rSACe metric created by Dormann (2007), LagResp and LagPred generated the strongest shifts in standardized coefficients (Fig. 2a). It is noteworthy that variation in effects within each method was greater than the variation across methods, making it difficult to know a priori how strong a relative SAC effect will be in any particular data set. However, the simple differences between standardized coefficients (diff) indicated that the absolute levels of the shifts were low when comparing spatial and non-spatial methods (Fig. 2b). Thus, many of the differences detected by rSACe were in variables with relatively low explanatory power. We also found that ranked coefficients typically were strongly correlated (Fig. 3), although, as with other metrics, some methods generated stronger shifts than others and there was a great deal of variation among data sets. SEVM-v2, SEVM-v3, CAR and MA in particular generated results generally concordant with OLS (strong positive correlations in most data sets), as also suggested by their rSACe and diff patterns (Fig. 2). SEVM-v1, LagResp, LagPred and SAR, however, were more variable in their consistency with OLS results (Fig. 3).

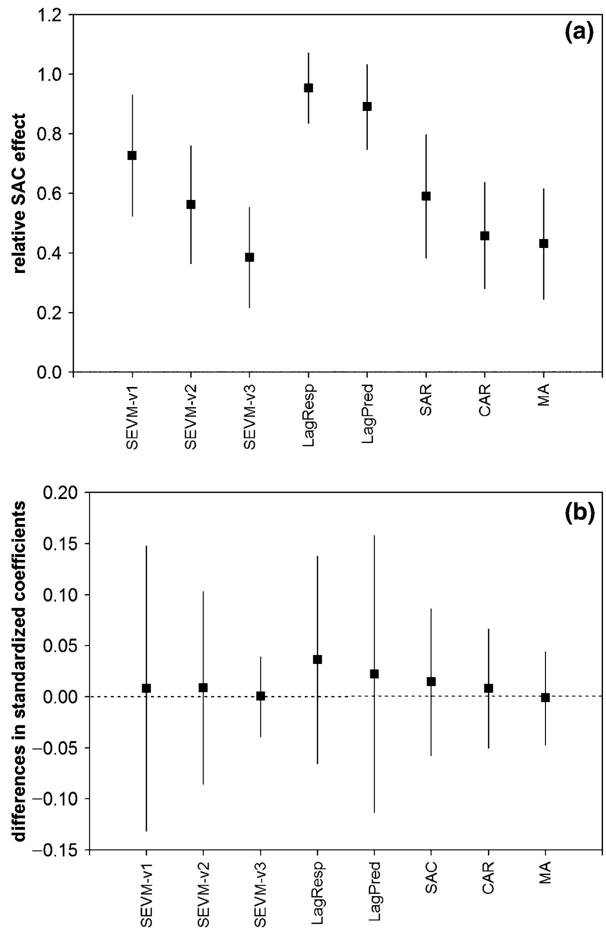

Figure 2. Effects of correcting for spatial autocorrelation on coefficient shifts: (a) "relative SAC effect" (rSACe; see text) and (b) differences between standardized coefficients estimated by OLS and spatial models. For method codes see Fig. 1.

Correlates of coefficient shifts

Akaike weighted, averaged multiple regression models of the three metrics of coefficient shift against the data-set predictors had low coefficients of determination, never >0.31 (even with up to six explanatory variables) and typically <0.20 (Table 2; see also Supplementary material for a detailed output with individual models). The multiple regression approach suggested that coefficient shifts are largely idiosyncratic, irrespective of the metric used, although some effects of residual Moran's I, albeit weak, are found for all metrics of coefficient shift, and mainly for rSACe.

Our attempt to synthesize the results using canonical correlation also failed to identify clear causes of correlation shifts. The correlations between the canonical variates produced by the two sets of variables were statistically significant (Table 3), but explanatory power was low, as expected based on the results of the multiple regressions (Table 2). The first pair of canonical variates extracted 30.1% of the variance from the eight correlates and 11% of the variance from rSACe. Of greatest importance to our analysis, the first canonical variate summarizing the set of eight data-set characteristics accounted for 19% of the variance in the rSACe values, as indicated by the coefficient of redundancy, meaning that over 80% of the variance in rSACe remained unexplained. The first canonical variate estimated using the average differences between standardized coefficients (diff) explained 24.3% of the variance from the suite of variables associated with data-sets' characteristics and 14% from the sets of variables indicating differences in coefficients. The first canonical variate extracted from the data-set characteristics accounted for 8% of the variance in the differences between non-spatial and spatial models. Finally, the variances extracted for each set of variables from the canonical correlation using the Spearman rank correlation were 35.7 and 10.5%, and the characteristics of the data sets summarized by the first canonical variate explained 18.9% of the variance in the Spearman set. Thus, in no case was it possible to explain the

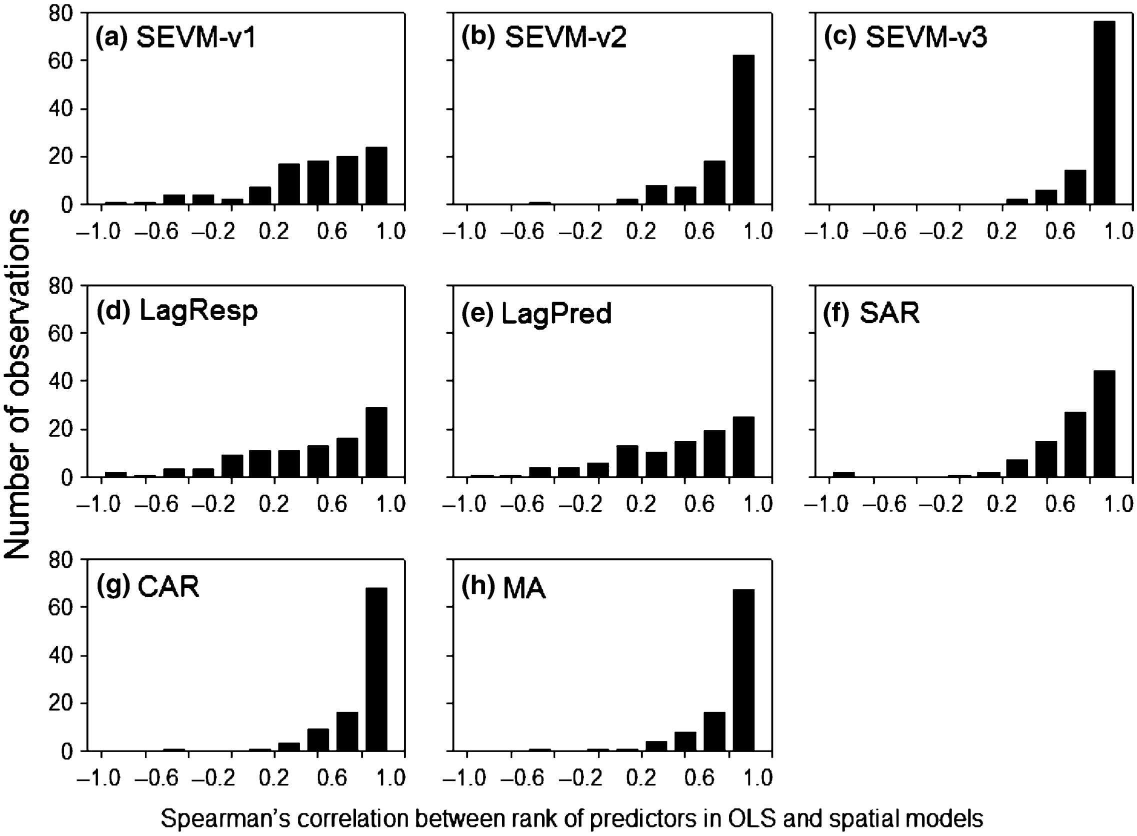

Figure 3. Frequency distribution (n = 97) of Spearman correlations between the ranks of the standardized partial coefficients estimated by OLS models and the ranks of these coefficients estimated by spatial methods. For method codes see Fig. 1.

Table 2. Averaged models estimated according to the Akaike weights of the models explaining three metrics of coefficient shifts (rSACe, difference and Spearman) between OLS and spatial regression techniques. Shown are the R^2 of the averaged model and the standardized regression coefficients of eight dataset characteristics. The strongest predictor of the metrics of coefficients shifts for each spatial method is indicated in bold.

Metric	Method	R^2	Predictors of coefficient shift							
			1	2	3	4	5	6	7	8
rSACe	SEVM-v1	0.08	**−0.28**	0.13	0.04	−0.04	−0.12	0.15	−0.02	−0.13
	SEVM-v2	0.31	−0.10	0.23	0.08	0.06	−0.26	**0.70**	−0.17	−0.45
	SEVM-v3	0.24	−0.10	0.21	−0.18	−0.04	−0.09	0.51	−0.16	**−0.60**
	LagResp	0.29	−0.03	0.06	0.17	0.15	**0.41**	−0.07	−0.19	0.07
	LagPred	0.19	−0.23	0.21	0.13	−0.26	−0.24	**0.30**	−0.09	−0.18
	SAR	0.28	−0.24	0.26	−0.06	−0.03	0.17	**0.61**	−0.11	−0.12
	CAR	0.16	−0.20	0.19	−0.18	0.08	0.19	**0.47**	0.03	−0.01
	MA	0.21	−0.43	0.26	−0.16	0.02	0.11	**0.55**	−0.08	−0.22
Difference	SEVM-v1	0.06	0.07	−0.02	0.15	0.01	0.12	**0.25**	0.00	−0.06
	SEVM-v2	0.08	0.12	0.00	−0.14	−0.14	−0.10	**−0.27**	0.06	0.05
	SEVM-v3	0.10	0.07	0.02	**0.31**	0.05	0.02	0.17	0.07	−0.25
	LagResp	0.17	**0.28**	−0.03	0.19	−0.04	0.26	0.10	0.09	0.09
	LagPred	0.11	0.17	0.06	0.15	**−0.22**	0.20	0.21	0.01	0.05
	SAR	0.15	0.11	−0.03	0.10	−0.06	0.23	**0.31**	0.08	−0.12
	CAR	0.10	0.08	0.01	0.03	−0.11	0.13	**0.32**	0.02	0.12
	MA	0.09	0.12	0.05	0.10	−0.09	0.08	**0.29**	0.01	−0.14
Spearman	SEVM-v1	0.03	**0.18**	−0.06	−0.10	0.02	−0.03	0.13	−0.08	0.12
	SEVM-v2	0.11	0.24	−0.09	0.07	−0.15	0.08	**−0.49**	0.11	0.40
	SEVM-v3	0.17	0.17	−0.06	0.06	0.20	0.10	−0.24	−0.05	**0.41**
	LagResp	0.21	0.06	0.05	−0.06	0.01	**−0.32**	−0.23	−0.03	−0.26
	LagPred	0.03	0.14	0.01	−0.11	−0.05	−0.06	−0.15	0.05	**−0.17**
	SAR	0.13	−0.01	−0.07	−0.10	0.06	−0.19	**−0.31**	−0.09	0.08
	CAR	0.07	0.09	−0.06	0.02	0.08	0.00	**−0.23**	−0.13	−0.16
	MA	0.04	**0.28**	0.06	−0.05	0.07	−0.09	−0.27	−0.02	−0.16

Variable 1: n; variable 2: m; variable 3: mean Moran's I (I_{ave}); variable 4: VIF; variable 5: adjusted R^2; variable 6: Moran's I–OLS residuals; variable 7: g1; variable 8: F (GWR versus OLS).

Table 3. Canonical correlations (R), chi-square (χ^2) values and significance levels (p) of the first three canonical variates (CV) extracted from analyses relating the metrics of coefficient shifts to data-set characteristics.

Metric	CV	R	χ^2	p
rSACe	CV 1	0.78	184.93	<0.001
	CV 2	0.64	102.48	<0.001
	CV 3	0.53	55.66	0.019
Differences	CV 1	0.59	108.29	<0.001
	CV 2	0.50	70.96	0.022
	CV 3	0.42	45.11	0.142
Spearman	CV 1	0.73	129.35	<0.001
	CV 2	0.50	62.55	0.093
	CV 3	0.37	37.47	0.402

bulk of variation in the magnitude of coefficient shifts using a wide range of data-set properties.

Although the canonical correlations derived from all three shift metrics were weak, shifts in standardized coefficients estimated by rSACe were relatively high in data sets with high adjusted coefficients of determination, high levels of non-stationarity, strong autocorrelation in the explanatory variables, high autocorrelation in the residuals of OLS models and large numbers of observations (Fig. 4a). LagResp, SAR and CAR were most strongly influenced by these data-set characteristics (Fig. 4b). Similar results were obtained using simple differences between standardized coefficients as a measure of coefficient shifts (Fig. 4c, d). The canonical correlation derived from Spearman correlations measuring the similarity in the importance of explanatory variables showed that LagResp and SAR differ the most from OLS. As with rSACe and diff, low correlations among ranks were mainly associated with the magnitude of autocorrelation in the residuals of OLS models, the level of non-stationarity of the data, numbers of observations, adjusted coefficient of determination, and the mean level of autocorrelation in the predictors (Fig. 4e, f). Variation across the second variate for this last analysis was discarded as the canonical correlation was not statistically significant. In sum, the multivariate approach found that it was difficult to predict when or why coefficients shift between spatial and non-spatial regression. And even when limited prediction was possible it was due to weak interactions among numerous characteristics of the data sets rather than a strong effect of any specific property of data. Notably, levels of SAC in the data (i.e. variables 1, 4 and 6) had little if any predictive power.

Discussion

A major goal of macroecological research is to identify the underlying causes of broad-scale ecological patterns, and standardized regression coefficients are widely believed to allow us to evaluate the relative importance of explanatory variables and to test hypotheses about drivers underlying these patterns. However, to date the extent to which coefficients differ between non-spatial OLS and spatially explicit methods and what data characteristics explain why such shifts occur is unclear. The most salient feature of our analysis is the lack of a clear signal related to the reasons why spatial regression techniques sometimes cause coefficient shifts. Two aspects of our results seem clear. First, although we were not able to find strong correlates of coefficient shifts, the extent to which shifts occur is highly sensitive to the spatial method used. This partly explains the vigorous discussion and conflicting findings in the literature regarding coefficient shifts, both in terms of their existence and their underlying drivers (Kühn 2007). Second, different metrics used to evaluate shifts in parameter estimates produced inconsistent results. Thus, the severity of coefficient shifts further depends on how they are measured and what is considered a shift (a minor change in estimates without changes in rank, or a major change in rank). Note that we are only considering the interpretation of regression coefficients and not other metrics, such as associated t-values, which should be more strongly affected by the standard errors of estimates.

Other workers have reported that some spatial methods generate parameter estimates that differ more from OLS than others (Dormann et al. 2007, Kissling and Carl 2008), and our findings support these results. Lagged models appear to be especially prone to generating outputs that differ from OLS models (Fig. 2a) and have been shown to generate unstable parameter estimates when tested with simple simulated data (Kissling and Carl 2008). SAR and spatial eigenvector mapping using all eigenvectors with spatial structure (SEVM variant 1) are also prone to change the ranks of explanatory variables compared to OLS regression. Thus, our results lend weak support for the conclusion that spatial methods in which new spatially structured variables are added to the regression as predictors tend to differ from OLS more than methods that focus on modelling spatial autocorrelation in the residuals (Dormann et al. 2007, Kissling and Carl 2008). This is reinforced somewhat by the fact that, although SEVM is based on adding variables to the model, the best performance of this method (as indicated by low levels of residual autocorrelation) was observed when eigenvectors were selected based directly on OLS residuals. This would be roughly analogous to fitting a GLS model based on spatial parameters estimated using a semi-variogram and to Griffith's and Peres-Neto (2006) method of selecting eigenvectors that minimize residual autocorrelation.

We used a single spatial structure for all spatial methods (i.e. Gabriel connections among sampling units), and it could be argued that some methods require different underlying spatial structures to take residual SAC fully into account. Although the method used to define the spatial structure of data is an important source of variation in results (Kissling and Carl 2008), we used a short-distance connection, which usually provides highly conservative results (Griffith 1996). For example, in the SAR models only seven data sets (out of 97) had residual Moran's Is $>$ 0.1, whereas 31 data sets had Moran's Is < -0.1, probably indicating overfitting in these cases. Nevertheless, to explore this issue we deleted the 38 data sets with Moran's Is of residual SAC >0.10 in the first distance class to determine whether it influenced our conclusions. We found that the frequencies of primary (37%) and secondary (49%) shifts were virtually identical to those when all data sets were included (39% for primary and 48% for secondary shifts). When repeating this analysis for all methods separately,

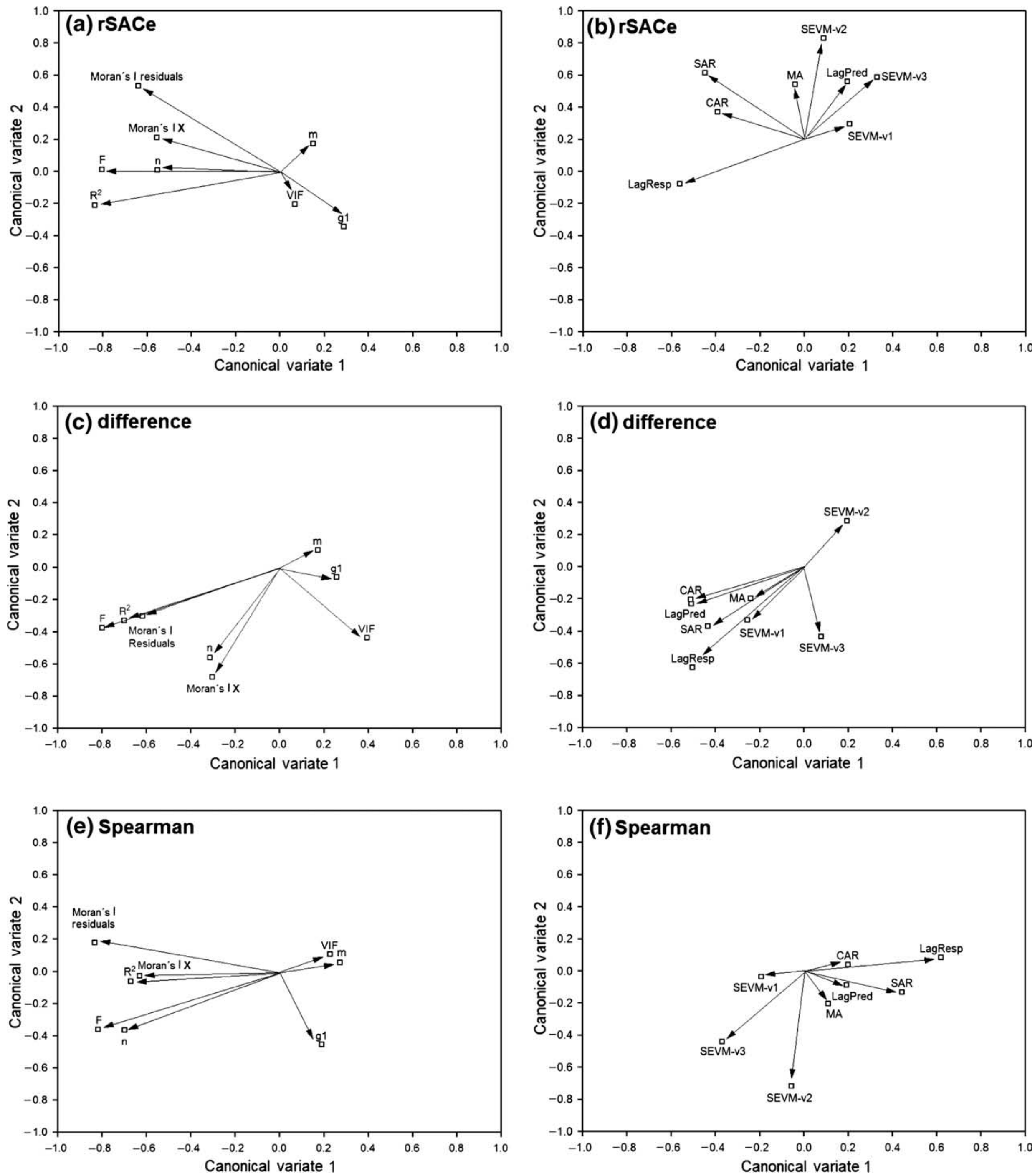

Figure 4. Pearson's correlations between the variables and their canonical variates for each of the three metrics of coefficient shift (a and b, for rSACe; c and d, for differences between standardized coefficients; e and f, for Spearman correlation between ranks of the standardized coefficients of the variables in OLS and spatial models). The two data matrices used in the canonical correlation analysis comprised 1) a given metric of coefficient shift estimated in relation to the different spatial models, and 2) the eight potential correlates of shifts calculated for each data set.

correlations between the frequency of primary and secondary shifts using the reduced and full data sets were 0.95 and 0.74, indicating that conclusions are quite similar even if we consider only data sets for which residual SAC was virtually eliminated using the Gabriel spatial structure. Furthermore, we found that changes in frequencies in primary and secondary shifts when removing the data sets with residual SAC occurred in the lagged models, reinforcing the conclusion that these methods are especially prone to differ from OLS. Lastly, although the definition of the optimal spatial structures underlying each spatial method deserves further analyses, this variability only reinforces the conclusion that results obtained using any spatial method are strongly data-dependent.

The spatial methods we used are generally sensitive to non-stationarity and non-linearity (Fotheringham et al. 2002), even though their performance in the face of this problem was not explicitly tested. The variable included in our analysis to measure the level of non-stationarity (F-stationarity) was not strongly correlated with coefficient shifts, indicating that the strength of violations in this assumption does not explain the differences between spatial and non-spatial methods. Despite this, it remains true that most macroecological data sets are non-stationary. As it is not good practice to use a statistical method when the data do not meet its underlying assumptions, we believe that most spatial methods should not be used without first confirming that the data are stationary, and any non-stationarity should be reported. In regions where the data are not stationary, GWR might be a good method to understand patterns at more local scales and evaluate changes in model structure within subsets of data, whereas global models can be useful for explaining general patterns across the full extent of the data.

Our analysis suggests that standardized coefficients may be extremely difficult to interpret, even if we ignore the causality issues that plague all multiple regression analyses whether spatially explicit or not (Shipley 2000). One common approach is to report models using OLS with equivalent models derived from one or two spatial methods (Tognelli and Kelt 2004, Kissling et al. 2008, Ramirez et al. 2008). If the coefficients are similar, authors conclude that the results are consistent, but if the coefficients differ, authors tend to discount the OLS results in favor of the spatial models (Lichstein et al. 2002, Bahn et al. 2006, Kühn 2007). However, neither of these strategies may be tenable. First, whether coefficients shift will depend on the spatial method selected and how spatial structure is modeled (Kissling and Carl 2008). Accordingly, robustness and stability of spatial regression models may be an illusion unless a large number of spatial methods are compared and are found to be consistent (which is not necessarily the case, as shown here). Second, and as seriously, we found no evidence that it is possible to predict when coefficients will shift, based on the low R^2 of our models (Table 2). Although we did not examine all possible sources of coefficient shifts, we tested a large number of obvious potential causes, including the inherent spatial structure of the variables, the explanatory power of the models in terms of both overall model fit and their ability to capture spatial structure at multiple scales, as well as different sources of model misspecification, including the quantity and distribution of spatial structure in model residuals, collinearity among predictors, and stationarity of slopes. This inability to understand the behavior of standardized regression coefficients in the face of spatial structure could be used to argue that the usefulness of multiple regression in geographical ecology is limited to purely descriptive exercises. We cannot evaluate the merits of this interpretation here, but it is not a good idea to accept uncritically the results of any multiple regression (spatial or non-spatial), nor to assume that the result from one method is more "correct" than another. Thus, we add one more item to the long list of problems related to the use of multiple regression in ecology (Olden and Jackson 2000, Graham 2003, Grace and Bollen 2005, Whittingham et al. 2006 and references therein).

If analysts agree that no spatial autocorrelation should remain in model residuals, and at the same time accept that it is not always possible to quantify all biological and ecological processes generating the spatial structure in data, one approach would be to focus on methods that minimize both spatial autocorrelation in residuals and coefficient shifts. Two candidate methods seem to satisfy this requirement reasonably well, SEVM-v3 and MA (Fig. 1, 2). However, although these methods may be better than randomly selecting a spatial model, it does not change the fact that different coefficients can be obtained from alternative spatial approaches for any particular dataset. If workers remain concerned that a single multiple regression approach, including OLS, is sufficient to understand the complex interactions found in data, an alternative approach is to fit different (carefully selected) models on the same data to evaluate the general congruence of results.

Our results are a reminder of the limits to the ecological interpretation of standardized coefficients (see also Grace and Bollen 2005). Although it is widely known that correlation does not reflect causation irrespective of how strong a relationship may be, standardized coefficients (with or without associated t-values) are almost universally used to rank explanatory variables in terms of their "importance". Since t-tests are affected by the presence of spatial autocorrelation, they should not be used to evaluate explanatory variables. However, our analyses suggest that the standardized coefficients themselves may be uninterpretable. If coefficients vary when different methods are used, and their values depend on complex interactions between multivariate data sets and the algorithms used to model spatial covariance structures, it becomes difficult to argue for the superiority of one model, or even one variable, over another. Some of these issues could perhaps be resolved by using simulation procedures, and indeed some workers (Dormann et al. 2007, Kissling and Carl 2008) have used simple statistical rules to create richness as a function of environmental variables plus autocorrelated errors. Although we do not claim that the results based on simulated data are without value, it is important to realize that these simulations are simplistic, because in empirical data sets richness (the response variable in the majority of macroecological studies) is most often generated by the range overlap of different species in a given spatial unit of analysis (e.g. a cell). A more realistic scenario might be achieved by simulating ranges in response to environmental variables and then calculating species richness as an emergent property. The main difficulty, however, is to translate the range models to richness models with known parameters values. We think that this could be an interesting avenue for further studies.

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.

Acknowledgements – The western Andalusia database generated by RGA and AA was supported by the Andalusian Regional Government (Consejería de Medio Ambiente and Consejería de Innovación, Ciencia y Empresa project P06-RNM-01499). MÁR, MR and MÁO-T were supported by the Spanish Ministry of Science and Innovation (grant CGL2006-03000/BOS and FPU fellowship AP2005-0636 to MÁO-T). Work by LMB, JAFDF, TFLVBR, PMJ, JML and MBA on species distribution modeling and spatial analysis was supported by the BBVA Foundation BIO-IMPACT project. LMB, JAFDF and PMJ have been supported by several CNPq grants, and TFLVBR was further supported by a CAPES/Fulbright PhD Fellowship.

References

Araújo, M. B. and New, M. 2007. Ensemble forecasting of species distributions. – Trends Ecol. Evol. 22: 42–47.

Araújo, M. B. et al. 2008. Quaternary climate changes explain diversity among reptiles and amphibians. – Ecography 31: 8–15.

Bahn, V. et al. 2006. Importance of spatial autocorrelation in modeling bird distributions at a continental scale. – Ecography 29: 835–844.

Bickford, S. A. and Laffan, S. W. 2006. Multi-extent analysis of the relationship between pteridophyte species richness and climate. – Global Ecol. Biogeogr. 15: 588–601.

Bini, L. M. et al. 2000. Local and regional species richness relationships in viperid snake assemblages from South America: unsaturated patterns at three different spatial scales. – Copeia 2000: 799–805.

Borcard, D. and Legendre, P. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. – Ecol. Model. 153: 51–68.

Borcard, D. et al. 1992. Partialling out the spatial component of ecological variation. – Ecology 73: 1045–1055.

Borcard, D. et al. 2004. Dissecting the spatial structure of ecological data at multiple scales. – Ecology 85: 1826–1832.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference. A practical information – theoretical approach. – Springer.

Cassemiro, F. A. S. et al. 2007. Non-stationary, diversity gradients and the metabolic theory of ecology. – Global Ecol. Biogeogr. 16: 820–822.

Cressie, N. A. C. 1993. Statistics for spatial data analysis. – Wiley.

Diniz-Filho, J. A. F. and Bini, L. M. 2005. Modelling geographical patterns in species richness using eigenvector-based spatial filters. – Global Ecol. Biogeogr. 14: 177–185.

Diniz-Filho, J. A. F. et al. 2003. Spatial autocorrelation and red herrings in geographical ecology. – Global Ecol. Biogeogr. 12: 53–64.

Diniz-Filho, J. A. F. et al. 2008. Model selection and information theory in geographical ecology. – Global Ecol. Biogeogr. 17: 479–488.

Dormann, C. F. 2007. Effects of incorporating spatial autocorrelation into the analysis of species distributional data. – Global Ecol. Biogeogr. 16: 129–138.

Dormann, C. F. et al. 2007. Methods to account for spatial autocorrelation in the analysis of distributional species data: a review. – Ecography 30: 609–628.

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 Ecol. Biogeogr. 13: 315–320.

Fortin, M.-J. and Payette, S. 2002. How to test the significance of the relation between spatially autocorrelated data at the landscape scale: a case study using fire and forest maps. – Ecoscience 9: 213–218.

Fortin, M.-J. and Dale, M. 2005. Spatial analysis: a guide for ecologists. – Cambridge Univ. Press.

Fotheringham, A. S. et al. 2002. Geographically weighted regression: the analysis of spatially varying relationships. – Wiley.

Grace, J. B. and Bollen, K. A. 2005. Interpreting the results from multiple regression and structural equation models. – Bull. Ecol. Soc. Am. 86: 283–294.

Graham, M. H. 2003. Confronting multicollinearity in ecological multiple regression. – Ecology 84: 2809–2815.

Griffith, D. A. 1996. Some guidelines for specifying the geographic weights matrix contained in spatial statistical models. – In: Arlinghaus, S. L. (ed.), Practical handbook of spatial statistics. CRC Press, pp. 65–82.

Griffith, D. A. and Peres-Neto, P. R. 2006. Spatial modeling in ecology: the flexibility of eigenfunction spatial analysis. – Ecology 87: 2603–2613.

Haining, R. 1990. Spatial data analysis in the social and environmental sciences. – Cambridge Univ. Press.

Haining, R. 2002. Spatial data analysis. – Cambridge Univ. Press.

Hawkins, B. A. et al. 2003. Productivity and history as predictors of the latitudinal diversity gradient of terrestrial birds. – Ecology 84: 1608–1623.

Hawkins, B. A. et al. 2007. Red herrings revisited: spatial autocorrelation and parameter estimation in geographical ecology. – Ecography 30: 375–384.

Hobbs, N. T. and Hilborn, R. 2006. Alternatives to statistical hypothesis testing in ecology: a guide to self teaching. – Ecol. Appl. 16: 5–19.

Hoeting, J. A. et al. 2006. Model selection for geostatistical models. – Ecol. Appl. 16: 87–98.

Kissling, W. D. and Carl, G. 2008. Spatial autocorrelation and the selection of simultaneous autoregressive models. – Global Ecol. Biogeogr. 17: 59–71.

Kissling, W. D. et al. 2008. Spatial patterns of woody plant and bird diversity: functional relationships or environmental effects? – Global Ecol. Biogeogr. 17: 327–339.

Kühn, I. 2007. Incorporating spatial autocorrelation may invert observed patterns. – Divers. Distrib. 13: 66–69.

Legendre, P. 1993. Spatial autocorrelation – trouble or new paradigm? – Ecology 74: 1659–1673.

Legendre, P. and Legendre, L. 1998. Numerical ecology, 2nd English ed. – Elsevier.

Legendre, P. et al. 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. – Ecography 25: 601–616.

Lennon, J. J. 2000. Red-shifts and red herrings in geographical ecology. – Ecography 23: 101–113.

Lichstein, J. W. et al. 2002. Spatial autocorrelation and autoregressive models in ecology. – Ecol. Monogr. 72: 445–463.

Olden, J. D. and Jackson, D. A. 2000. Torturing data for the sake of generality: how valid are our regression models? – Ecoscience 7: 501–510.
Ramirez, L. et al. 2008. Partitioning phylogenetic and adaptive components of the geographic body size pattern of New World birds. – Global Ecol. Biogeogr. 17: 100–110.
Rangel, T. F. L. V. B. et al. 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. – Global Ecol. Biogeogr. 15: 321–431.
Schabenberg, O. and Gotway, C. A. 2005. Statistical methods for spatial data analysis. – Chapman and Hall.
Sokal, R. R. and Rohlf, F. J. 1995. Biometry, 3rd ed. – W. H. Freeman.
Shipley, B. 2000. Cause and correlation in biology. – Cambridge Univ. Press.
Tognelli, M. F. and Kelt, D. A. 2004. Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. – Ecography 27: 427–436.
Wang, Q. et al. 2005. Application of a geographically-weighted regression analysis to estimate net primary production of Chinese forest ecosystems. – Global Ecol. Biogeogr. 14: 379–393.
Whittingham, M. J. et al. 2006. Why do we still use stepwise modelling in ecology and behaviour? – J. Anim. Ecol. 75: 1182–1189.

Download the Supplementary material as file E5717 from <www.oikos.ekol.lu.se/appendix>.