Gaspard et al. Journal of Ecology and Environment (2019) 43:19
https://doi.org/10.1186/s41610-019-0118-3

Journal of Ecology and Environment

REVIEW

Open Access

Residual spatial autocorrelation in macroecological and biogeographical modeling: a review

Guetchine Gaspard[1], Daehyun Kim[2,3*] and Yongwan Chun[4]

Abstract

Macroecologists and biogeographers continue to predict the distribution of species across space based on the relationship between biotic processes and environmental variables. This approach uses data related to, for example, species abundance or presence/absence, climate, geomorphology, and soils. Researchers have acknowledged in their statistical analyses the importance of accounting for the effects of spatial autocorrelation (SAC), which indicates a degree of dependence between pairs of nearby observations. It has been agreed that residual spatial autocorrelation (rSAC) can have a substantial impact on modeling processes and inferences. However, more attention should be paid to the sources of rSAC and the degree to which rSAC becomes problematic. Here, we review previous studies to identify diverse factors that potentially induce the presence of rSAC in macroecological and biogeographical models. Furthermore, an emphasis is put on the quantification of rSAC by seeking to unveil the magnitude to which the presence of SAC in model residuals becomes detrimental to the modeling process. It turned out that five categories of factors can drive the presence of SAC in model residuals: ecological data and processes, scale and distance, missing variables, sampling design, and assumptions and methodological approaches. Additionally, we noted that more explicit and elaborated discussion of rSAC should be presented in species distribution modeling. Future investigations involving the quantification of rSAC are recommended in order to understand when rSAC can have an adverse effect on the modeling process.

Keywords: Spatial autocorrelation, Residual spatial autocorrelation, Non-stationarity, Missing variables, Sampling design, Scale, Species distribution models

Background

Spatial autocorrelation

The use of spatial or geographical data entails learning about the properties of such data. Disciplines in which geographic data are used are all concerned with how such data are characterized, whether it be geography, ecology, or any related field where the space and time factors are involved. One of the most common issues regarding spatial data is the existence of structure or dependence among the observations. Often, processes, whether they be environmental or biological, are related spatially or temporally. This fact translates the notion of distance decay wherein the degree of dependence decreases over space. This is the basis of Tobler's (1970) first law of geography: *everything is related to everything else, but nearby things are more related than distant things*. This whole argument falls under the concept of spatial autocorrelation (SAC). This term, which was introduced around the late 1960s and early 1970s (Getis 2008), is loosely defined as follows:

> The property of random variables taking values, at pairs of locations a certain distance apart, that are more similar (positive autocorrelation) or less similar (negative autocorrelation) than expected for randomly associated pairs of observations (Legendre 1993, p. 1659).

* Correspondence: biogeokim@snu.ac.kr
[2]Department of Geography, Seoul National University, Seoul, South Korea
[3]Institute for Korean Regional Studies, Seoul National University, Seoul, South Korea
Full list of author information is available at the end of the article

© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Depending on the factors that drive natural processes, SAC is categorized into two major types: exogenous and endogenous SAC (Legendre 1993). The former is caused by external environmental (*physico-chemical*, *climatological*, *geomorphological*) factors such as temperature, soil, and terrain attributes (Dormann 2007a; Kissling and Carl 2008; Miller 2012; Václavík et al. 2012). It is generally associated with broad-scale spatial trends (Miller et al. 2007; Václavík et al. 2012). Endogenous SAC is induced by biological (or biology-related) processes (*geographic dispersal*, *predation*, *disturbance*, *inter-specific interactions*, *colonial breeding*, *home-range size*, *host availability*, *parasitization risk*, *metapopulation dynamics*, *history*) that are inherent to the species data (Dormann 2007a; Kissling and Carl 2008; Miller 2012; Crase et al. 2014). It reflects contagion effects in cases of positive autocorrelation or dispersion effects for negative autocorrelation (Lichstein et al. 2002; Griffith and Peres-Neto 2006; Crase et al. 2014). Such endogenous SAC is relevant at fine scales or to high-resolution stochastic biotic processes (Dormann 2007a; Miller et al. 2007; Chun and Griffith 2011; Václavík et al. 2012; Kim 2018).

Residual spatial autocorrelation

In the modeling context, residuals represent the differences between observed and predicted values. Hence, rSAC indicates the amount of SAC in the variance which is not explained by explanatory variables. Understanding residuals distribution is key to regression modeling, as assumptions such as linearity, normality, homoscedasticity (equal variance), and independence rely on the behavior of the errors.

Incorporating or ignoring rSAC has implications directly impacting the outcomes of species distribution modeling (SDM). In fact, failing to appropriately address rSAC will likely lead to three major statistical problems. First, the standard errors might be underestimated, leading to Type I error. This means that the existence of dependence between pairs of observations across space where independence is assumed can result in falsely rejecting, much more often than expected, the null hypothesis while it is true (Lennon 2000). Consequently, that will make the regression model itself unreliable (Legendre 1993; Anselin 2002; Kim et al. 2016). Second, parameter estimates, such as the regression coefficients and *F*-statistic, might be biased (Dormann 2007a; Václavík et al. 2012). The inflation or deflation of predictors' coefficients will induce the over- or under-estimation of their predictive power, respectively. Finally, model misspecification, related to variable selection, remains an important issue (Austin 2002; Lichstein et al. 2002; Miller et al. 2007; Václavík et al. 2012). The presence of SAC in model residuals is typical of spatial ecological data (Borcard et al. 1992; Lennon 2000; Dormann 2007a; Kissling and Carl 2008; Bini et al. 2009; Kim and Shin 2016); therefore, using these types of data usually violates the assumption of independence between pairs of observations, necessitating that the effects of rSAC be accounted for (Diniz-Filho and Bini 2005; Bahn et al. 2006).

Species distribution modeling

The views of previous species distribution modeling studies are mixed in regard to certain effects of SAC on the outcomes of spatial predictive models. In some articles (e.g., Lennon 2000; Dormann 2007a; Kim et al. 2016), the three statistical consequences briefly mentioned in the preceding section are well recognized. For example, Lennon (2000) urged ecologists to start integrating SAC in their modeling. Convinced of the ill effects of failing to incorporate SAC in ecological data modeling, he took a strong stance suggesting that such effects can invalidate previous works that used standard non-spatial models (e.g., ordinary least squares; OLS). In other research (Dormann 2007a; Kim et al. 2016), the voice was moderate. That is, despite the fact that spatially explicit models generally outperform their non-spatial counterparts (i.e., greater R^2 or lower rSAC), the final conclusions were rather tentative. In his review, Dormann (2007a) estimated, on average, a positive coefficient shift in favor of a spatial model as high as 25% and concluded that in certain methodological conditions, such models showed an edge over non-spatial models. Subsequent to Dormann's (2007a) review, two studies (Kim 2013; Kim et al. 2016) consistently witnessed a better performance of spatially explicit models over non-spatial ones. However, it was concluded that whether that superiority holds true for any spatial methods, sampling strategies or field designs remains to be seen. It was suspected that whether data were collected randomly, on a grid, in a nested or stratified fashion, or how densely the samples are distributed might make a difference in the modeling outcomes. As compelling and relevant as SAC appears to be, only a minority of published studies in the ecological field—for example, less than 20% (Dormann 2007a) or 3 out of 44 (Crase et al. 2014)—working with spatial data have addressed the issue.

On the other hand, there are other studies (e.g., Diniz-Filho et al. 2003; Hawkins et al. 2007; Bini et al. 2009; Miller 2012) in which the abovementioned claims were not agreed. To wit, the question concerning which parameter estimates in non-spatial modeling (models that do not account for SAC) are biased was not a critical issue. For example, Hawkins et al. (2007) warned about claiming the superiority of spatial models and the falseness of non-spatial ones as they found no significant differences between global OLS models and spatial models, especially when using gridded data. For them, the assumption that non-spatial models are automatically flawed, as argued by

Lennon (2000), in comparison with spatial models was a mistake. Moreover, changes in coefficients between spatial and non-spatial models were mainly idiosyncratic and depended on the type of method used (Bini et al. 2009), which suggests that modelers should be explicit and cautious in their claims. These conclusions were already drawn in previous studies where non-spatial regression models were found unbiased. Additionally, these conclusions recommended that the scale factor be considered when interpreting results (Hawkins et al. 2007). Therefore, claiming that models that ignore SAC are flawed is groundless (Diniz-Filho et al. 2003). In addition, mathematical analyses show that neither coefficients of spatial models nor those of non-spatial models are totally unbiased; in fact, precision decreases for spatial model coefficients as SAC increases (Miller 2012).

Justification

A substantial number of studies in biogeography and macroecology have broadly covered the topic of SAC, but little is known about how deeply those works have discussed the case of rSAC. Previous studies suspect that failing to include certain explanatory variables might be at the heart of the problem (Crase et al. 2014). This problem, when related to the endogenous rather than the external type of SAC, remains unexplored. An effort to identify potential missing variables and establish how much their omission increases the level of rSAC would potentially bring new knowledge and contribute to the SDM literature body. Along with environmental and biotic missing predictors, the type of sampling design will also be scrutinized. Sampling design is often mentioned as having the ability to potentially cause rSAC to increase (Lichstein et al. 2002; Bini et al. 2009; Crase et al. 2014). This present paper addresses sampling design in terms of sample size, data type, sampling technique, and the effect of small scales in particular. Analyzing data at very fine scales coupled with the inclusion of important spatially autocorrelated missing variables is believed to have the potential to significantly reduce or even remove rSAC in species distribution models. Assuming that environmental factors behave differently at distinct spatial scales, Diniz-Filho et al. (2003) suggest that the inclusion of relevant environmental factors acting at each scale in a regression model would eventually remove SAC from the residuals at different scales.

Our goal in this review article is to evaluate an umbrella research question: Under what circumstances can the magnitude of rSAC increase? This question is broken down into the following three sub-questions:

1. What are the causes of rSAC?
2. How much do missing variables account for rSAC?
3. How do different sampling designs influence the level of SAC in model residuals?

Completing this investigation is expected to accomplish the following: (1) establish the full picture of rSAC in the existing literature of macroecological and biogeographical modeling and (2) serve as a foundation to conduct further research on rSAC.

Articles search, selection, and categorization

In this review, we initially targeted articles from macroecology and biogeography that dealt with SDM in which SAC was explicitly incorporated. We used keywords such as *residual spatial autocorrelation*, *spatial autocorrelation*, *ecological*, or *biogeographical*, as well as *species distribution modeling*, to search for relevant articles via the Web of Science and Google Scholar engines. We also selected additional articles quoted and referred to by some of these original selections. Thus, some of the studies reviewed in this paper were not exactly from the macroecology and biogeography fields. The subjects of these additional articles belonged to the disciplines of hydrology, soil science, and geomorphology, but they still covered important aspects of SAC in terms of methods, functions, history, and modeling.

As a result, we have chosen a total of 97 articles dating from 1984 to 2017 (Table 1). These articles were carefully reviewed and then categorized based on the level of detail they discussed on rSAC. In the end, we attempted to understand the conditions under which SAC occurs—and magnifies—in model residuals.

In terms of approach, the articles reviewed were all unique with respect to SAC modeling in geographical ecology. However, SDM remained as the most studied topic across the board (61% of the articles), followed by habitat suitability modeling (22%) and methods (16%). The remaining proportion discussed other aspects of SAC modeling. The modeling included many species, such as birds, plants, mammals, and reptiles. Here are some proxies used as dependent variables: richness, occurrence, abundance, presence and absence, occupancy, composition, dispersal, diversity, and density. For habitat suitability, some surrogates were niche suitability, habitat distribution, climatic suitability, climatic forecast, or predictability.

Potential sources of residual SAC in SDM

Reviews of the existing literature revealed that accounting for SAC in SDM still has a long way to go, even though studies have increasingly strived to broadly incorporate the effect of spatial dependence in investigating ecological and biogeographical processes over the last three decades. We found that only a small proportion (less than 20%) of ecological and biogeographical modelers incorporated SAC in their research. This is due partly to the fact that the need to incorporate SAC has yet to become unanimous among modelers (Diniz-Filho et al. 2003;

Table 1 Literature review in macroecological and biogeographical modeling. *SAC* spatial autocorrelation, *rSAC* residual spatial autocorrelation

Number	Author	Year	Journal	rSAC	Subject
1	Bahn et al.	2006	*Ecography*	Elaborate	Bird distribution
2	Bini et al.	2009	*Ecography*	Elaborate	Spatial and non-spatial regression
3	Borcard et al.	1992	*Ecology*	Elaborate	Partialling out species abundance
4	Bonada et al.	2012	*Journal of Biogeography*	Elaborate	Richness and composition invertebrates
5	Crase et al.	2012	*Ecography*	Elaborate	rSAC in mangrove species distribution
6	Crase et al.	2014	*Global Change Biology*	Elaborate	Mangrove species distribution and forecast
7	Diniz-Filho et al.	2003	*Global Ecology and Biogeography*	Elaborate	Species richness of bird
8	Diniz-Filho and Bini	2005	*Global Ecology and Biogeography*	Elaborate	Bird species richness and SAC
9	Diniz-Filho et al.	2008	*Global Ecology and Biogeography*	Elaborate	Model selectin in mammal species
10	Dormann	2007a	*Global Ecology and Biogeography*	Elaborate	Spatial and non-spatial models in ecology
11	Griffith and Peres-Neto	2006	*Ecology*	Elaborate	Eigenfunction in ecological modeling
12	Griffith	2000	*Journal of Geographical Systems*	Elaborate	Regression modeling of geo-demographic data
13	Hawkins et al.	2007	*Ecography*	Elaborate	Analyzing coefficient shifts in bird species richness
14	Kühn	2007	*Diversity and Distributions*	Elaborate	Plant species richness and environmental correlates
15	Kim et al.	2013	*Physical Geography*	Elaborate	Multiple SAC in soil moisture and landscape
16	Kim et al.	2016	*Soil Science Society of America Journal*	Elaborate	Multiple SAC in soil–landform modeling
17	Kissling and Carl	2008	*Global Ecology and Biogeography*	Elaborate	SAC and model selection
18	Lichstein et al.	2002	*Ecological Monographs*	Elaborate	Models and breeding habitats of songbirds
19	de Oliveira et al.	2012	*Biodiversity Conservation*	Elaborate	Climatic suitability of biome in climate change
20	de Oliveira et al.	2014	*Ecography*	Elaborate	Ecological niche modeling of plant species
21	Sheehan et al.	2017	*Ecology and Evolution*	Elaborate	Bird species habitat
22	Ortiz-Yusty et al.	2013	*Caldesia*	Elaborate	Species richness and climate
23	Pickup and Chewings,	1986	*Ecological Modelling*	Elaborate	Prediction of erosion and deposition
24	Le Rest et al.	2014	*Global Ecology and Biogeography*	Elaborate	Variable selection in species abundance
25	Revermann et al.	2012	*Journal of Ornithology*	Elaborate	Bird species habitat and climate change
26	Václavík et al.	2012	*Journal of Biogeography*	Elaborate	Multi-scale SAC and invasive forest pathogen distribution
27	Veloz	2009	*Journal of Biogeography*	Elaborate	Niche modeling and plant species distribution
28	Wu and Zhang	2013	*Applied Geography*	Elaborate	Model comparison and occurrence of cloud cover
29	Siesa et al.	2011	*Biological Invasions*	Elaborate	SAC and crayfish distribution
30	Piazzini et al.	2011	*Journal of Herpetology*	Elaborate	SAC and presence of reptile species
31	Ishihama et al.	2010	*Ecological Resources*	Elaborate	Distribution of herbaceous species
32	Record et al.	2013a, b	*Global Ecology and Biogeography*	Elaborate	Plant species distribution projection and SAC
33	Naimi et al.	2011	*Journal of Biogeography*	Elaborate	SAC and species occurrence modeling
34	Ficetola et al.	2012	*Ecography*	Elaborate	SAC and reptile species dispersal

Table 1 Literature review in macroecological and biogeographical modeling. *SAC* spatial autocorrelation, *rSAC* residual spatial autocorrelation *(Continued)*

Number	Author	Year	Journal	rSAC	Subject
35	Dormannn	2007b	*Ecological Modelling*	Elaborate	SAC and species distribution
36	Wu et al.	2009	*Ecological Modelling*	Elaborate	SAC and landscape dynamics
37	Merckx et al.	2009	*Ecological Modelling*	Elaborate	SAC and prediction of marine nematode biodiversity
38	Dowd et al.	2014	*Ecological Applications*	Elaborate	Coastal marine benthic microfaunal distribution modeling
39	Hefley et al.	2017a, b	*Ecology*	Elaborate	Modeling SAC in ecological data
40	Betts et al.	2006	*Ecological Modelling*	Elaborate	SAC and forest bird occurrence
41	Mets et al.	2017	*Ecosphere*	Elaborate	SAC in deforestation modeling
42	Tallowin et al.	2017	*Journal of Biogeography*	Elaborate	Terrestrial vertebrate richness
43	Hindrikson et al	2017	*Biological Reviews*	Elaborate	Wolf species richness and distribution
44	Record et al.	2013a, b	*Ecosphere*	Elaborate	Climate change prediction
45	Austin	2002	*Ecological modelling*	Elaborate	Species distribution modeling
46	Carl and Kühn	2007	*Ecological Modelling*	Elaborate	SAC in Species distribution
47	Dirnböck and Dullinger	2004	*Journal of Vegetation Science*	Elaborate	Species distribution modeling
48	Zhang et al.	2009	*Forest Science*	Elaborate	Species model comparison
49	Gwenzi and Lefsky	2017	*IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*	Elaborate	SAC and plant biomass
50	Roth et al.	2016	*American naturalist*	Elaborate	Interactions among endangered species
51	Davis et al.	2016	*Ecosphere*	Elaborate	Urban plant invasion
52	Mattsson et al.	2013	*PLOS ONE*	Simple mention	Species assemblage
53	Chun and Griffith	2011	*Annals of the Associations of American Geographers*	Simple mention	Network SAC and migration flows
54	Cliff	1984	*Journal of the American Statistical Association*	Simple mention	Correlation estimation between scores
55	Getis	2008	*Geographical Analysis*	Simple mention	History of SAC
56	Miller et al.	2007	*Ecological Modelling*	Simple mention	SAC and predictive vegetation modeling
57	Lennon	2000	*Ecography*	Simple mention	SAC and geographical ecology
58	Zhu et al.	2012	*Journal of Geographical Science*	Simple mention	SAC and vegetation cover.
59	Poley et al.	2014	*Journal of Biogeography*	Simple mention	SAC and large mammals' occupancy
60	Jackson et al.	2015	*Biological Conservation*	Simple mention	Prediction of bird species habitat
61	Platts et al.	2008	*Ecological Modelling*	Simple mention	Model selection in tree distribution
62	Hefley et al.	2017a, b	*Ecology*	Simple mention	Functions in spatial ecological modeling
63	Estrada and Rodriguez-Estrella	2016	*Animal Conservation*	No mention	Biodiversity-bird species
64	Ali et al.	2010	*Water Resources Research*	No mention	Soil moisture and topographical modeling
65	Anselin and Bera	1998	*Handbook of Applied Statistics*	No mention	SAC and regression models
66	Santos et al.	2009	*Canadian Journal of Zoology*	No mention	SAC in pine species
67	Dorken et al.	2017	*Journal of Ecology*	No mention	Plant species density
68	Ennen et al.	2016	*Canadian Journal of Zoology*	No mention	Reptile pattern modeling
69	Weeks et al.	2017	*River Research and Applications*	No mention	Snail and aquatic vegetation

Table 1 Literature review in macroecological and biogeographical modeling. *SAC* spatial autocorrelation, *rSAC* residual spatial autocorrelation *(Continued)*

Number	Author	Year	Journal	rSAC	Subject
70	Dronova et al.	2016	*Remote Sensing*	No mention	Bird species diversity
71	Anselin et al.	2006	*Geographical Analysis*	No mention	Spatial effects in environmental economics
72	Augustin	2001	*Journal of Applied Ecology*	No mention	Succession in semi-natural vegetation
73	Chang et al.	2012	*PLOS ONE*	No mention	Genetic and bird species distribution
74	Seymour	2005	*Journal of the American Statistical Association*	No mention	Spatial data: theory and practice
75	Siderov	2005	*Austral Ecology*	No mention	SAC practice and theory
76	Hongoh et al.	2012	*Applied Geography*	No mention	Mosquito distribution
77	Miller	2012	*Progress in Physical Geography*	No mention	Species distribution modeling
78	Kleisner et al.	2010	*Marine Ecology Progress Series*	No mention	Pelagic fish modeling
79	Tarkhnishvili et al.	2012	*Biological Journal of the Linnean Society*	No mention	Distribution of forest species
80	Wiegand and Moloney	2004	*Oikos*	No mention	Point pattern analysis in ecology
81	Yu	2012	*Ph.D. Dissertation*	No mention	Tree growth modeling and seedling recruitment
82	Lloyd et al.	2005	*Diversity and Distributions*	No mention	SAC and Benthic invertebrates
83	Rodriguez et al.	2015	*Journal Insect Conservation*	No mention	Distribution of oak wasp species
84	Nicolaus et al.	2013	*Journal Evolution Biology*	No mention	Gastropod mollusk distribution
85	Warren et al.	2014	*Trends in Ecology and Evolution*	No mention	Species distribution modeling
86	Wieczorek and Bugaj-Nawrocka	2014	*Agricultural and Forest Entomology*	No mention	Ecological niche modeling
87	Epperson	2000	*Ecological Modelling*	No mention	Space-time and ecological modeling
88	Wulder et al.	2007	*Ecological Modelling*	No mention	Forest growth modeling
89	Büchi et al.	2009	*Ecological Modelling*	No mention	Meta-community and species distribution
90	Marmion et al.	2009	*Ecological Modelling*	No mention	Butterfly species distribution
91	Legendre	1993	*Ecology*	No mention	SAC trouble or paradigm in ecology
92	Guénard et al.	2016	*Ecosphere*	No mention	Fish-spatial modeling
93	Estrada et al.	2016	*PLOS ONE*	No mention	Habitat suitability
94	Ingberman et al.	2016	*PLOS ONE*	No mention	Muriquis distribution
95	Ciccarelli and Bacaro	2016	*Folia Geobotanica*	No mention	Spatial modeling and species diversity
96	Güler et al.	2016	*Journal of Vegetation Science*	No mention	Plant species richness
97	Komac et al.	2016	*PLOS ONE*	No mention	Habitat suitability

Hawkins et al. 2007; Bini et al. 2009; Miller 2012). The presence of SAC in ecological and biogeographical data has long been identified (as far back as the late 1970s), and statistical methods to address it were developed almost in the same period (Dormann 2007a). For example, Legendre (1993) defined and categorized the concept of SAC into endogenous and exogenous SAC in the field of ecological data modeling. However, modelers started substantially publishing studies that integrate SAC after 2000. This reality agrees with the reason why 92 out of the 97 articles we reviewed were published in the new millennium. Some of the earlier studies that acknowledged the effect of SAC prior to 2000 include, but are not limited to, Borcard et al. (1992) who looked at partialling out the total variance of species abundance into spatial and non-spatial components and Pickup and Chewings (1986) who investigated the prediction of erosion and deposition in alluvial landscapes of central Australia.

These discussions explain why rSAC, as a subcategory of SAC, remains relatively unexplored in ecological and biogeographical modeling. We categorized the articles into three groups (i.e., *no mention*, *simple mention*, and *elaborate*) based on the level of details at which a discussion is provided on rSAC (Table 2). In fact, 35 articles (36%) never mentioned the presence or influence of rSAC. The remaining 62 (*simple mention* plus *elaborate*)

Table 2 Summary of the reviewed articles with regard to the level of detail they provided on residual spatial autocorrelation

Category	No mention	Elaborate	Simple mention
Proportion	36%	53%	11%

articles somehow mentioned rSAC. Only 51 of the articles provided more in-depth discussions on the subject (i.e., the *elaborate* category which represents 53%). The fact remains, however, that these levels of information provided by the 62 articles are still insufficient for estimating which factors possibly induced the occurrence of rSAC during modeling procedures. It is worth noting that 11 (the *simple mention*) of these 62 articles only referred to the term residual spatial autocorrelation once or twice in their introductory sections. The remaining 51 articles provided more detailed and descriptive information about rSAC. Such details included the definition of rSAC, its origin, methods and suggestions on how to address it, and its quantification using Moran's *I* (Table 1). Below, we discuss five possible mechanisms or factors that potentially drive rSAC in ecological and biogeographical modeling.

Ecological data and processes

Conceptually speaking, SAC is likely to exist in any spatial data because observations from close locations are generally more related than would be expected on a random basis (Kissling and Carl 2008). The interactions between responses at these locations' zone of spatial influence result from, for example, contagious biotic processes, such as dispersal, growth, mortality, spatial diffusion, diseases, reproduction, and predation (Borcard et al. 1992; Lichstein et al. 2002). These processes can eventually create spatial patterns in species data without the influence of other external environmental data (Borcard et al. 1992). Furthermore, Kim (2013) mentioned the increase in size or a reduction of vegetation as another contagious biological process that can explain the presence of fine-scale intrinsic SAC in spatial environmental data (e.g., soil moisture). Another reason why SAC occurs in ecological data is the diffusive property across space in the movement of environmental and biotic processes, whether it be on the surface of the Earth or below the ground (Kim et al. 2016). Such environmental factors distributed continuously across the geographical space explain why, for example, species composition is similar among neighboring locations, as most species generally occupy the ranges that are greater than the cell size under study (Diniz-Filho et al. 2003). Consequently, Diniz-Filho et al. (2003) noted that using coarse scales to explain species richness would indubitably deemphasize variations at very fine scales. They suggested the use of diffusive ecological processes that act at small scales to capture information on species composition. In fact, other subsequent studies (e.g., Václavík and Meentemeyer 2009) sought to capture small-scale contagious processes leading to spatially dependent distributions and thereby violating the assumption of equilibrium between species and environmental controls (Václavík et al. 2012). These studies used multiple degrees of spatial dependence to investigate the effect of dynamic contagious processes in empirical data. Therefore, inherently, any field where such data are analyzed is subject to having to address the issue of SAC induced by contagious processes. In this context, spatial dependencies will probably show up in models using ecological data and processes (Kissling and Carl 2008; Bini et al. 2009; Crase et al. 2014). Models using spatial data are not only susceptible to having spatially autocorrelated residuals as Revermann et al. (2012) noted. In particular, working with grid data almost guarantees that SAC patterns be observed in the errors (de Oliveira et al. 2012). In some cases, this is labeled a mismatch between a process unit and an observational unit.

Scale and distance

In fact, several studies have reiterated that rSAC is closely related to distance. For Bini et al. (2009), rSAC was stronger at smaller distances in most empirical data sets. Some researchers have used terms similar to scale and distance presenting the circumstances in which model residuals show spatial dependencies. Lichstein et al. (2002) mentioned first proximity or distance and then defined the concept of appropriate neighborhood size. For these authors, distance among samples was a necessary condition for the presence of rSAC in regression models. Such patterns occurred within an "appropriate neighborhood size," or the maximum distance at which model residuals are autocorrelated. Therefore, when spatial data are analyzed, an inappropriate spatial resolution will often generate rSAC (Dormann 2007a). It is clear that more works acknowledge the type of scale as a determinant factor for rSAC. Crase et al. (2014) suggested that most of the SAC occurred at small scales (less than 1 km). It is worth pointing out that failing to account for small-scale environmental factors (Diniz-Filho et al. 2003) or only accounting for broad-scale spatial structures (Diniz-Filho and Bini 2005) will result in positive rSAC in species richness modeling at small scales. Thus, all these local-scale spatial structures (Wu and Zhang 2013) accumulated and caused spatial dependencies in the residuals of, for example, bird richness modeling (Bahn et al. 2006). Bahn et al. (2006) conceded that rSAC disappeared when using environmental predictors at large scales (> 100 km). They also admitted that the omission of important community-scale processes constituted another crucial factor of spatial dependence.

Missing variables

Variable selection is one of the characteristics that are used to compare traditional non-spatial models to spatial models which explicitly account for the presence of SAC. One explanation of the differences between non-spatial and spatial approaches in selecting variables is that non-spatial models tend to recover the missing spatial information by adding environmental variables that happen to be spatially autocorrelated (Bahn et al. 2006). In fact, failing to select relevant localized, spatially autocorrelated variables is one of the primary causes, if not the first, of rSAC. Leaving out important spatially autocorrelated predictors can directly lead to model misspecification (Bini et al. 2009; Miller 2012), which potentially generates rSAC and creates an instability associated with Lennon (2000)'s "red shift" problem (Bini et al. 2009). As supported by Bini et al. (2009), whenever such unmodeled spatially dependent predictor variables are included in the model, the degree of rSAC decreases. In contrast, when SAC is accounted for as in the case of a spatially explicit model, the relative importance likely decreases for spatially autocorrelated independent variables. Certain predictors affect the response of ecological and biogeographical processes only at local scales. Conducting broad-scale modeling will undermine such localized response variables, thus resulting in the creation of rSAC (Diniz-Filho et al. 2003). Studies suggest that failing to include important variables also causes positive rSAC, which may serve as an indicator for model misspecification (Lichstein et al. 2002; Diniz-Filho et al. 2008; Kissling and Carl 2008; Bini et al. 2009). Residual SAC is a sub-type of either exogenous or endogenous SAC. Therefore, there will be a possibility that residuals are also autocorrelated, provided that one of these two types of SAC exists in the data, as corroborated by Diniz-Filho and Bini (2005), Miller et al. (2007), Václavík et al. (2012), and Crase et al. (2014).

Sampling design

Under this "sampling design" group is considered sampling size, measurement, founder effect, sampling scheme, and sampling intensity. Each one of these factors is believed to lead to residual spatial dependencies as stated by previous studies. Bini et al. (2009) observed that a high level of rSAC is usually present in data sets with many observations. On the other hand, Lichstein et al. (2002) suggested that autocorrelated residuals can well be caused by poorly measuring an important autocorrelated predictor. In species assemblage data such as species richness and proportion of endemic species, to name a few, the sampling category is called "artifacts" in a sense that they are not due to the environment but rather from a researcher (Dormann 2007a; Crase et al. 2014). According to these authors, these artifacts are difficult to correct, and they eventually show rSAC. The artifacts are caused by species-specific bias or different recorder density. As an example, taxonomists may split plant species into more "species" than common botanists would, or a data recording team may sample one area more intensively than another, creating a bias unrelated to the environment. Additionally, a different sampling scheme would generate rSAC when regions of a known occurrence are sampled with higher intensity than regions of an unclear occurrence. Finally, ecological interactions between species (e.g., competitive exclusion and founder effects) in isolated habitat patches, such as fragmented landscapes and lakes, will add to SAC in assemblage data that are absent from individual species distribution data (Dormann 2007a; Crase et al. 2014).

Assumptions and methodological approaches

Falsely assuming linearity between two factors, using a wrong variable selection method, and ignoring the presence of non-stationarity in a data set can lead to model residuals being spatially autocorrelated. As Bini et al. (2009) noted, for example, fitting a linear model to a quadratic distribution or response would result in the residuals being spatially autocorrelated. Moreover, performing model selection requires modelers to go through several important steps including variable selection. Different approaches are used in variable selection. Le Rest et al. (2014) found that the Akaike information criteria, when used as a metric to select variables in the presence of rSAC, proved to pick up unnecessary variables to the detriment of important predictors, thereby ignoring the presence of structure in such residuals. Bini et al. (2009) defined non-stationarity as the non-consistency in the relationship between variables throughout the whole extent of the data. Non-stationarity is less intuitive and less used compared to SAC and has only lately been incorporated in SDM (Miller 2012). The concept can be viewed as the spatial variant of a constraint in correlation and regression modeling known as the Simpson's paradox (the linear trend of a sub-group is reverse of that of the overall group). It is the statistical formalization of spatial heterogeneity, which defines uneven distribution across space (like SAC, it is generally caused by sampling differences, another process in different locations of the study area or model misspecification such as missing variables). Bini et al. (2009) observed that high rSAC is usually present in data sets with high levels of non-stationarity. Similarly, Lichstein et al. (2002) argued that misspecifying a model form, such as assuming linearity when the relationship is nonlinear, may lead to spatially autocorrelated residuals. According to Wu and Zhang (2013), rSAC will probably result from linearity oversimplification. In sum, all these authors agree that residual dependencies may result from an assumption that one makes and the methodological approach that one chooses.

Conclusions

In macroecological and biogeographical modeling, multiple facets of SAC have extensively been investigated. In fact, incorporating SAC in modeling process, comparing spatial and non-spatial modeling, and identifying the potential consequences arising from the presence of spatial structure are relatively well addressed in previous studies. There seems to be a consensus that spatially explicit models in most cases outperform non-spatial models that ignore the effects of spatial dependence. However, understanding the reason why such differences in model performance exist and the circumstances under which they magnify has yet to be investigated (Crase et al. 2014; Kim et al. 2016; Miralha and Kim 2018). Most importantly, it is agreed that modeling outcomes and inferences are most affected when model residuals are spatially autocorrelated. Therefore, there has been a sense of urgency and a need to investigate rSAC in more detailed and explicit ways.

Our review of the major studies covering the topic of SAC allowed us to identify the potential sources of rSAC. In fact, a thorough review of the works reveals that the nature of the data, missing autocorrelated variables, scale, sampling design, and false methodological assumptions constitute the primary causes of SAC in model residuals. In addition to the causes of SAC, it turned out that SDM and habitat suitability modeling in birds, plants, mammals, and reptiles along with methods are the most studied topics. Despite being somewhat subjective, this categorization is an important finding, considering that it provides a better understanding of the circumstances under which model residuals are spatially autocorrelated.

The lack of quantifiable data, however, prevented us from assessing the magnitude to which rSAC is a real issue in SDM. In our review, the proportion of papers (64% including those *elaborate* and *simple mention* categories; Table 2) that mentions rSAC for the most part does so slightly and fails to contain quantitative information that would in turn allow any estimations. This review shows that rSAC in macroecological and biogeographical models are mainly intrinsic as inherent biotic processes drive the presence of spatial structure in the errors. Thus, it suggests a need for future investigations to aim at quantifying rSAC and analyzing its augmentation patterns. It is worth examining the role of missing variables, diverse sampling designs, and types of data along with model misspecification in inducing the presence of SAC in model residuals. Therefore, using combinations of such factors at multiple scales to model macroecological and biogeographical processes is strongly recommended.

Abbreviations

OLS: Ordinary least squares; rSAC: Residual spatial autocorrelation; SAC: Spatial autocorrelation; SDM: Species distribution modeling

Acknowledgements

The authors are particularly grateful to Heejun Chang and his collaborator Janardan Mainali for their informative comments, suggestions, and discussions on this review.

Funding

This research was supported by (1) the National Science Foundation (grant numbers 0825753 and 1560907), (2) the National Research Foundation of Korea (NRF-2017R1C1B5076922), and (3) the Research Resettlement Fund for the new faculty of Seoul National University.

Availability of data and materials

This is a review article, so we did not analyze any data.

Authors' contributions

GG and DK conceived the work, performed literature review, and wrote the manuscript. YC provided critical advice on statistical issues. All the authors approved the submission of the manuscript in its current form.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

[1]Department of Geography, University of Kentucky, Lexington, KY, USA. [2]Department of Geography, Seoul National University, Seoul, South Korea. [3]Institute for Korean Regional Studies, Seoul National University, Seoul, South Korea. [4]University of Texas at Dallas, Dallas, TX, USA.

Received: 11 March 2019 Accepted: 21 April 2019
Published online: 14 May 2019

References

Ali GA, Roy AG, Legendre P. Spatial relationships between soil moisture patterns and topographic variables at multiple scales in a humid temperate forested catchment. Water Resour Res. 2010;46:10.

Anselin L. Under the hood: issues in the specification and interpretation of spatial regression models. Agric Econ. 2002;27:247–67.

Anselin L, Bera AK. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah A, Giles DEA, editors. Handbook of applied economic statistics. New York: Marcel Dekker; 1998. p. 237–89.

Anselin L, Syabri I, Kho Y. GeoDa: an introduction to spatial data. Geogr Anal. 2006;38:5–22.

Augustin NH, Cummins RP, French DD. Exploring spatial vegetation dynamics using logistic regression and a multinomial logit model. J Appl Ecol. 2001;38:991–1006.

Austin MP. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecol Model. 2002;157:101–18.

Bahn V, O'Connor RJ, Krohn WB. Importance of spatial autocorrelation in modeling bird distributions at a continental scale. Ecography. 2006;29:835–44.

Betts MG, Diamond AW, Forbes GJ, Villard M-A, Gunn JS. The importance of spatial autocorrelation, extent and resolution in predicting forest bird occurrence. Ecol Model. 2006;191:197–224.

Bini L, Alexandre J, Diniz-Filho F, TFLVB R, TSB A, Albaladejo RG, Albuquerque FS, Aparicio A, Araújo MB, Baselga A, Beck J, Bellocq MI, Böhning-Gaese K, PAV B, Castro-Parga I, Chey VK, Chown SL, de Marco P Jr, Dobkin DS, Ferrer-Castán D, Field R, Filloy J, Fleishman E, Gómez JF, Hortal J, Iverson JB, Kerr JT, Kissling WD, Kitching IJ, León-Cortés JL, Lobo JM, Montoya D, Morales-Castilla I, Moreno JC, Oberdorff T, Olalla-Tárraga MÁ, Pausas JG, Qian H, Rahbek C, Rodríguez MÁ, Rueda M, Ruggiero A, Sackmann P, Sanders NJ, Terribile LC, Vetaas OR, Hawkins BA. Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. Ecography. 2009;32:193–204.

Bonada N, Dolédec S, Statzner B. Spatial autocorrelation patterns of stream invertebrates: exogenous and endogenous factors. J Biogeogr. 2012;39:56–68.

Borcard D, Legendre P, Drapeau P. Partialling out the spatial component of ecological variation. Ecology. 1992;73:1045–55.

Büchi L, Christin PA, Hirzel AH. The influence of environmental structure on the life-history traits and diversity of species in a metacommunity. Ecol Model. 2009;220:2857–64.

Carl G, Kühn I. Analyzing spatial autocorrelation in species distributions using Gaussian and logit models. Ecol Model. 2007;207:159–70.

Chang J, Chen D, Ye X, Li S, Liang W, Zhang Z, Li M. Coupling genetic and species distribution models to examine the response of the Hainan partridge (*Arborophila ardens*) to late Quaternary climate. PLoS One. 2012. https://doi.org/10.1371/journal.pone.0050286.

Chun Y, Griffith DA. Modeling network autocorrelation in space–time migration flow data: an eigenvector spatial filtering approach. Ann Assoc Am Geogr. 2011;101:523–36.

Ciccarelli D, Bacaro G. Quantifying plant species diversity in coastal dunes: a piece of help from spatially constrained rarefaction. Folia Geobot. 2016;51:129–41.

Cliff N. An improved internal consistency reliability estimate. J Educ Behav Stat. 1984;9:151–61.

Crase B, Liedloff A, Vesk PA, Fukuda Y, Wintle BA. Incorporating spatial autocorrelation into species distribution models alters forecasts of climate-mediated range shifts. Glob Change Biol. 2014;20:2566–79.

Crase B, Liedloff A, Wintle BA. A new method for dealing with residual spatial autocorrelation in species distribution models. Ecography. 2012;35:879–88.

Davis AJS, Singh KK, Thill J, Meentemeyer RK. Accounting for residential propagule pressure improves prediction of urban plant invasion. Ecosphere. 2016. https://doi.org/10.1002/ecs2.1232.

de Oliveira G, Araújo MB, Rangel TF, Alagador D, Diniz-Filho JAF. Conserving the Brazilian semiarid (Caattinga) biome under climate change. Biodivers Conserv. 2012;21:2913–26.

de Oliveira G, Rangel TF, Lima-Ribeiro MS, Terribile LC, JAF D-F. Evaluating, partitioning, and mapping the spatial autocorrelation component in ecological niche modeling: a new approach based on environmentally equidistant records. Ecography. 2014;37:637–47.

Diniz-Filho JAF, Bini LM. Modelling geographical patterns in species richness using eigenvector-based spatial filters. Glob Ecol Biogeogr. 2005;14:177–85.

Diniz-Filho JAF, Bini LM, Hawkins BA. Spatial autocorrelation and red herrings in geographical ecology. Glob Ecol Biogeogr. 2003;12:53–64.

Diniz-Filho JAF, Rangel TFLVB, Bini LM. Model selection and information theory in geographical ecology. Glob Ecol Biogeogr. 2008;17:479–88.

Dirnböck T, Dullinger S. Habitat distribution models, spatial autocorrelation, functional traits and dispersal capacity of alpine plant species. J Veg Sci. 2004;15:77–84.

Dorken ME, Freckleton RP, Pannell JR. Small-scale and regional spatial dynamics of an annual plant with contrasting sexual systems. J Ecol. 2017;105:1044–57.

Dormann C. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. Glob Ecol Biogeogr. 2007a;16:129–38.

Dormann C. Assessing the validity of autologistic regression. Ecol Model. 2007b; 207:234–42.

Dowd M, Grant J, Lu L. Predictive modeling of marine benthic macrofauna and its use to inform spatial monitoring design. Ecol Appl. 2014;24:862–76.

Dronova I, Beissinger SR, Burnham JW, Gong P. Landscape-level associations of wintering waterbird diversity and abundance from remotely sensed wetland characteristics of Poyang lake. Remote Sens-Basel. 2016. https://doi.org/10.3390/rs8060462.

Ennen JR, Agha M, Matamoros WA, Hazzard SC, Lovich JE. Using climate, energy, and spatial-based hypotheses to interpret macroecological patterns of North America chelonians. Can J Zool. 2016;94:453–61.

Epperson BK. Spatial and space–time correlations in ecological models. Ecol Model. 2000;132:63–76.

Estrada A, Delgado MP, Arroyo B, Traba J, Morales MB. Forecasting large-scale habitat suitability of European bustards under climate change: the role of environmental and geographic variables. PLoS One. 2016. https://doi.org/10.1371/journal.pone.0149810.

Estrada CG, Rodriguez-Estrella R. In the search of good biodiversity surrogates: are raptors poor indicators in the Baja California Peninsula desert? Anim Conserv. 2016;19:360–8.

Ficetola GF, Manenti R, De Bernard F, Padoa-Schioppa E. Can patterns of spatial autocorrelation reveal population processes? An analysis with the fire salamander. Ecography. 2012;35:693–703.

Getis A. A history of the concept of spatial autocorrelation: a geographer's perspective. Geogr Anal. 2008;40:297–309.

Griffith DA. A linear regression solution to the spatial autocorrelation problem. J Geogr Syst. 2000;2:141–56.

Griffith DA, Peres-Neto PR. Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. Ecology. 2006;87:2603–13.

Guénard G, Lanthier G, Harvey-Lavoie S, Macnaughton CJ, Senay C, Lapointe M, Legendre P, Boisclair D. A spatially- explicit assessment of the fish population response to flow management in a heterogeneous landscape Guillaume. Ecosphere. 2016. https://doi.org/10.1002/ecs2.1252.

Güler B, Jentsch A, Apostolova I, Bartha S, Bloor JMG, Campetella G, Canullo R, Házi J, Kreyling J, Pottier J, Szabó G, Terziyska T, Uğurlu E, Wellstein C, Zimmermann Z, Dengler J. How plot shape and spatial arrangement affect plant species richness counts: implications for sampling design and rarefaction analyses. J Veg Sci. 2016;27:692–703.

Gwenzi D, Lefsky MA. Spatial modeling of Lidar-derived woody biomass estimates collected along transects in a heterogeneous savanna landscape. IEEE J Sel Top Appl. 2017;10:372–84.

Hawkins BA, Diniz-Filho JAF, Bini LM, De Marco P, Blackburn TM. Red herrings revisited: spatial autocorrelation and parameter estimation in geographical ecology. Ecography. 2007;30:375–84.

Hefley TJ, Broms KM, Brost BM. The basis function approach for modeling autocorrelation in ecological data. Ecology. 2017a;98:632–46.

Hefley TJ, Hooten MB, Russell RE, Walsh DP, Powell JA. When mechanism matters: Bayesian forecasting using models of ecological diffusion. Ecol Lett. 2017b;20:640–50.

Hindrikson M, Remm J, Pilot M, Godinho R, Stronen AV, Baltrūnaité L, Czarnomska SD, Leonard JA, Randi E, Nowak C, Åkesson M, López-Bao JV, Álvares F, Llaneza L, Echegaray J, Vilà C, Ozolins J, Rungis D, Aspi J, Paule L, Skrbinšek T, Saarma U. Wolf population genetics in Europe: a systematic review, meta-analysis and suggestions for conservation and management. Biol Rev. 2017;92:1601–29.

Hongoh V, Berrang-Ford L, Scott ME, Lindsay LR. Expanding geographical distribution of the mosquito, *Culex pipiens*, in Canada under climate change. Appl Geogr. 2012;33:53–62.

Ingberman B, Fusco-Costa R, de Araujo Monteiro-Filho EL. A current perspective on the historical geographic distribution of the endangered Muriquis (*Brachyteles* spp.): Implications for conservation. PLOS ONE. 2016. https://doi.org/10.1371/journal.pone.0150906.

Ishihama F, Takeda T, Oguma H, Takenaka A. Comparison of effects of spatial autocorelation on distribution predictions of four rare plant species in the Watarase wetland. Ecol Res. 2010;25:1057–69.

Jackson MM, Gergel SE, Martin K. Citizen science and field survey observations provide comparable results for mapping Vancouver Island white-tailed ptarmigan (*Lagopus Leucura saxatillis*) distributions. Biol Conserv. 2015;181:162–72.

Kim D. Incorporation of multi-scale spatial autocorrelation in soil moisture–landscape modeling. Phys Geogr. 2013;34:441–55.

Kim D. Modeling spatial and temporal dynamics of plant species richness across tidal creeks in a temperate salt marsh. Ecol Indic. 2018;93:188–95.

Kim D, Hirmas DR, McEwan RW, Mueller TG, Park SJ, Šamonil P, Thompson JA, Wendroth O. Predicting the influence of multi-scale spatial autocorrelation on soil–landform modeling. Soil Sci Soc Am J. 2016;80:409–19.

Kim D, Shin Y. Spatial autocorrelation potentially indicates the degree of changes in the predictive power of environmental factors for plant diversity. Ecol Indic. 2016;60:1130–41.

Kissling WD, Carl G. Spatial autocorrelation and the selection of simultaneous autoregressive models. Glob Ecol Biogeogr. 2008;17:59–71.

Kleisner KM, Walter JF III, Diamond SL, Die DJ. Modeling the spatial autocorrelation of pelagic fish abundance. Mar Ecol Prog Ser. 2010;411:203–13.

Komac B, Esteban P, Trapero L, Caritg R. Modelization of the current and future habitat suitability of *Rhododendron ferrugineum* using potential snow accumulation. PLoS One. 2016. https://doi.org/10.1371/journal.pone.0147324.

Kühn I. Incorporating spatial autocorrelation may invert observed patterns. Divers Distrib. 2007;13:66–9.

Le Rest K, Pinaud D, Monestiez P, Chadoeuf J, Bretagnolle V. Spatial leave-one-out cross-validation for variable selection in the presence of spatial autocorrelation. Glob Ecol Biogeogr. 2014;23:811–20.

Legendre P. Spatial autocorrelation: trouble or new paradigm? Ecology. 1993;74:1659–73.
Lennon JJ. Red-shifts and red herrings in geographical ecology. Ecography. 2000;23:101–13.
Lichstein JW, Simons TR, Shriner SA, Franzreb KE. Spatial autocorrelation and autoregressive models in ecology. Ecol Monogr. 2002;72:445–63.
Lloyd NJ, Nally RM, Lake PS. Spatial autocorrelation of assemblages of benthic invertebrates and its relationship to environmental factors in two upland rivers in southeastern Australia. Divers Distrib. 2005;11:375–86.
Marmion M, Luoto M, Heikkinen RK, Thuiller W. The performance of state-of-the-art modelling techniques depends on geographical distribution of species. Ecol Model. 2009;220:3512–20.
Mattsson BJ, Zipkin EF, Gardner B, Blank PJ, Sauer JR, Royle JA. Explaining local-scale species distributions: relative contributions of spatial autocorrelation and landscape heterogeneity for an avian assemblage. PLoS One. 2013. https://doi.org/10.1371/journal.pone.0055097.
Merckx B, Goethals P, Steyaert M, Vanreusel A, Vincx M, Vanaverbeke J. Predictability of marine nematode biodiversity. Ecol Model. 2009;220:1449–58.
Mets KD, Armenteras D, Dávalos LM. Spatial autocorrelation reduces model precision and predictive power in deforestation analyses. Ecosphere. 2017. https://doi.org/10.1002/ecs2.1824.
Miller J, Franklin J, Aspinall R. Incorporating spatial dependence in predictive vegetation models. Ecol Model. 2007;202:225–42.
Miller JA. Species distribution models: spatial autocorrelation and non-stationarity. Prog Phys Geogr. 2012;36:681–92.
Miralha L, Kim D. Accounting for and predicting the influence of spatial autocorrelation in water quality modeling. ISPRS Int J Geo-Inf. 2018. https://doi.org/10.3390/ijgi7020064.
Naimi B, Skidmore AK, Groen TA, Hamm NAS. Spatial autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling. J Biogeogr. 2011;38:1497–509.
Nicolaus M, Brommer JE, Ubels R, Tinbergen M, Dingemanse NJ. Exploring patterns of variation in clutch size–density reaction norms in a wild passerine bird. J Evolution Biol. 2013;26:2031–43.
Ortiz-Yusty CE, Páez V, Zapata FA. Temperature and precipitation as predictors of species richness in northern Andean amphibian from Colombia. Caldasia. 2013;35:65–80.
Piazzini S, Caruso T, Favilli L, Favilli L, Manganelli G. Role of predators, habitat attributes, and spatial autocorrelation on the distribution of eggs in the northern spectacled salamander (*Salamandrina perspicillata*). J Herpetol. 2011;45:389–94.
Pickup G, Chewings VH. Random field modeling of spatial variations in erosion and deposition in flat alluvial landscapes in arid Central Australia. Ecol Model. 1986;33:269–96.
Platts PJ, McClean CJ, Lovett JC, Marchant R. Predicting three distributions in an east African biodiversity hotspot: model selection, data bias and envelope uncertainty. Ecol Model. 2008;218:121–34.
Poley LG, Pond BA, Schaefer JA, Brown GS, Ray JC, Johnson DS. Occupancy patterns of large mammals in the far north of Ontario under imperfect detection and spatial autocorrelation. J Biogeogr. 2014;41:122–32.
Record S, Charney ND, Zakaria RM, Ellison AM. Projecting global mangrove species and community distributions under climate change. Ecosphere. 2013b. https://doi.org/10.1890/ES12-00296.1.
Record S, Fitzpatrick MC, Finley AO, Veloz S, Ellison AM. Should species distribution models account for spatial autocorrelation? A test of model projections across eight millennia of climate change. Glob Ecol Biogeogr. 2013a;22:760–71.
Revermann R, Schmid H, Zbinden N, Spaar R, Schröder B. Habitat at the mountain tops: how long can rock Ptarmigan (*Lagopus muta helvetica*) survive rapid climate change in the Swiss Alps? A multi-scale approach. J Ornithol. 2012;153:891–905.
Rodriguez A, Gómez JF, Nieves-Aldrey JL. Modeling the potential distribution and conservation status of three species of oak gall wasps (Hymenoptera: Cynipidae) in the Iberian range. J Insect Conserv. 2015;19:921–34.
Roth T, Bühler C, Amrhein V. Estimating effects of species interactions on populations of endangered species. Am Nat. 2016;187:457–67.
Santos SM, Mira AP, Mathias ML. Factors influencing large-scale distribution of two sister species of pine voles (*Microtus lusitanicus* and *Microtus duodecimcostatus*): the importance of spatial autocorrelation. Can J Zool. 2009;87:1227–40.
Seymour L. Spatial data analysis: theory and practice. J Am Stat Assoc. 2005;100:353.
Sheehan KL, Esswein ST, Dorr BS, Yarrow GK, Johnson RJ. Using species distribution models to define nesting habitat of the eastern metapopulation of double-crested cormorants. Ecol Evol. 2017;7:409–18.
Siderov K. Spatial data analysis: theory and practice. Austral Ecol. 2005;30:237–41.
Siesa ME, Manenti R, Padoa-Schioppa E, de Bernardi F, Ficetola GF. Spatial autocorrelation and the analysis of invasion processes from distribution data: a study with the crayfish *Procambarus clarkia*. Biol Invasions. 2011;13:2147–60.
Tallowin O, Allison A, Algar AC, Kraus F, Meiri S. Papua New Guinea terrestrial-vertebrate richness: elevation matters most for all except reptiles. J Biogeogr. 2017;44:1734–44.
Tarkhnishvili D, Gavashelishvili A, Mumladze L. Palaeoclimatic models help to understand current distribution of Caucasian forest species. Biol J Linn Soc. 2012;105:231–48.
Václavík T, Kupfer JA, Meentemeyer RK. Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). J Biogeogr. 2012;39:42–55.
Václavík T, Meentemeyer RK. Invasive species distribution modeling (iSDM): are absence data and dispersal constraints needed to predict actual distributions? Ecol Model. 2009;220:3248–58.
Veloz SD. Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. J Biogeogr. 2009;36:2290–9.
Warren DL, Cardillo M, Rosauer DF, Bolnick DI. Mistaking geography for biology: inferring processing from species distributions. Trends Ecol Evol. 2014;29:572–80.
Weeks AM, de Jager NR, Haro RJ, Sandland GJ. Spatial and temporal relationships between the invasive snail *Bithynia tentaculata* and submersed aquatic vegetation in Pool 8 of the upper Mississippi River. River Res Appl. 2017;33:729–39.
Wieczorek K, Bugaj-Nawrocka A. Invasive aphids of the tribe Siphini: a model of potentially suitable ecological niches. Agr Forest Entomol. 2014;16:434–43.
Wiegand T, Moloney KA. Rings, circles, and null-models for point pattern analysis in ecology. Oikos. 2004;104:209–29.
Wu D, Liu J, Zhang G, Ding W, Wang W, Wang R. Incorporating spatial autocorrelation in cellular automata model: an application to the dynamics of Chinese tamarisk (*Tamarisk chinensis* Lour.). Ecol Model. 2009;220:3490–8.
Wu W, Zhang L. Comparison of spatial and non-spatial logistic regression models for modeling the occurrence of cloud cover in north-eastern Puerto Rico. Appl Geogr. 2013;37:52–62.
Wulder MA, White JC, Coops NC, Nelson T, Boots B. Using spatial autocorrelation to compare outputs from a forest growth model. Ecol Model. 2007;209:264–76.
Yu MH. Modeling tree growth and seedling recruitment in a selectively logged temperature forest. The City University of New York: PhD Dissertation; 2012.
Zhang L, Ma Z, Guo L. An evaluation of spatial autocorrelation and heterogeneity in the residuals of six regression models. For Sci. 2009;55:533–48.
Zhu W, Jia S, Lü A, Yan T. Analyzing and modeling the coverage of vegetation in the Qaidam basin of China: the role of spatial autocorrelation. J Geogr Sci. 2012;22:346–58.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

BMC