

University of Kentucky

UKnowledge

Theses and Dissertations--Geography

Geography


2021

RESIDUAL SPATIAL AUTOCORRELATION IN MACROECOLOGICAL AND BIOGEOGRAPHICAL MODELING: A REVIEW

Guetchine Gaspard

University of Kentucky, ggansfsac2016@gmail.com

Author ORCID Identifier:

 <https://orcid.org/0000-0001-8806-9771>

Digital Object Identifier: <https://doi.org/10.13023/etd.2021.164>

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Recommended Citation

Gaspard, Guetchine, "RESIDUAL SPATIAL AUTOCORRELATION IN MACROECOLOGICAL AND BIOGEOGRAPHICAL MODELING: A REVIEW" (2021). *Theses and Dissertations--Geography*. 75. https://uknowledge.uky.edu/geography_etds/75

This Master's Thesis is brought to you for free and open access by the Geography at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Geography by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Guetchine Gaspard, Student

Dr. J. Anthony Stallins, Major Professor

Dr. Matthew Zook, Director of Graduate Studies

RESIDUAL SPATIAL AUTOCORRELATION
IN MACROECOLOGICAL AND BIOGEOGRAPHICAL MODELING: A REVIEW

THESIS

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Arts in the
College of Arts and Sciences
at the University of Kentucky

By

Guetchine Gaspard

Lexington, Kentucky

Director: Dr. J. Anthony Stallins, Professor of Geography

Lexington, Kentucky

2021

Copyright © Guetchine Gaspard 2021
<https://orcid.org/0000-0001-8806-9771>

ABSTRACT OF THESIS

RESIDUAL SPATIAL AUTOCORRELATION IN MACROECOLOGICAL AND BIOGEOGRAPHICAL MODELING: A REVIEW

Macroecological and biogeographical modelers have predicted the distribution of species across space relying on the relationship between biotic processes and environmental variables. Such a method employs data associated, for instance, with species abundance or presence/absence, climate, geomorphology, and soils. Statistical analyses found in previous studies have highlighted the importance of accounting for the effects of spatial autocorrelation (SAC), which indicates a level of dependence between pairs of nearby observations. A consensus has existed that residual spatial autocorrelation (rSAC) can substantially impact modeling processes and inferences. However, more emphasis should be put on identifying the sources of rSAC and the degree to which rSAC becomes detrimental. In this thesis, we review previous studies to identify various factors that potentially engender the presence of rSAC in macroecological and biogeographical models. Additionally, special attention is paid to the quantification of rSAC by attempting to bring out the magnitude to which the presence of SAC in model residuals impedes the modeling process. The review identified that five categories of factors potentially drive the presence of SAC in model residuals: the type of ecological data and the processes underlying it, scale and distance, missing variables, sampling design, as well as the assumptions and methodological perspectives of the investigator. Furthermore, we concluded that more explicit discussion of rSAC should be carried out in species distribution modeling. We recommend further investigations involving the quantification of rSAC to understand when rSAC can have a negative effect on the modeling process.

KEYWORDS: Spatial autocorrelation, Residual Spatial Autocorrelation, Missing Variables, Sampling Design, Scale, Species Distribution Models

Guetchine Gaspard

04/30/2021

Date

RESIDUAL SPATIAL AUTOCORRELATION
IN MACROECOLOGICAL AND BIOGEOGRAPHICAL MODELING: A REVIEW

By
Guetchine Gaspard

J. Anthony Stallins

Director of Thesis

Matthew Zook

Director of Graduate Studies

04/30/2021

Date

DEDICATION

I dedicate this work to all who encouraged me one way or another and wanted to see me successfully finish my studies. However, I will allow myself to offer this work particularly to my daughter Lynn Kiara Gaspard who was with me night in and night out, who despite being a child encouraged me.

ACKNOWLEDGEMENTS

This thesis, while an individual achievement, is the product of the contribution from multiple people. First, my original Chair, Dr. Daehyun Kim provided the initial opportunity to pursue this research. My second Chair, Dr. J. Anthony Stallins provided support during my journey as a student. He was very supportive, showed understanding and patience which allowed me to thrive during the ups and downs of my graduate student life. In addition, Dr. Stallins always made sure that I was on track every step of the way so that I could complete the thesis in a timely manner. Next, I want to thank the other members on my Committee namely Dr. Alice Turkington, Dr. Ole Wendroth, and Dr. Liang Liang each of whom provided valuable insights that helped complete this work.

This thesis emanates from my early research on the effects of spatial autocorrelation in regression modeling as part of a project funded by the National Science Foundation under the direction of Dr. Daehyun Kim. The Research was conducted mainly at the University of Kentucky, University of Texas at Dallas, and Portland State University. Also, I want to thank Dr. Yongwan Chun who also provided valuable insight to the original published article.

Furthermore, I am grateful to the Department of Geography for hosting and assisting me with all the resources necessary to completing my program. Lastly, I am very appreciative of the support and motivation that I received from family and friends throughout the whole process.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
LIST OF TABLES	v
CHAPTER 1. INTRODUCTION	1
1.1 Background.....	1
1.2 Objectives	2
1.3 Conceptual framework.....	2
1.4 Justification.....	4
CHAPTER 2. Methods and DATA	5
2.1 Selection of articles.....	5
2.2 Spatial autocorrelation in the articles.....	5
2.3 Sources of SAC.....	5
CHAPTER 3. RESULTS and DISCUSSION.....	7
3.1 Subjects and species addressed.....	7
3.2 Ecological data and processes.....	9
3.3 Scale and distance	11
3.4 Missing Variables	12
3.5 Sampling Design.....	13
3.6 Assumptions and methodological approaches	14
CHAPTER 4. CONCLUSIONS	23
BIBLIOGRAPHY.....	25
VITA.....	31

LIST OF TABLES

TABLE 3. 1 LITERATURE REVIEW IN MACROECOLOGICAL AND BIOGEOGRAPHICAL MODELING. SAC SPATIAL AUTOCORRELATION, RSAC RESIDUAL SPATIAL AUTOCORRELATION	16
--	----

CHAPTER 1. INTRODUCTION

1.1 Background

Using spatial or geographical data involves learning about the properties of such data. Fields such as geography, ecology that use geographic data, where space and time matter, remain concerned with how such data are characterized. The presence of structure or dependence among the observations is one of the most common issues that is associated with spatial data. Frequently, processes be it environmental or biological, are related across space and time. This fact reverts to the notion of distance decay wherein the degree of dependence decreases over space. That was 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 reasoning can be attributed to the concept of spatial autocorrelation (SAC) which was introduced around the late 1960s and early 1970s (Getis, 2008) and which 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: 1659).

Contingent upon the variables that drive natural processes, SAC is categorized into two types: exogenous and endogenous SAC (Legendre, 1993). The former is driven 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). Usually, it is associated with broad-scale spatial trends (Miller et al., 2007; Václavík et al., 2012). Endogenous SAC, however, is caused 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 emphasizes the contagion effects in cases of positive autocorrelation or the dispersion effects for negative autocorrelation (Lichstein et al., 2002; Griffith and Peres-Neto, 2006; Crase et al., 2014). Such intrinsic SAC is prominent at fine scales or to high-resolution stochastic biotic processes (Dormann, 2007a ; Miller et al., 2007; Chun and Griffith, 2011; Václavík et al., 2012). The following sections state the scope and relevance of the study and provide further insight on the concept of residual spatial autocorrelation, hereafter, rSAC.

1.2 Objectives

The purpose of this review is to determine the circumstances in which the magnitude of residual spatial autocorrelation increases in species distribution modeling (SDM). More specifically, we sought to answer the following research questions:

1. What are the major sources of rSAC?
2. How much do missing variable explain rSAC?
3. How do various sampling schemes affect the level of structure in model residuals?

1.3 Conceptual framework

Understanding rSAC remains a big issue in the field of ecological modeling. In a modeling context, residuals represent the differences between observed values and predicted values. Hence, rSAC indicates the amount of SAC present in the variance that is not explained by the independent variables. Understanding the distribution of residuals is

critical to performing regression modeling analysis, as assumptions such as linearity, normality, homoscedasticity (equal variance), and independence rely on the behavior of the error terms. 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); therefore, the use of such data generally violates the assumption of independence between pairs of observations, demanding that the effects of rSAC be accounted for (Diniz-Filho and Bini, 2005; Bahn et al., 2006).

Integrating or leaving out rSAC has implications that directly affect the outcomes of species distribution modeling (SDM). Failing to adequately address rSAC will eventually lead to three major statistical problems. First, the standard errors might well be underestimated, leading to what is known as Type I error. This simply means that the presence of dependence between pairs of observations across space, where independence between such observations is assumed, can result in falsely rejecting, much more often than expected, the null hypothesis while it is true (Lennon, 2000). Consequently, that will render the regression model itself unreliable (Legendre, 1993; Anselin, 2002; Kim et al., 2016). Second, parameter estimates, namely the regression coefficients, might be biased (Dormann, 2007a; Václavík et al., 2012). The inflation or deflation of predictors' coefficients will lead to the over- or under-estimation of their predictive power, respectively. Lastly, model misspecification, a critical component of variable selection, remains an important problem (Austin, 2002; Lichstein et al., 2002; Miller et al., 2007; Václavík et al., 2012).

1.4 Justification

The notion of SAC is extensively discussed in biogeography and macroecology literature. However, those studies have not taken a systematic look at the contexts and factors that contribute to rSAC. Previous researchers suspect that failing to incorporate certain independent predictors might be the main problem (Crase et al. 2014). The authors suggest that this problem, when associated with the intrinsic rather than the extrinsic type of SAC, remains unexplored. Identifying potential missing variables and establishing how much their omission increases the level of rSAC would generate new knowledge and add to the SDM literature body. In addition to the environmental and biotic missing variables, the type of sampling design should also be scrutinized since the latter is often mentioned as having the ability to increase rSAC (Lichstein et al. 2002; Bini et al. 2009; Crase et al. 2014). This thesis addresses sampling design with respect to sample size, data type, sampling technique, and the effect of small scales. Analyzing data at very fine scales coupled with the inclusion of important spatially autocorrelated missing variables is thought to have the potential to significantly reduce or even remove rSAC in species distribution models. Diniz-Filho et al. (2003) suggest that including relevant environmental factors that act at each scale in a regression model would eventually remove SAC from the residuals at different scales, under the assumption that environmental factors behave differently at distinct spatial scales.

The bottom line is that by conducting this investigation, we expected to: (1) provide a holistic understanding of rSAC across the existing literature of macroecological and biogeographical modeling and (2) lay a foundation to conduct further research on rSAC.

CHAPTER 2. METHODS AND DATA

2.1 Selection of articles

The purpose of this step was to gather the necessary literature to meet the objectives set forth in the review. Initially, we targeted published peer-reviewed articles from the fields of biogeography and macroecology that dealt with SDM and in which SAC was explicitly incorporated. For the actual search, we used keywords such as residual spatial autocorrelation, spatial autocorrelation, ecological or biogeographical as well as species distribution modeling to acquire relevant articles via the Web of Science and Google Scholar search engines. To complete the list, we also selected articles cited or referenced in the original selections.

2.2 Spatial autocorrelation in the articles

From the results of the search, we determined the degree to which each article discusses the concept of rSAC or SAC more broadly. The articles were carefully reviewed and then grouped based on the level of detail they provided about rSAC. To achieve this categorization, we used following scale as metric: *no mention* in cases the article does not mention rSAC, *simple mention*, in the event that concept is loosely mentioned or discussed in the article, and *elaborate* in case the topic of rSAC is well discussed by the paper.

2.3 Sources of SAC

Finally, we meticulously reviewed each article to find out which factor or the circumstance that study mentioned or identified as a potential source of SAC in model errors. By repeating this process across all the articles, we were able to group the sources

into larger categories, which was the main goal of our review. In the end, we attempted to understand the conditions under which SAC occurred—and magnified—in model residuals. The findings and their interpretation and discussion are presented in Chapter 3.

CHAPTER 3. RESULTS AND DISCUSSION

3.1 Subjects and species addressed

We ended up selecting 97 articles dating from 1984 to 2017 (Table 2.1). Then, we reviewed the selected papers in relation to the concept of SAC. The review of the existing literature revealed that accounting for SAC in SDM is still in an early stage, despite studies having increasingly attempted to widely incorporate the effect of spatial dependence in investigating ecological and biogeographical processes over the course of the last thirty years. The results indicated that only a small proportion (less than 20%) of ecological and biogeographical modelers incorporated SAC in their research. This is partly attributed to the fact that the need to incorporate SAC is still contentious among modelers (Diniz-Filho et al. 2003; Hawkins et al. 2007; Bini et al. 2009; Miller 2012). The presence of SAC in ecological and biogeographical data has long been detected (since around the late 1970s), and statistical methods capable of addressing it were developed almost in the same period (Dormann 2007a). Legendre (1993) defined and categorized the concept of SAC into endogenous and exogenous SAC in the field of ecological data modeling. However, modelers did not start substantially publishing studies that incorporate SAC until after 2000.

Species distribution modeling stood out 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 are niche suitability, habitat distribution, climatic suitability, climatic forecast, or predictability.

This finding aligns with the fact that 92 out of the 97 articles we reviewed were published in the new millennium. Some of the early works that acknowledged the effect of SAC before 2000 include, but are not limited to, Borcard et al (1992) who sought to partition the total variance of species abundance into spatial and non-spatial components, and Pickup and Chewings (1986) who worked on the prediction of erosion and deposition in alluvial landscapes of central Australia.

Reading these discussions about the context of the current literature shows why rSAC, as a subcategory of SAC, remains relatively unexplored in ecological and biogeographical modeling. We divided the articles into three groups (i.e., *no mention*, *simple mention*, and *elaborate*) based on the level of details being provided from the discussion on rSAC (Table 2.1). We found that 35 articles (36%) never mentioned the presence or influence of rSAC. Of the remaining 62 (*simple mention* plus *elaborate*) articles 51 of them provided more in-depth discussions on the topic (i.e., the *elaborate*

category which represents 53%). Yet the levels of information found in the 62 articles are still insufficient for quantifying which factors possibly caused the occurrence of rSAC during the modeling processes. It is worth pointing out that 11 (the *simple mention*) of these 62 articles only mentioned the term residual spatial autocorrelation once or twice in their introductions. As far as the remaining 51 articles were concerned, they 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 3.1). In the following sections, we discuss five possible mechanisms or factors that potentially dictate rSAC in ecological and biogeographical modeling.

3.2 Ecological data and processes

Theoretically speaking, SAC is likely to exist in any spatial data because observations from nearby locations are normally more related than would be expected on a random basis (Kissling et al., 2008). The exchange between responses at these locations' zone of spatial influence results 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 underlying processes can eventually create spatial patterns in species data without the influence of other external environmental data (Borcard et al., 1992). Moreover, Kim et al. (2013) mentioned the increase in size or a reduction of vegetation as being another contagious biological process capable of explaining the presence of fine-scale intrinsic SAC in spatial environmental data.

Furthermore, SAC occurs in ecological data due to 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). These environmental factors distributed continuously across the geographical area explain why, for instance, species composition remains the same among neighboring locations, as most species generally occupy the ranges that are greater than the cell size under study (Diniz-Filho et al., 2003). As a consequence, Diniz-Filho et al. (2003) suggested that using coarse scales to explain species richness would certainly deemphasize variations at very fine scales. The authors suggested the use of diffusive ecological processes that are effective at small scales to capture information on species composition. Later, Václavík and Meentemeyer (2009) sought to capture small-scale contagious processes that lead to spatially dependent distributions and thereby violating the assumption of equilibrium between species and environmental controls (Václavík et al., 2012). Both works used multiple levels of spatial dependence to investigate the effect of dynamic contagious processes in empirical data. Inherently, any discipline where these data are analyzed is bound to address the issue of SAC generated by diffusive processes. Thus, spatial dependencies will likely appear in models that use ecological data and processes (Kissling et al., 2008; Bini et al., 2009; Crase et al., 2014: 2467). Models that use spatial data are not susceptible to having spatially autocorrelated residuals only, as Reverman et al. (2012) noted. Using grid data almost guarantees that SAC patterns will be observed in the residuals (Oliveira et al., 2012). Sometimes, this is labeled a mismatch between a process unit and an observational unit.

3.3 Scale and distance

Several studies have reiterated that rSAC is highly associated with distance. According to Bini et al. (2009: 196), rSAC was stronger at smaller distances in most empirical datasets. Certain researchers have used terms related to scale and distance to account for the circumstances in which model residuals show spatial autocorrelation. As for Lichstein et al. (2002: 449), they mentioned first proximity or distance and then defined the concept of appropriate neighborhood size. According to the 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. Consequently, when spatial data are analyzed, an inappropriate spatial scale will often produce rSAC (Dormann, 2007a). An increasing number of studies acknowledge that scale extent is a contributing factor for rSAC. Crase et al. (2014) found that most of the SAC occurred at small scales (less than 1 km). As it pertains to small scales, it is worth mentioning that failing to account for small-scale environmental factors (Diniz-Filho et al., 2003) or only accounting for broad-scale spatial dependencies (Diniz-Filho et al., 2005) will create positive rSAC in species richness at small scales. Thus, all these local-scale spatial structures (Wu and Zhang, 2013) accumulated and caused spatial autocorrelation in the residuals (Bahn et al., 2006). Barn et al. (2006) suggested that rSAC disappeared when using environmental predictors at large scales (> 100 km). The researchers also admitted that the omission of important community-scale processes constituted another crucial factor of spatial dependence.

3.4 Missing Variables

When it comes to comparing traditional non-spatial models to spatial models which explicitly account for the presence of SAC, variable selection proves necessary. One way to explain the differences between non-spatial and spatial approaches in selecting variables is that non-spatial models tend to recover the missing spatial information by including environmental variables that happen to be spatially autocorrelated (Bahn et al., 2006). Failing to incorporate relevant localized, spatially autocorrelated variables is one of the primary sources, if not the first, of rSAC. Leaving out important spatially autocorrelated explanatory variables will directly lead to model misspecification (Bini et al., 2009; Miller, 2012), which potentially produces rSAC and creates an instability associated with the Lennon (2000)'s 'red shift' problem (Bini et al., 2009). As corroborated by Bini et al. (2009), whenever such unmodeled spatially independent variables are included in the model, the level of rSAC goes down. On the contrary, when SAC is accounted for as in the case of a spatially explicit model, the relative importance likely decreases for non-spatially autocorrelated explanatory variables. Certain variables influence the response of biogeographical and ecological processes essentially at local scales. Performing broad-scale modeling will undermine such localized dependent variables, thus resulting in the creation of rSAC (Diniz-Filho et al., 2003). Similarly, studies suggest that failing to include important variables also causes positive rSAC, which may be an indicator for model misspecification (Lichstein et al., 2002; Diniz-Filho et al., 2008; Kissling et al., 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 supported by Diniz-Filho and Bini

(2005), Miller et al. (2007), Václavík et al. (2012), and Crase et al. (2014). Wu and Zhang (2013) similarly invoke missing spatially-structured covariates as factors that are responsible for rSAC.

3.5 Sampling Design

By the “sampling design” designation, we mean to consider sampling size, measurement, sampling scheme, and sampling intensity. Each one of these components can potentially lead to residual spatial autocorrelation as mentioned by previous studies. Bini et al. (2009) observed that a high degree of rSAC is often present in datasets with multiple observations. In contrast, Lichstein et al. (2002: 458) suggested that autocorrelated residuals can be caused by poor measurement of an important autocorrelated variable. In sampling, these are termed “artifacts” in that they are not a result of the environment but rather caused by the researcher (Dormann, 2007a; Crase et al., 2014). For these authors, such artifacts are difficult to correct, and they ultimately display rSAC. The artifacts are generated by species-specific bias or by differences in how species are lumped or split into groups. For instance, 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 would, thus creating a bias unrelated to the environment. Furthermore, a different sampling scheme would produce rSAC when regions of a known occurrence are sampled with higher intensity than regions of an unclear occurrence. Lastly, ecological interactions among species (e.g., competitive exclusion and founder effects) in isolated habitat patches, such as fragmented landscapes and lakes, will increase the level of SAC in assemblage data

that are absent from individual species distribution data (Dormann, 2007a; Crase et al., 2014).

3.6 Assumptions and methodological approaches

Spatially autocorrelated residuals can be the result of falsely assuming linearity between two factors, using a wrong variable selection method, or ignoring the presence of non-stationarity in a dataset. As Bini et al. (2009: 197) put it, as an illustration, fitting a linear model to a quadratic distribution or response leads to the residuals being spatially autocorrelated. In addition, performing model selection requires modelers to follow several key steps, including variable selection. Various methods are used in variable selection, such as *P*-value, Adjusted R^2 , Aike information criterion, prediction and cross-validation, to name a few. Le Rest et al. (2014) suggested that the Akaike information criteria, when used as a metric to select variables in the presence of rSAC, proved to include unwanted variables to the detriment of other relevant variables, thereby ignoring the presence of dependence in such residuals. Bini et al. (2009) viewed non-stationarity as the non-consistency in the relationship between variables throughout the whole extent of the data. For Miller (2012), non-stationarity is less intuitive and less used compared to SAC and has only lately been incorporated in SDM. The author suggests that 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 represents the statistical formalization of spatial heterogeneity, which defines uneven spatial distribution (like SAC, it is generally the result of sampling differences, another process in different locations of the study area or model misspecification such as missing

variables). Bini et al. (2009: 200) found that high rSAC usually exists in datasets with high levels of non-stationarity. Comparably, Lichstein et al. (2002: 449) contended that misspecifying a model form, such as assuming linearity when the relationship is nonlinear, may lead to spatially autocorrelated residuals. For Wu and Zhang (2013: 59, 60), rSAC will eventually be caused by linearity oversimplification. Finally, the consensus view from among these studies is that residual structures may result from an assumption one holds about the system under study or the methodological approach that one chooses.

Table 3. 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	Partialing 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 & Biogeography	Elaborate	Species richness of bird
8	Diniz-Filho et al.	2005	Global Ecology & Biogeography	Elaborate	Bird species richness and SAC
9	Diniz-Filho et al.	2008	Global Ecology & Biogeography	Elaborate	Model selectin in mammal species
10	Dormann	2007a	Global Ecology & Biogeography	Elaborate	Spatial and non-spatial models in ecology
11	Griffith et al.	2006	Ecology	Elaborate	Eigenfunction in ecological modelling
12	Griffith	2000	Journal of Geographical of Systems	Elaborate	Regression modelling 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

Table 3.1 (continued)

16	Kim et al.	2016	Soil Science Society of America Journal	Elaborate	Multiple SAC in Soil-landform modelling
17	Kissling et al.	2008	Global Ecology & Biogeography	Elaborate	SAC and Model selection
18	Lichstein et al.	2002	Ecological Monographs	Elaborate	Models and breeding habitats of songbirds
19	Oliveira et al.	2012	Biodiversity Conservation	Elaborate	Climatic suitability of Biome in climate change
20	Oliveira et al.	2014	Ecography	Elaborate	Ecological niche modeling of plant species
21	Sheehan et al.	2016	Ecology and Evolution	Elaborate	Bird species habitat
22	Ortiz-Yusty et al.	2013	Caldesia	Elaborate	Species richness and climate
23	Pickup et al.	1986	Ecological Modelling	Elaborate	Prediction of erosion and deposition
24	Le Rest et al.	2014	Global Ecology & 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 & Invasive forest pathogen distribution
27	Veloz	2009	Journal of Biogeography	Elaborate	Niche modeling and plant species distribution
28	Wu et al.	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

Table 3.1 (continued)

32	Record et al.	2013	Global Ecology and Biogeography	Elaborate	Plant species distribution projection and SAC
33	Naimi et al.	2011	Journal of Biogeography	Elaborate	SAC and species Occurrence modelling
34	Ficetola et al.	2012	Ecography	Elaborate	SAC and reptile species dispersal
35	Dormann	2007 b	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 Predictability Marine Nematode biodiversity
38	Dowd et al.	2014	Ecological Applications	Elaborate	Coastal marine benthic microfaunal distribution modelling
39	Hefley et al.	2017	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	Genetics-Wolf species richness and distribution
44	Record et al.	2013	Ecosphere	Elaborate	SAC-Climate change prediction
45	Austin	2002	Ecological modelling	Elaborate	Spatial species distribution modeling
46	Carl et al	2007	ecological Modelling	Elaborate	SAC in Species distribution

Table 3.1 (continued)

47	Dirnböck et al.	2004	Journal of Vegetation Science	Elaborate	SAC-SP habitat distribution
48	Zhang et al.	2009	Forest Science	Elaborate	Species model comparison-SAC
49	Gwenzi et al.	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-endangered species
51	Davis et al.	2016	Ecosphere	Elaborate	Urban plant invasion
52	Mattsson et al.	2013	PloS ONE	Simple mention	SP Assamblage-SAC
53	Chun et al.	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 modelling
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

Table 3.1 (continued)

62	Hefley et al.	2017	Ecology	Simple mention	Functions in spatial ecological modelling
63	Estrada et al.	2016	Animal Conservation	No mention	Biodiversity-Bird species
64	Ali et al.	2010	Water Resources Research	No mention	Soil moisture and topographical modelling
65	Anselin et al.	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 SP
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 modelling
69	Weeks et al.	2017	River Research and Applications	No mention	Snail-Aquatic vegetation
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 modelling

Table 3.1 (continued)

78	Kleisner et al.	2010	Marine Ecology Progress Series	No mention	Pelagic fish modelling
79	Tarkhnishvili et al.	2012	Biological Journal of the Linnean Society	No mention	Distribution of forest species
80	Wiegand et al.	2004	OIKOS	No mention	Point pattern analysis in ecology
81	Yu et al.	2012	ProQuest Dissertations and Theses	No mention	Tree growth modelling 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	Distributions of oak wasps 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 et al.	2014	Agricultural and Forest Entomology	No mention	Ecological niche modeling aphids
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

Table 3.1 (continued)

95	Ciccarelli et al.	2016	Folia Geobotanica	No mention	Spatial modeling 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

CHAPTER 4. CONCLUSIONS

Macroecological and biogeographical modelers are aware that there are multiple facets of spatial autocorrelation. Incorporating SAC in the modeling process, comparing spatial and non-spatial modeling, and identifying the potential issues arising from the presence of spatial dependence were often recognized in the studies surveyed in this research. There appears to be a consensus among modelers that spatially explicit models in most cases outperform non-spatial models that ignore the effects of spatial structure. Understanding, however, why models show such differences in performance and the circumstances under which they amplify remains unclear (Cruse et al., 2014; Kim et al., 2016; Miralha and Kim, 2018).

Our review of the prominent works addressing the topic of SAC allowed us to identify and categorize the potential sources of rSAC. The nature of the data, missing autocorrelated variables, scalar extent of the study and sampling design, as well as the kinds of methodological assumptions represent the primary causes of SAC in model residuals. This categorization is a critical finding given that it provides a better understanding of the circumstances under which model residuals are spatially structured.

However, the scarcity in quantifiable parameters prevented us from evaluating the magnitude to which rSAC becomes problematic in SDM. In our review, the percentage of the papers (64% comprising those *elaborate* and *simple mention* categories in Table 3.1) that allude to rSAC for the most part do so slightly and lack quantitative information that would in turn facilitate any kind of quantitative comparisons. This review shows that rSAC in macroecological and biogeographical models remains predominantly endogenous, in that intrinsic biotic processes drive the presence of spatial autocorrelation in the residuals.

This suggests a need for further investigations that aim to quantify rSAC and analyze how it accumulates. It is critical to establish the role of missing variables, various sampling designs and types of data along with model misspecification in generating the presence of SAC in model residuals. Consequently, we strongly recommend using combinations of these factors at multiple scales to model macroecological and biogeographical processes.

BIBLIOGRAPHY

- Ali GA, Roy AG, and Legendre P (2010) Spatial relationships between soil moisture patterns and topographic variables at multiple scales in a humid temperate forested catchment. *Water Resource Research*.
- Anselin L, Syabri I, and Kho Y (2006) GeoDa: An Introduction to spatial data. *Geographical Analysis* ISSN 0016-7363
- Anselin L (2002) Under the hood: Issues in the specification and interpretation of spatial regression models. *Agricultural Economics* 27: 247-267
- Anselin L and AK Bera (1998) *Spatial dependence in linear regression models with an introduction to spatial econometrics*. Statistics Textbooks and Monographs
- Augustin NH, Cummins RP, and French DD (2001) Exploring spatial vegetation dynamics using logistic regression and a multinomial logit model. *Journal of Applied Ecology* 38: 991–1006
- Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling* 157 (2): 101-118
- Bahn V, O'Connor RJ, and Krohn WB (2006) Importance of spatial autocorrelation in modeling bird distributions at a continental scale. *Ecography* 29:835-844
- Betts MG, Diamond AW, Forbes GJ, et al. (2006) The importance of spatial autocorrelation, extent and resolution in predicting forest bird occurrence. *Ecological modelling* 191:197-224
- Bini ML, Diniz-Filho JAF, Rangel TFLVB, et al. (2009) Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. *Ecography* 32: 193-204
- Borcard D, Legendre P, and Drapeau P (1992) Partialling out the spatial component of ecological variation. *Ecology* 73 (3):1045-1055
- Bonada N, Sylvain Dolédec S, and Stutzner B (2012) Spatial autocorrelation patterns of stream invertebrates: exogenous and endogenous factors. *Journal of Biogeography* 39: 56–68
- Büchi L, Christin PA, and Hirzel AH (2009) The influence of environmental structure on the life-history traits and diversity of species in a metacommunity. *Ecological modelling* 220: 2857-2864
- Carl G and Kühn I (2007) Analyzing spatial autocorrelation in species distributions using Gaussian and logit models. *Ecological modelling* 207:159–170
- Chang J, Chen D, Ye X, et al. (2012) Coupling Genetic and Species Distribution Models to Examine the Response of the Hainan Partridge (*Arborophila ardens*) to Late Quaternary Climate. *PloS ONE*
- Chun Y and Griffith DA (2011) Modeling Network Autocorrelation in Space–Time Migration Flow Data: An Eigenvector Spatial Filtering Approach. *Annals of the Association of American Geographers* 101(3): 523–536
- Ciccarelli D and Bacaro G (2016) Quantifying plant species diversity in coastal dunes: a piece of help from spatially constrained rarefaction. *Folia Geobot* 51:129–141 DOI 10.1007/s12224-016-9249-9
- Cliff N (1984) An improved internal consistency reliability estimate. *Journal of Educational Statistics* 9 (2): 151-161

- Cruse B, Liedloff A, Vesik PA, et al. (2014) Incorporating spatial autocorrelation into species distribution models alters forecasts of climate-mediated range shifts. *Global Change Biology* 20: 2566–2579
- Cruse B, Liedloff A, and Wintle BA (2012) A new method for dealing with residual spatial autocorrelation in species distribution models. *Ecography* 35: 879–888
- Davis AJS, Singh KK, Thill J, and Meentemeyer RK (2016) Accounting for residential propagule pressure improves prediction of urban plant invasion. *Ecosphere* 7(3): e01232. 10.1002/ecs2.1232
- Diniz-Filho JAF, Bini LM, Hawkins BA (2003) Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology & Biogeography* 12: 53-64
- Diniz-Filho JAF and Bini LM. (2005) Modelling geographical patterns in species richness using eigenvector-based spatial filters. *Ecography and Biogeography* 14: 177-185
- Diniz-Filho JAF, Rangel TFLVB, and Bini LM (2008) Model selection and information theory in geographical ecology. *Global Ecology and Biogeography* 17: 479–488
- Dirnböck T and Dullinger S (2004) Habitat distribution models, spatial autocorrelation, functional traits and dispersal capacity of alpine plant species. *Journal of Vegetation Science* 15: 77-84
- Dorken ME, Freckleton RP, and Pannell JR (2017) Small-scale and regional spatial dynamics of an annual plant with contrasting sexual systems. *Journal of Ecology* 105: 1044–1057
- Dormann C (2007a) Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography*, 29–138
- Dormann C (2007b) Assessing the validity of autologistic regression. *Ecological Modelling* 207(2-4): 234-242
- Dowd M, Grant J, and Lu L (2014) Predictive modeling of marine benthic macrofauna and its use to inform spatial monitoring design. *Ecological Applications* 24(4): 862-876
- Dronova I, Beissinger SR, Burnham JW, and Gong P (2016) Landscape-Level Associations of Wintering Waterbird Diversity and Abundance from Remotely Sensed Wetland Characteristics of Poyang Lake. *Remote Sensing* 8(6), 462: 1-22 doi:10.3390/rs8060462
- Ennen JR, Agha M, Matamoros WA, et al. (2016) Using climate, energy, and spatial-based hypotheses to interpret macroecological patterns of North America chelonians. *Canadian Journal of Zoology* 94: 453–461
- Epperson BK (2000) Spatial and space–time correlations in ecological models. *Ecological Modelling* 132: 63–76
- Estrada A, Delgado MP, Arroyo B, et al. (2016) Forecasting Large-Scale Habitat Suitability of European Bustards under Climate Change: The Role of Environmental and Geographic Variables. *PloS ONE* /DOI:10.1371/journal.pone.0149810
- Estrada CG and Rodriguez-Estrella R (2016) In the search of good biodiversity surrogates: are raptors poor indicators in the Baja California Peninsula desert? *Animal Conservation* 19: 360–368
- Ficetola GF, Manenti R, De Bernard F, and Padoa-Schioppa E (2012) Can patterns of spatial autocorrelation reveal population processes? An analysis with the fire salamander. *Ecography* 35: 693–703

- Getis A (2008) A History of the Concept of Spatial Autocorrelation: A Geographer's Perspective. *Geographical Analysis* ISSN 0016-7363
- Griffith DA and Peres-Neto PR (2006) Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. *Ecology* 87(10): 2603–261
- Griffith DA (2000) A linear regression solution to the spatial autocorrelation problem. *Geographical System* 2: 141-156
- Guénard G, Lanthier G, Harvey-Lavoie S, et al. (2016) A spatially- explicit assessment of the fish population response to flow management in a heterogeneous landscape Guillaume. *Ecosphere* 7(5)
- Güler B, Jentsch A, Apostolova I, et al. (2016) How plot shape and spatial arrangement affect plant species richness counts: implications for sampling design and rarefaction analyses. *Journal of Vegetation Science* 27: 692-703
- Gwenzi D and Lefsky MA (2017) Spatial Modeling of Lidar-Derived Woody Biomass Estimates Collected Along Transects in a Heterogeneous Savanna Landscape. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10 (1)
- Hawkins BA, Diniz-Filho JAF, Bini LM, et al. (2007) Red herrings revisited: spatial autocorrelation and parameter estimation in geographical ecology. *Ecography* 30: 375-384
- Hefley TJ, Broms KM, Brost BM (2017) The basis function approach for modeling autocorrelation in ecological data. *Ecology* 98(3): 632–646
- Hefley TJ, Hooten MB, Russell RE, et al. (2017) When mechanism matters: Bayesian forecasting using models of ecological diffusion. *Ecology Letters* 20: 640–650
- Hindrikson M, Remm J, Pilot M, et al. (2017) Wolf population genetics in Europe: a systematic review, meta-analysis and suggestions for conservation and management. *Biological Reviews* 92: 1601–1629.
- Hongoh V, Berrang-Ford I, Scott ME, and Lindsay IR (2012) Expanding geographical distribution of the mosquito, *Culex pipiens*, in Canada under climate change. *Applied Geography* 33: 53-62
- Ingberman B, Fusco-Costa R, and Monteiro-Filho ELA (2016) A Current Perspective on the Historical Geographic Distribution of the Endangered Muriquis (*Brachyteles spp.*): Implications for Conservation. *PloSONE*
|DOI:10.1371/journal.pone.0150906
- Ishihama F, Takeda T, and Takenaka HOA (2010) Comparison of effects of spatial autocorrelation on distribution predictions of four rare plant species in the Watarase wetland. *Ecological Research* 25: 1057–1069
- Jackson MM, Gergel SE, and Martin K (2015) Citizen science and field survey observations provide comparable results for mapping Vancouver Island White-tailed Ptarmigan (*Lagopus Leucura saxatillis*) distributions. *Biological conservation* 181: 162-172
- Kim D (2013) Incorporation of multi-scale spatial autocorrelation in soil moisture–landscape modeling. *Physical Geography*
- Kim D, Hirmas DR, McEwan RW, et al. (2016) Predicting the Influence of Multi-Scale Spatial Autocorrelation on Soil–Landform Modeling. *Soil Science Society of America*.
- Kissling WD and Carl G (2008) Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography* 17: 59-71

- Kleisner KM, Walter III JF, Diamond SL, and Die DJ (2010) Modeling the spatial autocorrelation of pelagic fish abundance. *Marine Ecology Progress Series* 411: 203-213
- Komac B, Esteban P, Trapero L, and Caritg R (2016) Modelization of the Current and Future Habitat Suitability of *Rhododendron ferrugineum* Using Potential Snow Accumulation. *PloS ONE* /DOI:10.1371/journal.pone.0147324
- Kühn I (2007) Incorporating spatial autocorrelation may invert observed patterns. *Diversity and Distributions* 13: 66–69
- Lennon JJ (2000) Red-shifts and red herrings in geographical ecology. *Ecography* 23: 101-113.
- Legendre P (1993) Spatial autocorrelation: Trouble or new paradigm? *Ecology* 74: 1659–1673
- Le Rest K, Pinaud D, Monestiez P, et al. (2014) Spatial leave-one-out cross-validation for variable selection in the presence of spatial autocorrelation. *Global Ecology and Biogeography* 23: 811–820
- Lichstein JW, Simons TR, Shriner SA, and Franzreb KE (2002) Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs*, 72(3), pp. 445-463
- Lloyd NJ, Nally RM, and Lake PS (2005) Spatial autocorrelation of assemblages of benthic invertebrates and its relationship to environmental factors in two upland rivers in southeastern Australia. *Diversity and Distributions* 11, 375–386
- Marmion M, Luoto M, Heikkinen RK, and Thuiller W (2009) The performance of state-of-the-art modelling techniques depends on geographical distribution of species. *Ecological modelling* 220: 3512-3520
- Mattsson BJ, Elise F, Zipkin SF, Gardner B, et al. (2013) Explaining Local-Scale species Distributions: Relative Contributions of Spatial Autocorrelation and Landscape Heterogeneity for an Avian Assemblage. *PloS ONE*
- Merckx B, Goethals P, Steyaert M, et al. (2009) Predictability of marine nematode biodiversity. *Ecological Modelling* 220:1449-1458
- Mets KD, Armenteras D, and Dávalos LM (2017) Spatial autocorrelation reduces model precision and predictive power in deforestation analyses. *Ecosphere* 8(5): e01824.10. 1002/ecs2.1824
- Miller J, Franklin J, and Aspinall R (2007) Incorporating spatial dependence in predictive vegetation models. *Ecological Modelling* 225-242
- Miller JA (2012) Species distribution models: Spatial autocorrelation and non-stationarity. *Progress in Physical Geography* 681-692
- Miralha L and Kim D (2018) Accounting for and predicting the influence of spatial autocorrelation in water quality modeling. *ISPRS International Journal of Geo-Information* 7: 64. doi:10.3390/ijgi7020064
- Naimi B, Skidmore AK, Groen TA, and Hamm NAS (2011) Spatial autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling. *Journal of Biogeography* 38: 1497-150
- Nicolaus M, Brommer JE, Ubels R, et al. (2013) Exploring patterns of variation in clutch size–density reaction norms in a wild passerine bird. *Journal of Evolutionary Biology* 26: 2031-2043
- Oliveira G, Rangel TF, Lima-Ribeiro MS, et al. (2014) Evaluating, partitioning, and mapping the spatial autocorrelation component in ecological niche modeling: a new approach based on environmentally equidistant records. *Ecography* 37: 637-647

- Oliveira G, Araújo MB, Rangel TF, et al. (2012) Conserving the Brazilian semiarid (Caatinga) biome under climate change. *Biodiversity and Conservation* 21(11): 2913–2926
- Ortiz-Yusty CE, Paez V and Zapata FA (2013) Temperature and precipitation as predictors Species richness in Northern Andean amphibian from Colombia/ Temperatura y precipitación como elementos de predicción de la riqueza de especies de anfibios del norte de los Andes en Colombia. *Caldasia* 35(1): 65-80
- Piazzini S, Caruso T, Favilli L, and Manganelli G (2011) Role of Predators, Habitat Attributes, and Spatial Autocorrelation on the Distribution of Eggs in the Northern Spectacled Salamander (*Salamandrina perspicillata*). *Journal of Herpetology* 45(3):389-394.
- Pickup G and Chewings VH (1986) Random Field Modeling of spatial variations in erosion and deposition in flat alluvial landscapes in arid Central Australia. *Ecological Modelling* 33: 269-296
- Platts PJ, McClean CJ, Lovett JC, and Marchant R (2008) Predicting three distributions in an East African biodiversity hotspot: model selection, data bias and envelope uncertainty. *Ecological Modelling* 218: 121-134
- Poley LG, Pond BA, Schaefer JA, et al. (2014) Occupancy patterns of large mammals in the Far North of Ontario under imperfect detection and spatial autocorrelation. *Journal of Biogeography* 41: 122-132
- Record S, Fitzpatrick MC, Finley AO, et al. (2013) Should species distribution models account for spatial autocorrelation? A test of model projections across eight millennia of climate change. *Global Ecology and Biogeography* 12017
- Record S, Charney ND, Zakaria RM et al. (2013) Projecting global mangrove species and community distributions under climate change. *Ecosphere* 4(3):34. <http://dx.doi.org/10.1890/ES12-00296.1>
- Revermann R, Schmid H, and Zbinden N (2012) 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. *Journal of Ornithology* 153:891–905
- Rodriguez A, Gómez JF, and Nieves-Aldrey (2015) Modeling the potential distribution and conservation status of three species of oak gall wasps (Hymenoptera: Cynipidae) in the Iberian range. *Journal of Insect Conservation* 19:921-934
- Roth T, Bühler C, and Amrhein V (2016) Estimating Effects of Species Interactions on Populations of Endangered Species. *American naturalist* 187(4)
- Santos SM, Mira AP, and Mathias ML (2009) Factors influencing large-scale distribution of two sister species of pine voles (*Microtus lusitanicus* and *Microtus duodecimcostatus duodecimcostatus*): the importance of spatial autocorrelation. *Canadian Journal of Zoology* 87
- Seymour L (2005) Spatial Data Analysis: Theory and Practice. *Journal of the American Statistical Association* 100:469, 353-353, DOI: 10.1198/jasa.2005.s8
- Sheehan KL, Esswein ST, Dorr BS, et al. (2016) Using species distribution models to define nesting habitat of the eastern metapopulation of double- crested cormorants. *Ecology and Evolution* 7: 409-418
- Siderov K (2005) Spatial data analysis: theory and practice. *Austral Ecology* 30: 237-241
- Siesa ME, Manenti R, Padoa-Schioppa E, et al. (2011) Spatial autocorrelation and the

- analysis of invasion processes from distribution data: a study with the crayfish *Procambarus clarkia*. *Biol Invasions* 13:2147–2160
- Tarkhnishvili D, Gavashelishvili A, and Mumladze L (2012) Palaeoclimatic models help to understand current distribution of Caucasian forest species. *Biological Journal of the Linnean Society* 105: 231-248.
- Tallowin O, Allison A, Algar Ac, et al. (2017) Papua New Guinea terrestrial-vertebrate richness: elevation matters most for all except reptiles. *Journal of Biogeography* 44: 1734-1744
- Václavík T and Metemeyer RK (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling* 220 (23): 3248-3258
- Václavík T, Kupfer JA, and Meentemeyer RK (2012) Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM)
- Veloz SD (2009) Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography* 36: 2290-2299
- Warren DL, Cardillo M, Rosauer DF, and Bolnick DI (2014) Mistaking geography for biology: Inferring processing from species distributions. *Trends in Ecology & Evolution* 29 (10)
- Weeks AM, De Jager NR, Haro RJ, and Sandland GJ (2017) Spatial and Temporal Relationships between the Invasive Snail *Bithynia tentaculata* and Submersed Aquatic Vegetation in Pool 8 of the Upper Mississippi River. *River Research and Applications* 35: 729-739
- Wiegand T and Moloney KA (2004) Rings, circles, and null-models for point pattern analysis in ecology. *OIKOS* 104: 209-229
- Wieczorek K and Bugaj-Nawrocka A (2014) Invasive aphids of the tribe Siphini: a model of potentially suitable ecological niches. *Agricultural and Forest Entomology* 16: 434-443
- Wu D, Liu J, Zhang G, et al. (2009) Incorporating spatial autocorrelation in cellular automata model: An application to the dynamics of Chinese tamarisk (*Tamarisk chinensis* Lour.). *Ecological Modelling* 220: 3490-3498
- Wu W and Zhang L (2013) Comparison of spatial and non-spatial logistic regression models for modeling the occurrence of cloud cover in north-eastern Puerto Rico. *Applied Geography* 37: 52e62
- Wulder MA, White JC, Coops NC, et al. (2007) Using spatial autocorrelation to compare compare outputs from a forest growth model. *Ecological Modelling* 209: 264-276
- Yu MHP, Charles M, Canham C, et al. (2012) Modeling tree growth and seedling recruitment in a selectively logged temperature forest. *ProQuest Dissertations and Theses*
- Zhang L, Ma Z, and Guo L (2009) An Evaluation of Spatial Autocorrelation and Heterogeneity in the Residuals of Six Regression Models. *Forest Science* 55(6)
- Zhu W, JIA S, LÜ A, and Yan T (2012) Analyzing and modeling the coverage of Vegetation in the Qaidam Basin of China: The role of spatial autocorrelation. *Journal of Geographical Sciences* 22(2): 346-358

VITA

I, Guetchine Gaspard, holds a Bachelor's degree Agricultural Science, one Master's degree in Geography and Environmental Resources as a Fulbright grantee, and another Master's degree in Physical Geography. I have extensive work experience in Haiti, where I worked as Program Manager, Professor, and Consultant. I also worked full-time in the United States as Teaching and Research Assistant, while working part-time during intercessions first as a GIS intern and as indexing assistant in oral history. Below is a listing of my education, professional positions, honors, and publications.

- 1. MA.** Geography. University of Kentucky, KY, USA. Expected May 2021
MS: Geography and Environmental Resources. Southern Illinois University. IL, USA, 2011-2013.
BS: Agricultural Science. Faculty of Agronomy and Veterinary Medicine, State University of Haiti (FAMV-UEH). 2001-2006
- 2.** Gaspard, G., Kim, D. & Chun, Y. Residual spatial autocorrelation in macroecological and biogeographical modeling: a review. *j ecology environ* 43, 19 (2019). <https://doi.org/10.1186/s41610-019-0118-3>.
- 3. Teaching & Research Assistant.** Geography. University of Kentucky, KY, USA 8/16 – 5/21
Indexing Assistant. Oral History Center/University of Kentucky. 2/17-8/17; 6/18- 8/19
Professor. University of Jeremy. Haiti, 8/2015 - 9/2015
GIS Intern. City Planning of Carbondale, IL, USA. 2012.
Program Manager. Association of Mayors (AMAGA) / Nantes Métropole. 02 /14 – 08/16.
Program Manager. Concert-Action, Haiti. November 2010-May 2011
Training Manager. CARE-Haiti, Haiti. May 2008- December 2009
Coordinator. (IFOS): Training Institute from South / Ministry of Education(MENFP) /World Bank. 10/2015 - 12/2015
Consultant. New Grand'Anse Foundation /Lutheran World Federation. Haiti. 2014.
Consultant. Lutheran World Federation (LWF). Gressier, Haïti, 2013.