Review Article

Species Distribution Modeling in Latin America: A 25-Year Retrospective Review

Nicolás Urbina-Cardona[1] ⓘ, Mary E. Blair[2], Maria C. Londoño[3], Rafael Loyola[4,5], Jorge Velásquez-Tibatá[6], and Hernan Morales-Devia[7]

Abstract

Species distribution modeling (SDM) is a booming area of research that has had an exponential increase in use and development in recent years. We performed a search of scientific literature and found 5,533 documents published from 1993 to 2018 using SDM, representing a global network of 4,329 collaborating institutions from 155 countries, with Brazil and Mexico being in the top 10 of the most prolific countries globally. National Autonomous University of Mexico, Chinese Academy of Sciences, University of Kansas, and U.S. Geological Survey are the most prolific institutions worldwide. Latin American institutions ($n = 556$) participated in 1,000 (18% of global productivity) documents published in collaboration with 591 institutions outside Latin American countries, from which the National Autonomous University of Mexico, Federal University of Goiás, Institute of Ecology A.C., National Scientific and Technical Research Council in Argentina, University of São Paulo, and University of Brasilia were the most productive. From this body of literature, the most frequently modeled taxonomic groups were Chordata and Insecta, and the most common realms of application were conservation planning and management, climate change, species conservation, epidemiology, evolutionary biology, and biological invasions. From the 36 modeling methods identified to generate SDMs, MaxEnt is used in 73.5% of the papers, followed by Genetic Algorithm for Rule-Set Prediction (GARP) with 18.7%, and just 7.4% of the papers compared between 3 and 10 modeling methods. In Latin American countries, productivity in SDM research could be improved as the network of collaborations diversifies and connects with other productive countries (such as United Kingdom, China, Spain, Germany, Australia, and France). The scientific collaboration between Latin American countries should be increased, as the most prolific countries (Brazil, Mexico, Argentina, and Colombia) share less than 10% of its productivity. Some of the main challenges for SDM development in Latin America include bridging the gaps from (a) software use to research productivity and (b) translation to decision-making. To address these challenges, we propose to strengthen communities of practice where modelers, species experts, and decision-makers come together to discuss and develop SDM to shift and enhance current paradigms on how science and decision-making are linked.

[1]Departamento de Ecología y Territorio, Facultad de Estudios Ambientales y Rurales, Pontificia Universidad Javeriana, Bogotá, Colombia
[2]Center for Biodiversity and Conservation, American Museum of Natural History, New York, NY, USA
[3]Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, Colombia
[4]Conservation Biogeography Lab, Departamento de Ecologia, Universidade Federal de Goiás, Goiânia, Brazil
[5]Fundação Brasileira para o Desenvolvimento Sustentável, Rio de Janeiro, Brazil

[6]NASCA Conservation Program, The Nature Conservancy, Bogotá, Colombia
[7]Biblioteca General Alfonso Borrero Cabal, Pontificia Universidad Javeriana, Bogotá, Colombia

Corresponding Author:
Nicolás Urbina-Cardona, Departamento de Ecología y Territorio, Facultad de Estudios Ambientales y Rurales, Pontificia Universidad Javeriana, Carrera 7 No. 40-62, Bogotá, Colombia.
Email: urbina-j@javeriana.edu.co

Introduction

Species distribution modeling (SDM) is one of the most recent and robust tools used in ecology, biogeography, and macroecology (Soberón, Osorio-Olivera, & Peterson, 2017). SDM is a method of assessing areas that provide suitable abiotic environments (often, climate) for a given species, using the relationship between observed points of occurrence and environmental variables, to generate a spatial prediction of regions within which environmental conditions are suitable for species survival and population growth. Previous reviews of SDM have focused on their use to predict the distribution of invasive species (Barbosa, Schneck, & Melo, 2012), to determine MaxEnt model usage in wildlife research (Baldwin, 2009), and to inform conservation practitioners in decision-making (Villero, Pla, Camps, Ruiz-Olmo, & Brotons, 2017), among others. Cayuela et al. (2009) pointed out that 39% of the papers they reviewed on SDM were mainly focused on the development of new methodologies and the evaluation of performance of different methods, while the other 61% of the studies applied SDM in contexts such as species conservation, biological invasions, climate change, autecology, and biogeography.

The state of knowledge of biodiversity at the species' level tends to be lower in megadiverse countries (Meyer, Kreft, Guralnick, & Jetz, 2015), in which the so-called Wallacean shortfall is evident (i.e., the incomplete knowledge of species distributions; Lomolino, 2004). Latin America harbors 58.3% of the world's megadiverse countries (United Nations Environmental Program/Convention on Biological Diversity, 2016), and in this region, a very high proportion of endemic species is specially problematic for overall biodiversity knowledge as there is scarce data on geographical distribution for some highly rare species that occur in specialized microhabitats (Urbina-Cardona & Loyola, 2008). While uncertainty in species distributions is highly unknown for Latin America (but see modeling methods to estimate uncertainty in Diniz-Filho et al., 2009; Loyola, Lemes, Faleiro, Trindade-Filho, & Machado, 2012; Sales, Neves, De Marco, & Loyola, 2017), countries in the region are increasing their economic dependence on extraction activities, profoundly impacting their ecosystems (Rosales, 2008; Villarroya, Barros, & Kiesecker, 2014). As pinpointed by Villero et al. (2017), SDM should inform decision-making for unknown and rare species, but the absence of information should be viewed as an opportunity for collaboration between institutions more than a weakness of the region. Networks of scientific collaboration could pave the road to improve species' knowledge not just on their distributions but on habitat requirements, tolerance to disturbance, and popula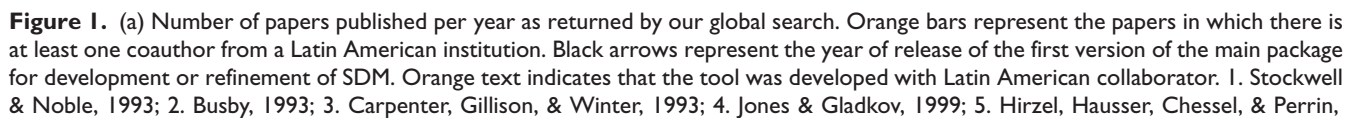tion dynamics, to have robust forecasting on the impacts of environmental change on species' conservation status (Franklin, 2010).

In this regard, it is crucial to evaluate and understand publication patterns to identify possible taxonomic and ecosystem bias, as well as opportunities for collaboration both within the region and with institutions from other regions, to accelerate the generation of new knowledge on SDM and to reduce the Wallacean shortfall for Latin American countries. In recent years, an exponential increase in online biodiversity data (e.g., GBIF.org), environmental data (Worldclim v1: Hijmans, Cameron, Parra, Jones, & Jarvis, 2005; CliMond; Kriticos et al., 2012; Microclim: Kearney, Isaac, & Porter, 2014; Worldclim v2: Fick & Hijmans, 2017; Chelsa: Karger et al., 2017; Envirem: Title & Bemmels, 2017; MERRAclim: Vega, Pertierra, & Olalla-Tárraga, 2017), and open software to run models (R programming language) have resulted in increased productivity of SDMs with several new modeling methods developed (Figure 1(a)). Thus, it is essential to have an updated review of SDM usage from Latin American institutions.

Here, we aim to (a) identify the spatial and temporal pattern of publications using SDM from Latin American countries, from a global perspective; (b) identify the network of collaborations at both national and institutional level that resulted in SDM publications from Latin American institutions in the scientific literature; and (c) characterize SDM research conducted by Latin American institutions according to taxonomic groups, realm of application, geographic region, ecosystem type, and modeling methods used. We then expand upon our results to identify the main challenges for SDM development in Latin America and discuss ways to overcome them.

Methods

One of the authors (H.M.-D.) did a structured literature search on Web of Science (WoS) and SCOPUS for scientific literature up to December 31, 2018, using a query string for each database (see online Appendix A). The metadata for each document was compiled and standardized using tech mining techniques (Porter & Cunningham, 2004) and structured into a database. For data processing, we used Vantagepoint version 10.0 (Porter & Cunningham, 2004), a text mining software that allows incorporating and standardizing large volumes of information from multiple online research sources, such as SCOPUS and WoS, to analyze and visualize information to find patterns and relationships (Search Technology, 2018). With Vantagepoint, we deleted duplicates (there was 69% duplicity between WoS and SCOPUS) and obtained 13,388 unique documents. The title of each document was read by the first author (N.U.-C.) excluding 7,109 documents in thematic

Figure 1. (a) Number of papers published per year as returned by our global search. Orange bars represent the papers in which there is at least one coauthor from a Latin American institution. Black arrows represent the year of release of the first version of the main package for development or refinement of SDM. Orange text indicates that the tool was developed with Latin American collaborator. 1. Stockwell & Noble, 1993; 2. Busby, 1993; 3. Carpenter, Gillison, & Winter, 1993; 4. Jones & Gladkov, 1999; 5. Hirzel, Hausser, Chessel, & Perrin,

areas outside disciplines of interest (online Appendix B). From the resulting 6,279 documents, N.U.-C. read the title, abstract, and keywords sections and excluded 746 documents that did not employ SDM (online Appendix B). The final database of 5,533 documents represents the global studies on SDM included in qualitative synthesis (online Appendix B).

From this compilation of global literature, we obtained the country and institutional affiliation of each coauthor. We then normalized the country and institution names, filtered Latin American institutions, and created co-occurrence matrices (online Appendix C). The Latin American database was composed of 1,000 documents with the participation of at least one coauthor from a Latin American institution. One of the authors (H.M.-D.) extracted 4,481 words from the title, abstract, and keywords, then N.U.-C. identified 1,661 pertinent terms, with a frequency of occurrence of more than 3 times, and conducted a more in-depth qualitative analysis, classifying each keyword into four different categories: taxonomic group, realm of application, geographic region, and ecosystem type. Each category was visualized through a foam tree (Carrot Search FoamTree, 2019) in which the available space is divided into polygons of different color and size, depending on their class and frequency, respectively.

To identify the modeling methods used on each document to model species distributions, N.U.-C. read the title and abstract sections of each document, from the Latin American database. In addition, J.V.-T. classified the different modeling methods into three categories: profile methods (i.e., methods that use only presence data to estimate geographic distributions; this includes envelope methods such as BIOCLIM, distance methods—DOMAIN, and multivariate methods—ecological niche factor analysis); statistical methods and machine learning methods. A total of 687 documents reported the modeling method in title and abstract sections (68.6% of the whole Latin American database) and were migrated to VOSviewer (van Eck & Waltman, 2010). VOSviewer

allows to build and visualize bibliometric maps of keywords based on co-occurrence data, using the modularity-based cluster and mapping techniques that minimize a weighted sum of squares of Euclidean distances between all pairs of items (categories) through an optimization process (van Eck & Waltman, 2013).

Based on the body of global and Latin American documents, we built and visualized collaboration networks at both national and institutional levels in which node-based centrality (number of links that each country or institution has) was applied to measure the importance of a node within a network (Bonacich, 1987; Freeman, 1977). We visualized the number of papers published by country in a map, using the COM complement Power Pivot for Excel—Office 2016.

We also used data from the American Museum of Natural History on the number of downloads of MaxEnt SDM software between December 12, 2016, and June 8, 2018 (178 days after the new open source version of MaxEnt became available; Phillips, Anderson, Dudik, Schapire, & Blair, 2017), to compare patterns of MaxEnt software downloads with SDM research publication patterns by country.

Results

Global Network of SDM Research

The global network of collaboration was composed by 155 countries, of which the 10 most productive, in descending order of productivity, were United States, United Kingdom, China, Spain, Germany, Australia, France, Brazil, Mexico, and Canada (Figure 1(b)). The United States led the global scientific production extensively with 1,863 papers (Figure 1(b)). Further, the 15 countries that represented the nodes with the highest centrality degree (with collaborations with more than 30 countries) were, in descending order, United States, Germany, United Kingdom, France, Spain, Australia, Canada, Italy, Netherlands, Switzerland, Brazil, China,

Figure 1. Continued.
2002; 6. Thuiller, 2003; 7. Phillips et al., 2006; 8. Freeman & Moisen, 2008; 9. Rangel, Diniz-Filho, & Araújo, 2009; 10. Guo & Liu, 2010; 11. Midgley et al., 2010; 12. Warren, Glor, & Turelli, 2010; 13. de Souza Muñoz et al., 2011; 14. Engler, Hordijk, & Guisan, 2012; 15. Qiao, Lin, Ji, & Jiang, 2012; 16. Hijmans & Elith, 2013; 17. Muscarella et al., 2014; 18. Aiello-Lammens, Boria, Radosavljevic, Vilela, & Anderson, 2015; 19. Duan, Kong, Huang, Wu, & Wang, 2015; 20. Mazzoni, Halvorsen, & Bakkestuen, 2015; 21. Barbosa, Brown, Jimenez-Valverde, & Real, 2016; 22. Brewer, O'Hara, Anderson, & Ohlemüller, 2016; 23. Cánovas, Magliozzi, Mestre, Palazón, & González-Wangüemert, 2016; 24. Leroy, Meynard, Bellard, & Courchamp, 2016; 25. Naimi & Araújo, 2016; 26. Qiao et al., 2016; 27. Brown, Bennett, & French, 2017; 28. Carl & Kühn, 2017; 29. Dedman, Officer, Clarke, Reid, & Brophy, 2017; 30. Di Cola et al., 2017; 31. Guillaumot, MarHn, Eléaume, & Saucède, 2017; 32. Schmitt, Pouteau, Justeau, de Boissieu, & Birnbaum, 2017; 33. Golding et al., 2018; 34. Kass et al., 2018; 35. Title & Bemmels, 2018. (b) Global production of SDM literature: Circle size represents the number of published documents per country. (c) Global collaboration network between countries to produce SDM scientific literature. Orange circles represent Latin American countries. Circle size represents the number of papers published by country. The thickness of the connections (links or edges) indicates the strength of collaboration that exists, between countries, at the level of scientific production on Species distribution modeling.

Belgium, Denmark, and South Africa, indicating their roles as a collaboration hubs (Figure 1(c)). Worldwide, the most important collaboration networks occurred between Portugal and Spain ($n = 84$ published papers in collaboration), United Kingdom and Spain ($n = 69$), Germany and United Kingdom (60), as well as between United States and United Kingdom ($n = 144$), Australia ($n = 135$), Mexico ($n = 112$), China ($n = 95$), Canada ($n = 94$), Spain ($n = 89$), Brazil ($n = 86$), France ($n = 86$), and Germany ($n = 85$; Figure 1(c)). From this network of collaboration, the most productive Latin American countries were Brazil, Mexico, Argentina, Colombia, Chile, and Ecuador, with 356, 350, 145, 103, 57, and 39 papers, respectively (Figure 1(b, c)).

Latin American Networks and Their Articulation With the World

The Latin American collaboration network was composed of 26 countries that mainly interact with 60 countries worldwide, from which the most productive collaborations (represented by links known as edges) are with United States, Spain, United Kingdom, Germany, and Australia (Figure 2(a, b)). The strongest network of collaboration was evidenced between United States with Mexico ($n = 112$ collaborations), Brazil ($n = 86$), Colombia ($n = 44$), Chile ($n = 21$), and Ecuador ($n = 21$); followed by the interactions of Brazil with United Kingdom ($n = 23$) and Germany ($n = 21$); as well as between Spain and Mexico ($n = 28$), Brazil ($n = 15$), Argentina ($n = 11$), Ecuador ($n = 11$), Colombia ($n = 8$), and Chile ($n = 8$; Figure 2(b)). Further, the 10 countries that represented the nodes with the highest centrality degree (with collaborations with more than 30 countries) between Latin America and the rest of the world are, in descending order, United States, Brazil, Australia, United Kingdom, Mexico, Colombia, Spain, Germany, France, and Canada (Figure 2(a, b)).

The top 10 of the most prolific Latin American countries are Brazil ($n = 356$), Mexico ($n = 350$), Argentina ($n = 145$), Colombia ($n = 103$), Chile ($n = 57$), Ecuador ($n = 39$), Costa Rica ($n = 20$), Peru ($n = 20$), Uruguay ($n = 19$), and Bolivia ($n = 18$; Figure 2(a, b)). However, when considering the number of published documents per country population density (population/km^2), a new ranking pattern appears. In descendant order, the top 10 prolific countries are Brazil, Argentina, Mexico, Colombia, Chile, Bolivia, Uruguay, Peru, Ecuador, and Venezuela. In contrast, when considering the number of published documents per country population, the ranking pattern changes: Uruguay, Costa Rica, Argentina, Chile, Mexico, Guyana, Panama, Ecuador, Colombia, and Suriname.

The most robust collaboration edges between Latin American countries were evidenced between Brazil with Argentina ($n = 18$), Mexico ($n = 17$), and Colombia ($n = 11$) and between Mexico with Colombia ($n = 17$), Ecuador ($n = 10$), and Argentina ($n = 8$); these collaborations represent just between 2.2% and 10.7% of the country's publications. From Central American institutions, Costa Rica, Panama, and Nicaragua were the most prolific with 60, 27, and 11 published documents, respectively, followed by Honduras, Guatemala, and El Salvador (with 7, 6, and 2 publications, respectively). From the Caribbean islands, Cuba and Puerto Rico were the most prolific with 16 and 5 publications, respectively, followed by Jamaica, Antigua and Barbuda, Dominican Republic, and Haiti (with 3, 3, 2, and 1 publications, respectively; Figure 2(a)).

From an institutional perspective, the global collaboration network was composed of 4,329 institutions, from which 51 published more than 34 papers each (mean 61.6), shaping a major component of the network (Figure 3(a)). Top 10 institutions included two from Latin America and, in descending order of production, were National Autonomous University of Mexico (UNAM), Chinese Academy of Sciences, University of Kansas, U.S. Geological Survey (USGS), Consejo Superior de Investigaciones Científicas, Federal University of Goiás (Brazil), University of Melbourne, University of Porto, U.S. Forest Service, and Colorado State University. Based on the amount of productivity, it was possible to identify that the UNAM had the highest centrality degree among institutions worldwide, being the most prolific institution, with 199 published papers, collaborating with other 225 institutions (38.2% from Latin American countries) in the network. In contrast, the second most prolific institution, the Chinese Academy of Sciences, collaborated with 17 institutions but published 70% of the documents in collaboration with a single institution (University of Chinese Academy of Sciences). There was no collaboration between the UNAM with the Chinese Academy of Sciences (Figure 3(a)).

University of Kansas and USGS (USA) are in the third and fourth level of productivity, with 161 and 125 published documents, respectively. Within the major component of the global network, University of Kansas has collaborated with 28 institutions of which five are from Latin America (UNAM, Institute of Ecology A.C.—INECOL, University of São Paulo (USP), Federal University of Goiás, and University of Brasilia), while USGS collaborated with just 18 institutions showing no interaction with Latin American institutions (Figure 3(a)).

The Latin American network was composed of 556 Latin American institutions that collaborate with 519 other institutions outside Latin American countries
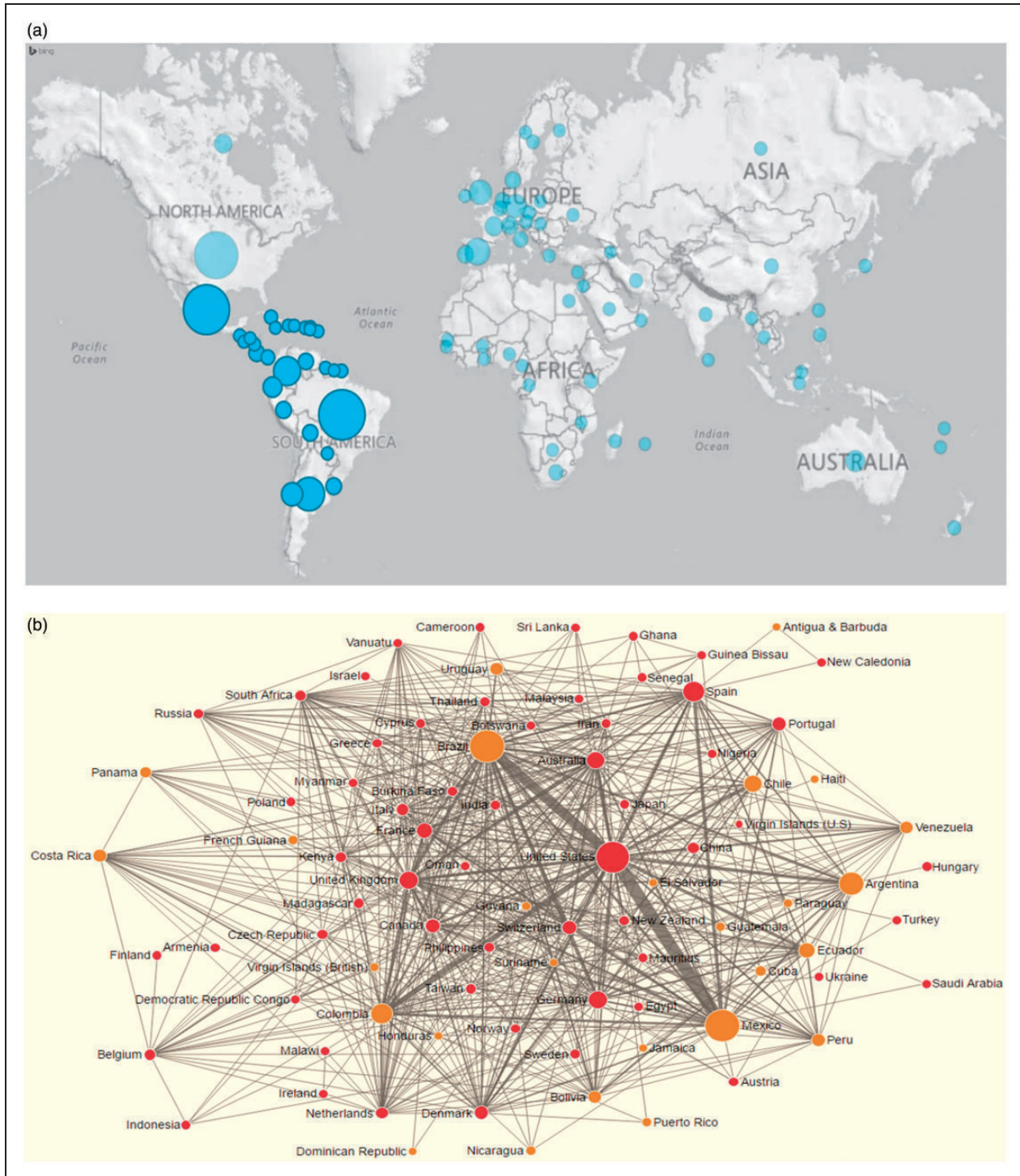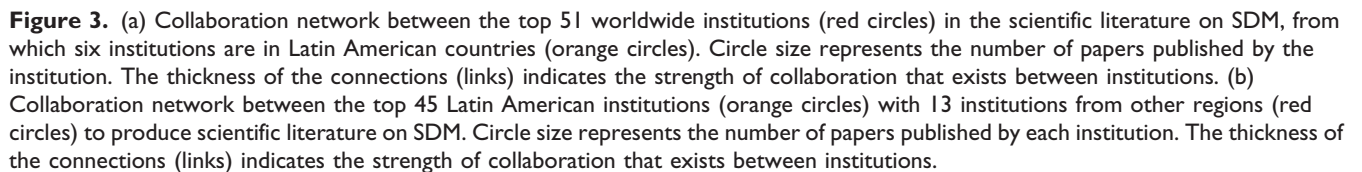
Figure 2. (a) Productivity of SDM literature in which Latin American institutions participate (blue circles) with other countries worldwide (pale blue circles). Circle size represents the number of published documents per country. (b) Collaboration network between the top 26 Latin American countries (orange circles) with other 60 countries (red circles) to produce documents on SDM in the scientific literature. Circle size represents the number of papers published by country. The thickness of the connections (links or edges) indicates the strength of collaboration that exists between countries.

Figure 3. (a) Collaboration network between the top 51 worldwide institutions (red circles) in the scientific literature on SDM, from which six institutions are in Latin American countries (orange circles). Circle size represents the number of papers published by the institution. The thickness of the connections (links) indicates the strength of collaboration that exists between institutions. (b) Collaboration network between the top 45 Latin American institutions (orange circles) with 13 institutions from other regions (red circles) to produce scientific literature on SDM. Circle size represents the number of papers published by each institution. The thickness of the connections (links) indicates the strength of collaboration that exists between institutions.

(online Appendix C). A total of 45 Latin American institutions and 13 institutions from outside, published more than 10 papers each (mean 22.8) shaping the major component of the network (Figure 3(b)). From those, the top institution is UNAM, mainly collaborating with University of Kansas, INECOL, and University of Texas at Austin (17%, 10.5%, and 6.5% of UNAM publications, respectively). The second most productive Latin American institution was Federal University of Goiás collaborating with 24 institutions, mainly with State University of Goias, Consejo Superior de Investigaciones Científicas, Federal University of Paraná, UNAM, and National Institute of Amazonian Research (between 5 and 7 shared papers). The other four most prolific Latin American institutions were INECOL (55 papers published with 17 institutions), National Scientific and Technical Research Council in Argentina—CONICET (52 papers published with 19 institutions), USP (51 papers published with 28 institutions), and University of Brasilia (36 papers published with 20 institutions; online Appendix C).

Thematic Characterization of Latin American Research on SDM

From the characterization of keywords reported among the 1,000 SDM papers published with collaboration by Latin American institutions, the main patterns were as follows:

1. Taxonomic group: From the 25 taxonomic groups of species modeled (from Kingdom to Class or Division), the most frequent were Chordata (45%), Insecta (17%), and Magnoliophyta (7%; Figure 4(a)).
2. Realm of application: From 30 types of thematic areas of research, 70% of the keywords mainly were related with studies in the following topics: Conservation planning and management, Climate change, Species conservation, Epidemiology, Evolutionary biology, and Biological invasions. Each category represents a contribution between 4% and 13% of the categorized words (Figure 4(b));
3. Geographic region of the SDM: From the 25 types of geographic regions, the most frequent were South America and Mesoamerica with contributions of 37% and 28% categorized words, respectively (online Appendix D). The most frequent countries were Mexico (29%), Brazil (14%), Argentina (6%), and Colombia (5%; online Appendix D).
4. Ecosystem type of the SDM: From the 17 types of ecosystems, the most frequent were rainforest, mountain ecosystems, and tropical dry forest with contributions between 14% and 29% of the categorized words (online Appendix E).

Modeling Methods Used to Develop SDM in Latin American Research

From the detailed reading (title, abstract, and keywords) of the 1,000 papers in which Latin American institutions collaborated, we identified 673 papers (excluding 14 review papers that did not conduct any SDM modeling and were excluded from the bibliometric mapping) that used 36 methods (in three categories) to model species distributions (see online Appendix F). The first modeling method to model distributions in which Latin American institutions participated, was GARP (Sánchez-Cordero & Martínez-Meyer, 2000; Soberón, Golubov, & Sarukhán, 2001).

It is well recognized that there is no guarantee that a single modeling method will be optimal to model species distributions for all data sets and geographical realms (Mateo, Felicísimo, & Muñoz, 2011). However, It was only after 2006 that researchers began to compare SDM with more than two modeling methods (generalized additive models, GARP, BIOCLIM, generalized linear models—GLM, MaxEnt, DOMAIN, FloraMap, Weights of evidence, Mahalanobis Distance and Random Forests, among others; online Appendix G). Until 2018, only 50 articles (7.4%) compared between 3 and 10 modeling methods, from which only 6 compared more than 7 methods (Escobar, Qiao, Cabello, & Peterson, 2018; Ochoa-Ochoa, Flores-Villela, & Bezaury-Creel, 2016; Queiroz et al., 2013; Reiss, Cunze, Konig, Neumann, & Kroncke, 2011; Tessarolo, Rangel, Araújo, & Hortal, 2014; Tognelli, Roig-Juñent, Marvaldi, Flores, & Lobo, 2009; online Appendix F, G).

We found that 73.5% of the analyses used MaxEnt, followed by GARP (18.7%), GLM (6.4%), BIOCLIM (4.6%), Random Forests (5%), and generalized additive models (3.1%; Figure 5(a)). The most common combination of performed analyses are MaxEnt models with GARP ($n = 47$), BIOCLIM ($n = 22$), GLM ($n = 17$), and Random Forests ($n = 16$; Figure 5(a), online Appendix F, G). Our results highlight MaxEnt as the most commonly used software to model the distribution of species from Latin American institutions. This finding is further supported by MaxEnt download data showing that between December 12, 2016, and June 8, 2018, 178 days after the new open source version of MaxEnt became available (Phillips et al., 2017), 27,472 downloads were made from 156 countries, of which 19.7% were made in the United States, 9.2% in Mexico, 7.5% in Brazil, 6% in China, and 5.5% in Colombia (Figure 5 (b); data from the American Museum of Natural History). Latin American countries made 8,861 downloads, 32.2% of total downloads, during this period. Also, after controlling for the population size of countries, 5 Latin American countries emerged among the top 10 countries with the highest number of downloads

Table 1. Top Five Most Prolific Countries in Different Subject Areas and Categories of Knowledge.

All subject areas[a]	Biological sciences[a]	Ecology, evolution, behavior, and systematics[a]	Animal sciences[a]	Plant sciences[a]	SDM of invasive species (Barbosa et al., 2012)	SDM worldwide (current study)
United States	United States	United States	United States	United States	United States	United States
China	China	United Kingdom	Brazil	China	Australia	United Kingdom
United Kingdom	United Kingdom	China	United Kingdom	Germany	New Zealand	China
Germany	Germany	Germany	Germany	Japan	Spain	Spain
Japan	Japan	Canada	Canada	India	South Africa	Germany

Note. SDM = species distribution modeling.
[a]Based upon SCImago SJR country ranking of productivity from 1996 to 2017 (SCImago, 2019).

per million people during this period: Costa Rica, French Guiana, Colombia, Ecuador, and Chile.

Discussion

According to our results, research on SDM started in the early 90s and evolved with more than 5,533 published papers and 35 modeling tools used as of 2018 (Figure 1 (a)). We propose the following strategies to boost the generation of new knowledge on SDM, to reduce the Wallacean shortfall and to overcome this great challenge of environmental degradation within the Latin American countries.

Interaction Opportunities Visualized From Institutional Collaboration Patterns

We conclude that the United States is the country with the highest scientific production in SDM, as it is for other scientific areas as shown in Table 1 (SCImago, 2019). However, the United States is important for collaborations between countries as well, but our study presents a different per country publication pattern than the one reported by Barbosa et al. (2012) for SDM of invasive species (Table 1).

Countries in Latin America collaborate in 18% of the global production of SDM academic publications including the development of new modeling tools (Figure 1(a)). Brazil and Mexico are among the top 10 countries worldwide and become a reference at the international level, especially among other Latin American countries. This correlates with the results of the SCImago Journal & Country Rank, a portal that provides scientific indicators to assess scientific domains, where at a global level Brazil is in the top 15 and Mexico in the top 29 for academic production in all subject areas category (Table 1; SCImago, 2019).

Within the Latin American region, the most productive institutions on SDM is the UNAM, which is also number one worldwide, followed by Federal University of Goiás, USP, and University of Brasilia, those

institutions are ranked by SCImago (2019) in Latin America as the 3rd, 45th, 1st, and 25th positions, respectively.

The geographic scope of the SDM research in Latin America coincides with the four more scientifically productive countries Mexico, Brazil, Argentina, and Colombia (Figure 2(a, b); online Appendix D). Although the collaboration network in Latin America is complex, and involves 12 countries, research in the region might be reducing the Wallacean shortfall only on those countries. It would be desirable that future collaboration reduces the Wallacean shortfall not only in the most productive countries but expands to other less studied geographic areas.

It is interesting how the number of collaborations with other institutions has a contrasting pattern between the top 1 and 2 global ranking institutions. While UNAM collaborated with 225 other institutions, the Chinese Academy of Sciences collaborates with only 17 other institutions. This suggest that in contrast to China, collaboration is a crucial issue for ranking universities in Latin America. The Chinese Academy of Sciences is the second institution that globally has the most publications on SDM. Based on its research performance, innovation outputs and societal impact, the Chinese Academy of Sciences ranks first among global institutions as measured by their web visibility (SCImago, 2019). However, in Latin America, this institution only collaborates with Guatemala and Chile (online Appendix C).

The implementation of collaboration strategies between Latin American institutions and those outside Latin America should continue to be enhanced through fellowships, capacity building, sabbatical years between researchers, or shared advisors of postgraduate students. Latin American collaboration networks will tend to grow as more researchers return to their country of origin, maintaining collaboration networks created from postgraduate studies and fellowships, and beginning new collaborations with other Latin American countries.

Closing the Gap From Software User to Coauthor: A MaxEnt Study Case

MaxEnt is the most commonly used software to model the distribution of species from Latin American institutions (Figure 5(a)). The first nine documents published, by Latin American institutions, using MaxEnt software were from 2007 to 2008 (online Appendix F, G), between 1 and 2 years after the release of the first version of the software (Phillips, Anderson, & Schapire, 2006; Figure 1 (a)). MaxEnt users see this tool as one of those that requires the fewest actions needed to implement the analysis being at the top among modeling tools in terms of usefulness, learning curve, system capabilities, and overall user satisfaction (Ahmed et al., 2015). As mentioned earlier, after the new open source version of MaxEnt was made available (Phillips et al., 2017), it was followed by an enormous number of downloads (Figure 5(b)); 32.3% of which took place in Latin America.

Also, half of the top 10 countries downloading the software (when controlling for population size) were in Latin America. So, if Latin America is indeed powerful within the global community of MaxEnt users, why do Latin American authors and institutions represent only 18% of the global scientific literature on SDM? Looking deeper into download trends, we can examine the 13,155 or 48% of global downloads that included optional institution information. Of this number, 3,385 or 38.2% of total downloads from Latin America (8,861) included institutional information. Of those, the vast majority (2,944 or 87%) of MaxEnt downloads were from academic institutions (universities or university centers/institutes). Downloads from government institutions including research institutes, agencies, national park system administrations, and ministries comprised 384 or 11%, nongovernmental organizations (NGOs) including conservation organizations, private not-for-profit research institutes, and foundations accounted for 102 or 3%. Only 5 downloads or 0.1% were from for-profit companies including consultants.

As mentioned earlier, the most prolific Latin American institutions in global scientific literature were universities (online Appendix C). The software download information also reflects this trend, but the number of MaxEnt downloads from government institutions is also notable, which is not as evident in the observed scientific literature trends. We might assume that if we examined beyond the scientific literature to include white papers and government reports, we might find additional written contributions from these institutions and probably several examples of translations of SDM to decision-making related to biodiversity conservation at NGOs and government agencies (see some examples later). We can also probably assume

that there is a great deal of SDM development and research ongoing at universities that is not being published in the global scientific literature, or that is being used for training only. This research may be available instead as student theses or reports in university databases. There may also be further challenges to publishing results such as costs, language, and other factors.

Implementing Best Practices and Integrating SDM Into Monitoring Biodiversity Change

Strengthening of collaborative networks between academic institutions and institutions outside of academia in Latin America would further enhance the translation of SDM research to biodiversity conservation decision-making (as discussed later) and would likely increase the visibility of ongoing and future research using SDM in the global scientific literature. Open source SDM training opportunities and online networks in the Spanish language (such as the Colombian initiative BioModelos—see later—and more informally, the Facebook group "Modelado de Nichos Ecológicos y Distribuciones Geográficas") are a robust and well-appropriated strategy to facilitate online training and the formation of communities of practice across researchers and practitioners from different sectors and different Latin American countries (Cuervo-Robayo et al., 2017).

We identified that some Latin American countries have great potential to become world leaders in the area of SDM, to the extent that the collaboration between institutions increases and diversifies. As SDM research areas increase in Latin America, good practices must also acquire by researchers, input data should be open access, codes should be shared with complete documentation, and prediction layers should be made available (Breiman, 2001; Guisan & Thuiller, 2005; Humphries & Huettmann, 2018). Latin American countries are moving into open science culture, the number of open access mandates adopted has grown from 1 in 2005 to 54 in 2018 (ROARMAP http://roarmap.eprints.org/, last visit on April 18, 2019).

Megadiverse countries in Latin America, have important reservoirs of mining resources such as gold, nickel, chromium, and copper, among others (USGS, 2018). Those resources attract higher pressure from big companies, that together with oil and timber extraction, push governments to give priority to extractive activities over preservation of biodiversity (Loyola, 2014; Villarroya et al., 2014). Under such scenarios, SDM could provide critical information to make a difference in the decision-making process by improving the understanding and monitoring of biodiversity, as well as focusing spatial conservation strategies and their effective implementation by responsible authorities (Scheldeman & van

Zonneveld, 2010); for example, by combining SDMs with other key information as a part of a systematic conservation planning framework (e.g., as in Marxan: Ball, Possingham, & Watts, 2009; or Zonation: Moilanen et al., 2005). Enhancing the training of new generations of biologists and ecologists in computational techniques is needed so that machine learning methods can play an essential role in understanding whole ecosystems by holistic modeling (Humphries & Huettmann, 2018).

New technology can easily be misused when they are not correctly applied, as explicitly seen for SDM (Guillera-Arroita et al., 2015; Humphries & Huettmann, 2018) and model quality should be evaluated (Araújo et al., 2019). SDM are particularly prone to problems arising from a mismatch between data type and intended purpose. Hence for conservation practice is critical to know if the output from an SDM will be appropriate for the intended application (Guillera-Arroita et al., 2015). The majority of the SDM produced in Latin America comes from presence-background data, meaning that their outputs represent in most cases relative likelihoods of observation, with subsequent limitations in their applications. As such Latin America needs to move toward the collection of more informative survey data of the occupancy-detection type so that better inference of species distribution can be made (Guillera-Arroita et al., 2015).

SDM in Latin America have been used in multiple realms of application but with a considerable bias toward vertebrates and insects, and in three main ecosystems (rainforests, tropical dry forests, and mountain ecosystems). Researchers in Latin America have explored multiple applications for SDM; we encourage that this diverse level of application continues in an interdisciplinary manner. Is also important to notice that one of the most threatened ecosystems worldwide (tropical dry forest: Hoekstra, Boucher, Ricketts, & Roberts, 2005) is also one of the best represented in SDM research (online Appendix E). Also, the most critical global change drivers, land use change, biological invasions, and climate change (Sala et al., 2000) are among the most studied thematic from SDM perspective (Figure 4(b)). Those results show indeed the awareness by the research community of the importance of SDM for conservation in the regions.

For research results to have a higher incidence in the decision-making process, the importance of SDM needs to be also understood by decision-makers. There is no clear sustainable future in some megadiverse neotropical countries in which biodiversity is not a priority for central governments that are committed with leading economy countries (such as China and US), while local poverty increases in some regions, in which illegal drug business and big companies exhibit local power on land

decisions (Brocket, 1990; Huettmann, 2015). As noticed by Humphries and Huettmann (2018), SDM study is rarely linked to effective conservation management, as they are not referred to in most policy or legal decisions, and SDM studies widely lack a reflective component that advances ethical and societal questions. How can we move to translate scientific results into effective conservation management? We propose that in Latin America, SDM must be placed in a broader context between monitoring data and specific management and conservation contexts. New biodiversity monitoring frameworks may provide robust tools to articulate user needs with scientific outputs, Biodiversity Observation Networks could be developed in Latin America and facilitate explicit knowledge to transit into tacit knowledge, addressing societal and economic needs, and thus increase research incidence on conservation (Navarro et al., 2017).

Building Communities of Practice Around SDM: The Case of BioModelos in Colombia

SDM ideally must rely on enough occurrence data as well as knowledge regarding species' environmentally limiting factors and dispersal limitations (Anderson, 2012; Araújo & Peterson, 2012). Although occurrence data are increasingly available in open access platforms such as Global Biodiversity Information Facility (Meyer et al., 2015), albeit, with questionable quality (Anderson et al., 2016), knowledge on species ecology and evolution is also insufficient but crucial for the development of biologically meaningful models (Diniz-Filho, Loyola, Raia, Mooers, & Bini, 2013). To remedy this shortfall, the Colombian initiative BioModelos brings together modelers with experts who can inform model development and assess the reliability of models (Velásquez-Tibatá et al., 2019). We present this initiative here as an example of collaborative knowledge building for improving SDM.

In BioModelos, experts are arranged into groups according to thematic areas of interest, and each group has a defined set of species of interest. Experts contribute through an online platform (biomodelos.humboldt.org.co) to any of the following activities for each species in their group: (a) occurrence data cleaning; (b) delineation of accessible area; (c) identification of suitable land covers; (d) identification of areas of model over/under-prediction; (e) selection of suitability thresholds to create binary models; and (f) qualitative evaluation of model accuracy. A core team develops and edits models according to expert inputs. Once this collaborative modeling process is completed, models become available for visualization and download in standard GIS formats along with metadata documenting the modeling process on the BioModelos webpage. Importantly, the metadata
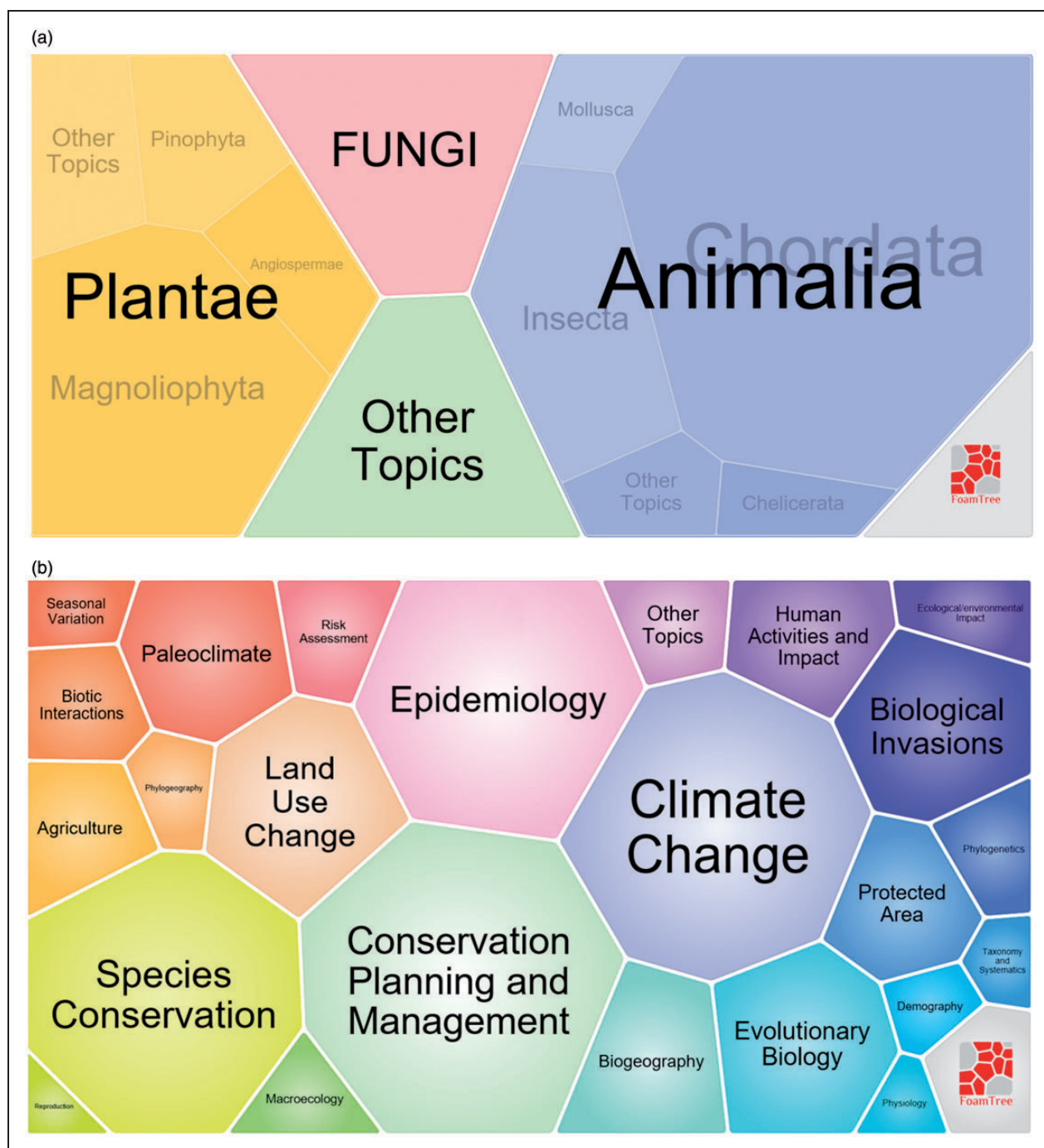
Figure 4. Carrot Search FoamTree visualization of pertinent words classified by categories: (a) taxonomic group and (b) realm of application.

acknowledges the contribution of everyone involved in model development.

Thus far, 475 experts have joined at least one of the 20 expert groups in BioModelos, where they contribute to the development of models for 980 species (Velásquez-Tibatá et al., 2019). These models have been particularly crucial for the Colombian government in the evaluation of species extinction risk (Renjifo, Amaya-Villareal, Burbano-Girón, & Velásquez-Tibata, 2016) and the development of national conservation action plans (e.g., cycads and primates: http://www.minambiente. gov.co/index.php/bosques-biodiversidad-y-servicios-eco sistematicos/fauna-y-flora/programas-de-conservacion# documentos). Along those assessments and plans,
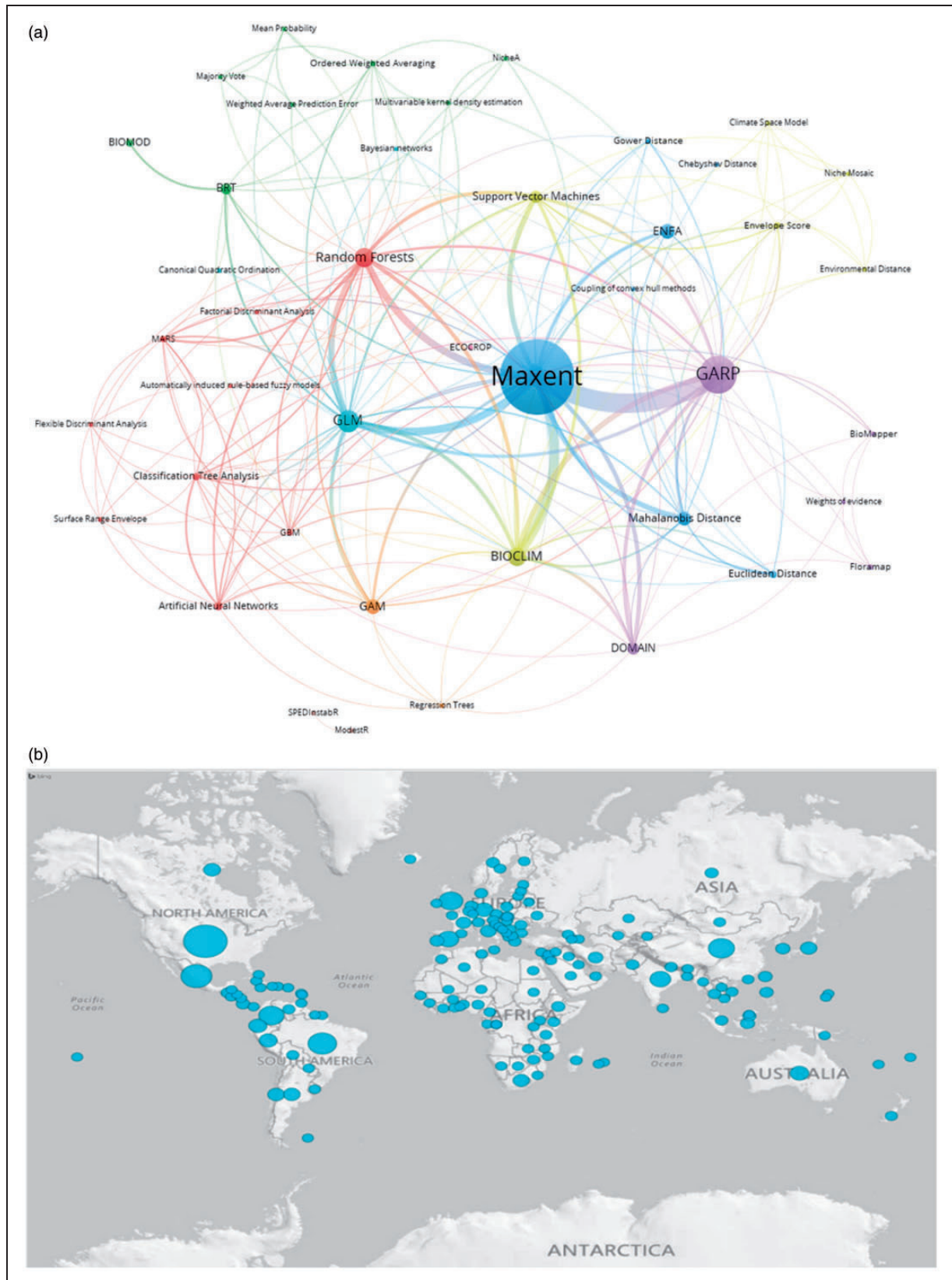
Figure 5. (a) Network of modeling methods used to develop SDM in documents including collaborators from Latin American institutions. (b) Global downloads of MaxEnt software from December 12, 2016, to June 8, 2018 ($n = 27,472$; data from the American Museum of Natural History). Circle size represents the number of downloads per country, in 178 days.
ENFA = ecological niche factor analysis; GARP = Genetic Algorithm for Rule-Set Prediction; GLM = generalized linear models; GAM = generalized additive models.

qualitative assessments of the biological realism of spatial predictions by qualified experts are essential to minimize costly commission errors (Guisan et al., 2013; Lobo, Jiménez-Valverde, & Hortal, 2010) that cannot be measured unbiasedly with presence-only data (Lobo, Jiménez-Valverde, & Real, 2008). However, model users also report their use for basic research (33 %) and educational activities (32%) and varied activities such as environmental impact assessments, bioprospecting, and applied research (Velásquez-Tibatá et al., 2019).

The implementation of BioModelos in Colombia shows that it is possible to go a long way toward addressing data incompleteness and quality issues by involving experts in model development. By using transparent workflows and making data freely available, this approach contributes to diminishing the duplication of efforts in data cleaning and modeling, thereby accelerating the use of modeling outcomes in basic and applied research. Building SDM in a collaborative way is the first step to create communities of practice that discuss and overcome the different caveats of SDM. A next step is to integrate stakeholders from governmental institutions, NGOs, and productive sectors to develop derived products from SDM that can be directly incorporated into decision-making.

Implications for Conservation

SDM is used to support decision-making in Latin America; however, we must increase the incidence of SDM derived products for decision-making. Boundary institutions between science and academia, such as CONABIO in Mexico (Koleff & Urquiza-Hass, 2011; Ochoa-Ochoa, Urbina-Cardona, Flores-Villela, Vázquez, & Bezaury-Creel, 2009; Urbina-Cardona & Flores-Villela, 2010), or Conservation International and Instituto Alexander von Humboldt in Colombia (Conservación Internacional-Colombia & Secretaría Distrital de Ambiente, 2010; Londoño-Murcia, Zárate, & Ruiz-Agudelo, 2011a; Londoño-Murcia et al., 2011b), have invested in the development of SDM as a critical input for (sub)national assessments of biodiversity patterns, conservation gap analysis, and conservation planning, as well as for the development of official cartography such as the National Ecosystem Map of Colombia and their use in biodiversity offsets (Ministerio de Ambiente y Desarrollo Sostenible, 2018) and species extinction risk assessments (Renjifo et al., 2016).

Non-governmental and other institutions in Latin America have used SDM for supporting species conservation programs and land-use planning. This is the case in Brazil, for example, in which the Ministry of Environment partners with the Chico Mendes Institute for Biodiversity Conservation to use SDMs to guide and propose actions to monitor and protect threatened species (in species recovery plans: Ferraz et al., 2012) and to identify priority areas for conservation and sustainable use of Brazil's biodiversity (e.g., Ministério do Meio Ambiente, 2007). Use of SDM in Latin America has focused on the conservation of native species, but there are other areas such as control of invasive species (e.g., Sales et al., 2017), agriculture, and valuation of ecosystem services (Manhães et al., 2018) or ecological restoration (Zwiener et al., 2017), where SDM is still mainly an academic interest.

Although highly recognized, the potential of SDM for guiding field surveys (Raxworthy et al., 2003) so that data gaps are reduced has not yet been fully realized in Latin America to guide national monitoring or inventory schemes, perhaps due to persistent data limitations. In many cases, field work is done opportunistically, leading to some questioning: Is the use of SDM to guide field expeditions limited by a lack of understanding, a gap in communication of results between academia and government, or lack of confidence because of inherent uncertainties and bias? Alternatively, is there a deeper reason related to traditional institutional cultures? We argue that all issues are going on simultaneously.

To use data science products more often in decision-making, both scientists and practitioners may need to shift paradigms. Scientists should be comfortable and transparent when communicating the uncertainty and bias of their products, and not be afraid of delivering the information as the best knowledge that they have available. They need to be aware that decisions will not wait for perfect science; many decisions will be taken with or without scientific evidence, and as such may benefit from the best information available on species distributions at the time of the decision-making process. On the other hand, practitioners should understand that science is a process and derived products are not perfect. As proposed by Redford, Groves, Medellín, and Robinson (2012), scientists and decision-makers should build knowledge together and exchange information on how science works and how real-life decisions are taken so that suitable products are generated on time to fit user needs. In this sense, we propose that communities of practice around SDM could integrate different visions from academy and government, increasing the credibility and use of models in decision-making scenarios.

Final Remarks

The results of this paper provide clear guidelines for initiatives such as BioBridge that aim to enhance the technical cooperation between parties of the Convention on Biological Diversity, and the International Union of Biological Sciences, where one of the objectives is to initiate, facilitate, and coordinate research, capacity

building, and other scientific activities that involve international and interdisciplinary cooperation.

Although the scientific community is aware of data gaps, such as the Wallacean or Linnaean shortfall (Lomolino, 2004), scientists on their own will not be able to address these gaps, as it requires shared effort from governments and politicians to invest in data acquisition, management, and integration at national levels. The implementation of information infrastructures, long-term monitoring, and citizen science data is a priority in Latin America. Biodiversity and Ecological Informatics research fields need further support, for example, through new, targeted graduate academic programs in Latin America so that research and collaboration networks can evolve. Upscaling results from scientific journals to decision-makers is an urgent need.

We hope that data gaps and the use of modeling techniques to address information for supporting decision-making have a place in the CBD post-2020 agenda. Strategic partnerships within Latin American countries between academia and governments can also help to establish and scale-up the need for and the high potential use of SDM in decision-making.

Acknowledgments

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

ORCID iD

Nicolás Urbina-Cardona https://orcid.org/0000-0002-4174-8467

References

Ahmed, S. E., McInerny, G., O'Hara, K., Harper, R., Salido, L., Emmott, S., & Joppa, L. N. (2015). Scientists and software – Surveying the species distribution modelling community. *Diversity and Distributions*, 21, 258–267.

Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, 38, 541–545.

Anderson, R. P. (2012). Harnessing the world's biodiversity data: Promise and peril in ecological niche modeling of species distributions. *Annals of the New York Academy of Sciences*, 1260, 66–80.

Anderson, R. P., Araújo, M. B., Guisan, A., Lobo, J. M., Martínez-Meyer, E., Peterson, A. T., & Soberón, J. (2016). *Final report of the task group on GBIF data fitness for use in distribution modeling. Ver 1.1 (29 p)*. New York, NY: Global Biodiversity Information Facility.

Araújo, M. B., Anderson, R. P., Barbosa, A. M., Beale, C. M., Dormann, C. F., Early, R., ... Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. *Science Advances*, 5, eaat4858.

Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, 93, 1527–1539.

Baldwin, R. A. (2009). Use of maximum entropy modeling in wildlife research. *Entropy*, 11, 854–866.

Ball, I. R., Possingham, H. P., & Watts, H. (2009). Marxan and relatives: Software for spatial conservation prioritization. In Moilanen A., Wilson K. A., & Possingham H. P. (Eds.), *Spatial conservation prioritization: Quantitative methods and computational tools* (Chapter 14, pp. 185–195). Oxford, England: Oxford University Press.

Barbosa, A. M., Brown, J. A., Jimenez-Valverde, A., & Real, R. (2016). *modEvA: Model evaluation and analysis. R package version 1.3.2.* Retrieved from https://CRAN.R-project.org/package=modEvA

Barbosa, F. G., Schneck, F., & Melo, A. S. (2012). Use of ecological niche models to predict the distribution of invasive species: A scientometric analysis. *Brazilian Journal of Biology*, 72, 821–829.

Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92, 1170–1182.

Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, 16, 199–231.

Brewer, M. J., O'Hara, R. B., Anderson, B. J., & Ohlemüller, R. (2016). Plateau: A new method for ecologically plausible climate envelopes for species distribution modelling. *Methods in Ecology and Evolution*, 7, 1489–1502.

Brocket, C. (1990). *Land, power and poverty* (2nd ed.). New York, NY: HarperCollins Publishers.

Brown, J. L., Bennett, J. R., & French, C. M. (2017). SDMtoolbox 2.0: The next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ*, 5, e4095.

Busby, J. R. (1993). BIOCLIM – A bioclimate analysis and prediction system. *Plant Protection Quarterly*, 6, 8–9.

Cánovas, F., Magliozzi, C., Mestre, F., Palazón, J. A., & González-Wangüemert, M. (2016). ENiRG: R-GRASS interface for efficiently characterizing the ecological niche of species and predicting habitat suitability. *Ecography*, 39, 593–598.

Carl, G., & Kühn, I. (2017). Spind: A package for computing spatially corrected accuracy measures. *Ecography*, 40, 675–682.

Carpenter, G., Gillison, A. N., & Winter, J. (1993). Domain: A flexible modeling procedure for mapping potential distributions of animals and plants. *Biodiversity and Conservation*, *2*, 667–680.

Carrot Search FoamTree. (2019). Retrieved from https://carrot search.com/foamtree/

Cayuela, L., Golicher, D. J., Newton, A. C., Kolb, M., de Alburquerque, F. S., Arets, E. J. M. M., . . . Pérez, A. M. (2009). Species distribution modeling in the tropics: Problems, potentialities, and the role of biological data for effective species conservation. *Tropical Conservation Science*, *2*, 319–352.

Conservación Internacional-Colombia, & Secretaría Distrital de Ambiente. (2010). *Ecological Connectivity Guidelines for Bogota. Association Contract 01141 de (2009) (234 p)*. Bogotá, Colombia: Author.

Cuervo-Robayo, A. P., Escobar, L. E., Osorio-Olvera, L. A., Nori, J., Varela, S., Martínez-Meyer, E., . . . Peterson, A. T. (2017). Introduction to Spatial Analysis with Emphasis on Ecological Niche Models. *Biodiversity Informatics*, *12*, 45–57.

Dedman, S., Officer, R., Clarke, M., Reid, D. G., & Brophy, D. (2017). Gbm.auto: A software tool to simplify spatial modelling and Marine Protected Area planning. *Plos One*, *12*, e0188955.

de Souza Muñoz, M. E., De Giovanni, R., de Siqueira, M. F., Sutton, T., Brewer, P., Scachetti-Pereira, R., . . . Perez-Canhos, V. (2011). openModeller: A generic approach to species' potential distribution modelling. *Geoinformatica*, *15*, 111.

Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F. T., D'Amen, M., Randin, C., . . . Guisan, A. (2017). ecospat: An R package to support spatial analyses and modeling of species niches and distributions. *Ecography*, *40*, 774–787.

Diniz-Filho, J. A. F., Bini, L. M., Fernando Rangel, T., Loyola, R. D., Hof, C., Nogués-Bravo, D., & Araújo, M. B. (2009). Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography*, *32*, 897–906.

Diniz-Filho, J. A. F., Loyola, R. D., Raia, P., Mooers, A. O., & Bini, L. M. (2013). Darwinian shortfalls in biodiversity conservation. *Trends in Ecology & Evolution*, *28*, 689–695.

Duan, R., Kong, X., Huang, M., Wu, G., & Wang, Z. (2015). SDMvspecies: A software for creating virtual species for species distribution modelling. *Ecography*, *38*, 108–110.

Engler, R., Hordijk, W., & Guisan, A. (2012). The MIGCLIM R package – Seamless integration of dispersal constraints into projections of species distribution models. *Ecography*, *35*, 872–878.

Escobar, L. E., Qiao, H., Cabello, J., & Peterson, A. T. (2018). Ecological niche modeling re-examined: A case study with the Darwin's fox. *Ecology and Evolution*, *8*, 4757–4770.

Ferraz, K. M. P. M. B., Ferraz, S. F. B., Paula, R. C., Beisiegel, B., & Breitenmoser, C. (2012). Species distribution modeling for conservation purposes. *Natureza & Conservação*, *10*, 214–220.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*, 4302–4315.

Franklin, J. (2010). Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, *16*, 321–330.

Freeman, L. C. (1977). A set of measure of centrality based on betweenness. *Sociometry*, *40*, 35–41.

Freeman, E. A., & Moisen, G. (2008). PresenceAbsence: An R package for presence absence analysis. *Journal of Statistical Software*, *23*, 11.

Golding, N., August, T. A., Lucas, T. C. D., Gavaghan, D. J., Loon, E. E., & McInerny, G. (2018). The zoon r package for reproducible and shareable species distribution modelling. *Methods in Ecology and Evolution*, *9*, 260–268.

Guillaumot, C., MarHn, A., Eléaume, M., & Saucède, T. (2017). *SDMPlay: Species distribution modelling playground. R package version 1.0*. Retrieved from https://cran.r—project.org/web/packages/SDMPlay/

Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P.E., . . . Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, *24*(3), 276–292.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., . . . Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*, 1424–1435.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, *8*, 993–1009.

Guo, Q., & Liu, Y. (2010). ModEco: An integrated software package for ecological niche modeling. *Ecography*, *33*, 637–642.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978.

Hijmans, R. J., & Elith, J. (2013). *dismo: Species distribution modeling with R*. Retrieved from http://cran.r-project.org/web/packages/dismo/index.html

Hirzel, A. H., Hausser, J., Chessel, D., & Perrin, N. (2002). Ecological-niche factor analysis: How to compute habitat suitability maps without absence data? *Ecology*, *83*, 2027–2036.

Hoekstra, J. M., Boucher, T. M., Ricketts, T. H., & Roberts, C. (2005). Confronting a biome crisis: Global disparities of habitat loss and protection. *Ecology Letters*, *8*, 23–29.

Huettmann, F. (2015). *Central American biodiversity: Conservation, ecology, and a sustainable future (805 pp)*. New York, NY: Springer.

Humphries, G. R., & Huettmann, F. (2018). A Perspective on the Future of Machine Learning: Moving Away from 'Business as Usual' and Towards a Holistic Approach of Global Conservation. In *Machine Learning for Ecology and Sustainable Natural Resource Management* (pp. 411–430). Springer, Cham.

Jones, P. G., & Gladkov, A. (1999). *FloraMap, version 1: A computer tool for predicting the distribution of plants and other organisms in the wild*. Cali, Colombia: Centro Internacional de Agricultura Tropical.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., ... Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Nature – Scientific Data*, 4, 170122.

Kass, J. M., Vilela, B., Aiello-Lammens, M. E., Muscarella, R., Merow, C., & Anderson, R. P. (2018). Wallace: A flexible platform for reproducible modeling of species niches and distributions built for community expansion. *Methods in Ecology and Evolution*, 9, 1151–1156.

Kearney, M. R., Isaac, A. P., & Porter, W. P. (2014). microclim: Global estimates of hourly microclimate based on long-term monthly climate averages. *Nature – Scientific Data*, 1, 140006.

Koleff, P., & Urquiza-Haas, T. (2011). *Planeación para la conservación de la biodiversidad terrestre en México: Retos en un país megadiverso*. México D.F., México: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad–Comisión Nacional de Áreas Naturales Protegidas.

Kriticos, D. J., Webber, B. L., Leriche, A., Ota, N., Macadam, I., Bathols, J., & Scott, J. K. (2012). CliMond: Global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution*, 3, 53–64.

Leroy, B., Meynard, C. N., Bellard, C., & Courchamp, F. (2016). virtualspecies, an R package to generate virtual species distributions. *Ecography*, 39, 599–607.

Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, 33, 103–114.

Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2008). AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17, 145–151.

Lomolino, M. V. (2004). Conservation biogeography. In Lomolino M. V. & Heaney L. R. (Eds.), *Frontiers of biogeography: New directions in the geography of nature* (pp. 293–296). Sunderland, MA: Sinauer Associates.

Londoño-Murcia, M. C., Zárate, M., & Ruiz-Agudelo, C. A. (2011a). *Environmental zonation of the Ranchería River Basin*. In CORPOGUAJIRA, NATIONAL PARKS AND CONSERVATION INTERNATIONAL – COLOMBIA (Eds.), Ranchería River Basin Management Plan. Riohacha, Colombia.

Londoño-Murcia, M. L., Corzo, G., Fonseca-Tobián, C., Ramírez, W., Salamanca, B., Alcázar, C., ... Lasso, C. A. (2011b). Priority areas identification for in situ biodiversity conservation (Scale 1:100,000). In Corzo G., Londoño-Murcia M. C., Ramírez W., García H., Lasso C., & Salamanca B. (Eds.), *Environmental planning for biodiversity conservation in the operational areas of Ecopetrol S.A. located in the Magdalena Medio and the Eastern Plains of Colombia* (pp. 25–42). Bogotá D.C., Colombia: Instituto Alexander von Humboldt y Ecopetrol S.A.

Loyola, R. (2014). Brazil cannot risk its environmental leadership. *Diversity and Distributions*, 20, 1365–1367.

Loyola, R. D., Lemes, P., Faleiro, F. V., Trindade-Filho, J., & Machado, R. B. (2012). Severe loss of suitable climatic conditions for marsupial species in Brazil: Challenges and opportunities for conservation. *PLoS One*, 7, e46257.

Manhães, A. P., Loyola, R., Mazzochini, G. G., Ganade, G. Oliveira-Filho, A. T., & Carvalho, A. R. (2018). Low-cost strategies for protecting ecosystem services and biodiversity. *Biological Conservation*, 217, 187–194.

Mateo, R. G., Felicísimo, Á. M., & Muñoz, J. (2011). *Species distribution models: A synthetic review*. Revista Chilena de Historia Natural, 84, 217–240.

Mazzoni, S., Halvorsen, R., & Bakkestuen, V. (2015). MIAT: Modular R-wrappers for flexible implementation of MaxEnt distribution modelling. *Ecological Informatics*, 30, 215–221.

Meyer, C., Kreft, H., Guralnick, R., & Jetz, W. (2015). Global priorities for an effective information basis of biodiversity distributions. *Nature Communications*, 6, 8221.

Midgley, G. F., Davies, I. D., Albert, C. H., Altwegg, R., Hannah, L., Hughes, G. O., ... Thuiller, W. (2010). BioMove – An integrated platform simulating the dynamic response of species to environmental change. *Ecography*, 33, 612–616.

Ministerio de Ambiente y Desarrollo Sostenible. (2018). *Resolution 0256 By which the updating of the Biodiversity Component of the Environmental Offsets Manual and other determinations are made*. Bogotá, Colombia: Author.

Ministério do Meio Ambiente. (2007). *Priority areas for conservation, sustainable use and sharing of the benefits of Brazilian biodiversity: Update – Portaría*. Brasilia, Brazil: Author.

Moilanen, A., Franco, A. M., Early, R. I., Fox, R., Wintle, B., & Thomas, C. D. (2005). Prioritizing multiple-use landscapes for conservation: Methods for large multi-species planning problems. *Proceedings of the Royal Society of London B*, 22, 1885–1891.

Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for MaxEnt ecological niche models. *Methods in Ecology and Evolution*, 5, 1198–1205.

Naimi, B., & Araújo, M. B. (2016). sdm: A reproducible and extensible R platform for species distribution modelling. *Ecography*, 39, 368–375.

Navarro, L. M., Fernández, N., Guerra, C., Guralnick, R., Kissling, W. D., Londoño, M. C., ... Pereira, H. M. (2017). Monitoring biodiversity change through effective global coordination. *Current Opinion in Environmental Sustainability*, 29, 158–169.

Ochoa-Ochoa, L., Flores-Villela, O., & Bezaury-Creel, J. (2016). Using one vs. many, sensitivity and uncertainty analyses of species distribution models with focus on conservation area networks. *Ecological Modelling*, 320, 372–382.

Ochoa-Ochoa, L., Urbina-Cardona, J. N., Flores-Villela, O., Vázquez, L.-B., & Bezaury-Creel, J. (2009). The effects of governmental protected areas and social initiatives for land protection on the conservation of Mexican amphibians. *PLoS One*, 4, e6878.

Phillips, S., Anderson, R. P., Dudik, M., Schapire, R., & Blair, M. E. (2017). Opening the black box: An open-source release of MaxEnt. *Ecography*, 40, 887–893.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*, 231–259.

Porter, A. L., & Cunningham, S. W. (2004). *Tech mining: Exploiting new technologies for competitive advantage*. Hoboken, NJ: John Wiley & Sons, Inc.

Qiao, H., Lin, C., Ji, L., & Jiang, Z. (2012). mMWeb – An online platform for employing multiple ecological niche modeling algorithms. *PLoS One*, *7*, e43327.

Qiao, H., Peterson, A. T., Soberón, J., Campbell, L. P., Ji, L., & Escobar, L. E. (2016). NicheA: Creating virtual species and ecological niches in multivariate environmental scenarios. *Ecography*, *39*, 805–813.

Queiroz, D. L., Majer, J., Burckhardt, D., Zanetti, R., Fernandez, J. I., Queiroz, E. C., ... Anjos, N. (2013). Distribution of Glycaspis in Brazil. *Australian Journal of Entomology*, *52*, 20–30.

Rangel, T. F. L. V., B., Diniz-Filho, J. A. F., & Araújo, M. B. (2009). *BIOENSEMBLES 1.0. Software for computer intensive ensemble forecasting of species distributions under climate change*. Goias, Madrid: Evora.

Raxworthy, C. J., Martinez-Meyer, E., Horning, N., Nussbaum, R. A., Schneider, G. E., Ortega-Huerta, M. A., & Peterson, A. T. (2003). Predicting distributions of known and unknown reptile species in Madagascar. *Nature*, *426*, 837–841.

Redford, K. H., Groves, C., Medellin, R. A., & Robinson, J. G. (2012). Conservation stories, conservation science, and the role of the intergovernmental platform on biodiversity and ecosystem services. *Conservation Biology*, *26*, 757–759.

Reiss, H., Cunze, S., Konig, K., Neumann, H., & Kroncke, I. (2011). Species distribution modelling of marine benthos: A North Sea case study. *Marine Ecology Progress Series*, *442*, 71–86.

Renjifo, L. M., Amaya-Villareal, A., Burbano-Girón, J., & Velásquez-Tibata, J. (2016). *Red Book of Birds of Colombia, Volume II*. Bogotá, D.C., Colombia: Editorial Pontificia Universidad Javeriana e Instituto Alexander von Humboldt.

Rosales, J. (2008). Economic growth, climate change, biodiversity loss: Distributive justice for the global north and south. *Conservation Biology*, *22*, 1409–1417.

Sala, O. E., Chapin, F. S., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., ... Wall, D. H. (2000). Global biodiversity scenarios for the year 2100. *Science*, *287*(5459), 1770–1774.

Sales, L. P., Neves, O. V., De Marco, P., & Loyola, R. (2017). Model uncertainties do not affect observed patterns of species richness in the Amazon. *PLoS One*, *12*, e0183785.

Sánchez-Cordero, V., & Martínez-Meyer, E. (2000). Museum specimen data predict crop damage by tropical rodents. *Proceedings of the National Academy of Sciences*, *97*, 7074–7077.

Scheldeman, X., & van Zonneveld, M. (2010). *Training manual on spatial analysis of plant diversity and distribution*. Rome, Italy: Bioversity International.

Schmitt, S., Pouteau, R., Justeau, D., de Boissieu, F., & Birnbaum, P. (2017). ssdm: An r package to predict distribution of species richness and composition based on stacked species distribution models. *Methods in Ecology and Evolution*, *8*, 1795–1803.

SCImago. (2019). *SJR-SCImago Journal & Country Rank*. Retrieved from https://www.scimagojr.com/countryrank.php

Search Technology. (2018). *Vantagepoint: Turn information into knowledge*. Retrieved from https://www.thevantagepoint.com/

Soberón, J., Golubov, J., & Sarukhán, J. (2001). The importance of Opuntia in Mexico and routes of invasion and impact of Cactoblastis cactorum (Lepidoptera: Pyralidae). *The Florida Entomologist*, *84*, 486–492.

Soberón, J., Osorio-Olivera, L., & Peterson, T. (2017). Conceptual differences between ecological niche modeling and species distribution modeling. *Revista Mexicana de Biodiversidad*, *88*, 437–441.

Stockwell, D. R. B., & Noble, I. R. (1993). Induction of sets of rules from animal distribution data: A robust and informative method of data analysis. *Mathematical and Computer Simulation*, *33*, 385–390.

Tessarolo, G., Rangel, T. F., Araújo, M. B., & Hortal, J. (2014). Uncertainty associated with survey design in species distribution models. *Diversity and Distributions*, *20*, 1258–1269.

Thuiller, W. (2003). BIOMOD – Optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, *9*, 1353–1362.

Title, P. O., & Bemmels, J. B. (2018). ENVIREM: An expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*, *41*, 291–307.

Tognelli, M. F., Roig-Juñent, S. A., Marvaldi, A. E., Flores, G. E., & Lobo, J. M. (2009). An evaluation of methods for modelling distribution of Patagonian insects. *Revista Chilena de Historia Natural*, *82*, 347–360.

United Nations Environmental Program/Convention on Biological Diversity. (2016). *Like-minded mega-diverse countries Carta to achieve Aichi Biodiversity Target 11. UNEP/CBD/COP/13/INF/45*. Cancun, Mexico: Author.

Urbina-Cardona, J. N., & Flores-Villela, O. (2010). Ecological-niche modeling and prioritization of conservation-area networks for Mexican herpetofauna. *Conservation Biology*, *24*, 1031–1041.

Urbina-Cardona, J. N., & Loyola, R. D. (2008). Applying niche-based models to predict endangered-hylid potential distributions: Are neotropical protected areas effective enough? *Tropical Conservation Science*, *1*, 417–445.

U.S. Geological Survey. (2018). *Mineral, commodity summaries 2018*. Reston, VA: Author.

van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. doi:10.1007/s11192-009-0146-3

van Eck, N. J., & Waltman, L. (2013). *VOSviewer manual*. Retrieved from http://www.vosviewer.com/download/f-z2x2.pdf

Villarroya, A., Barros, A. C., & Kiesecker, J. (2014). Policy development for environmental licensing and biodiversity offsets in Latin America. *PLoS One*, *9*, e107144.

Vega, G. C., Pertierra, L. R., & Olalla-Tárraga, M. A. (2017). MERRAclim, a high-resolution global dataset of remotely sensed bioclimatic variables for ecological modelling. *Nature – Scientific Data*, 4, 170078.

Velásquez-Tibatá, J., Olaya-Rodríguez, M. H., López-Lozano, D., Gutiérrez, C., González, I., & Londoño-Murcia, M. C. (2019). BioModelos: A collaborative online system to map species distributions. *PLoS One*, 14, e0214522.

Villero, D., Pla, M., Camps, D., Ruiz-Olmo, J., & Brotons, L. (2017). Integrating species distribution modelling into decision making to inform conservation actions. *Biodiversity and Conservation*, 26, 251–271.

Warren, D. L., Glor, R. E., & Turelli, M. (2010). ENMTools: A toolbox for comparative studies of environmental niche models. *Ecography*, 33, 607–611.

Zwiener, V. P., Padial, A. A., Marques, M. C. M., Faleiro, F. V., Loyola, R., & Peterson, A. T. (2017). Planning for conservation and restoration under climate and land use change in the Brazilian Atlantic Forest. *Diversity and Distributions*, 23, 955–966.

Appendix 1. Search terms used to return documents on SDM in two academic databases.

Database	Web of Science (Core collection)	SCOPUS
Search equation	TS=(("specie* distribution" or (specie* NEAR/5 distribution) or "specie* distribution model*" or (specie* NEAR/3 model*) or "niche* model*" or (niche* NEAR/5 model*) or "ecologic* niche" or (ecologic* NEAR/5 niche*) or species distribution models or (SDM) or (ENM) or environmental niche modeling or (environmental NEAR/5 model) or habitat suitability modeling or (habitat NEAR/5 model*) or environmental favourability or climatic niche model* or bioclimatic niche model*) AND (maximum entropy method or "maximum entropy" or (Maxent) or Genetic Algorithm for Rule-Set Prediction or (GARP) or bioclimatic modelling or (Bioclim) or receiver operating characteristic or (ROC) or area under the curve or (AUC) or (DISMO) or (BIOMOD) or (DOMAIN) or (ENMEval) or (MigClim) or (SDMPLAY) or (ModEco) or (WALLACE) or (ZOON) or (GLM) or general additive modeling or generalized	TITLE-ABS-KEY(("specie* distribution" or (specie* w/5 distribution) or "specie* distribution model*" or (specie* W/3 model*) or "niche* model*" or (niche* W/5 model*) or "ecologic* niche" or (ecologic* w/5 niche*) or {species distribution models} or (SDM) or (ENM) or {environmental niche modeling} or (environmental W/5 model) or "habitat suitability modeling" or (habitat W/5 model*) or "environmental favourability" or "climatic niche model*" or "bioclimatic niche model*") AND ({maximum entropy method} or "maximum entropy" or (Maxent) or {Genetic Algorithm for Rule-Set Prediction} or (GARP) or {bioclimatic modelling} or (Bioclim) or {receiver operating characteristic} or (ROC) or {area under the curve} or (AUC) or (DISMO) or (BIOMOD) or (DOMAIN) or (ENMEval) or (MigClim) or (SDMPLAY) or (ModEco) or (WALLACE) or (ZOON) or (GLM) or {general additive modeling} or {generalized additive

	additive models or (GAM) or boosted regression trees or (BRT) or random forests or niche mapper or (DIVA) or ecological niche factor analysis or (ENFA))) AND PY=(1970-2018)	models} or (GAM) or {boosted regression trees} or (BRT) or {random forests} or {niche mapper} or (DIVA) or { ecological niche factor analysis} or (ENFA))) PUBYEAR < 2019 AND (LIMIT-TO (PUBSTAGE,"final")) AND (LIMIT-TO (DOCTYPE,"ar") OR LIMIT-TO (DOCTYPE,"re") OR LIMIT-TO (DOCTYPE,"ch") OR LIMIT-TO (DOCTYPE,"le") OR LIMIT-TO (DOCTYPE,"bk"))
Number of documents	11471	7877

Identification

19348 of records identified through database searching

Screening

13388 of records after duplicates removed

6279 of records (global database) after search term filtering and title reading

7109 of documents outside disciplines of interest. e.g. pollution, psychiatry, neuroscience, hydrology, edaphology, chemistry, acoustics, morphometric, enzymatic activity, spectroscopy, land use/cover changes, electrochemical noise measurements (ENM), shared decision making (SDM), site-directed mutagenesis (SDM), signal-domain mixing (SDM), sulfadimethoxine (SDM), space-division multiplexing (SDM) techniques

Eligibility

5533 of full-text articles assessed for eligibility (Title, abstract and keywords reading)

746 of documents excluded as they doesn't model species distributions (e.g. mid-domain effect along elevation gradients, occupancy models, species diversity-richness modelling, species distribution based upon long/latitude raw data, species´ habitat selection, animal-plant interactions, aboveground biomass, bioacoustics, co-occurrence and nestedness patterns, notes on species´ geographic range expansion).

Included

5533 studies included for global analyses and filtered for those including co-authors with institutional affiliation pertaining to Latin American countries

Information obtained:
1. Co-author institutional affiliation
2. Country of this institution

1000 studies from Latin American institutions included in qualitative synthesis

Information obtained:
1. Latin American country of this institution
2. Co-author institutional affiliation
3. Keyword thematic category (taxonomic group, realm of application, geographic region and ecosystem type)
4. Methods to model species distributions obtained from abstract section of 687 document (68,6% of Latin American database)
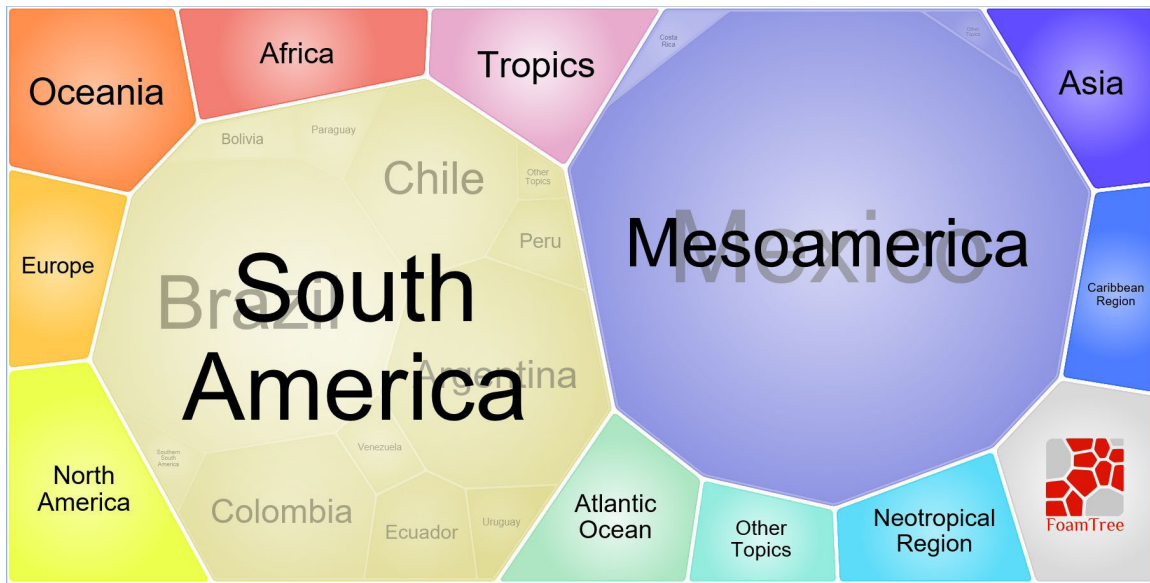
Appendix 2. Flow diagram for study selection along different phases of this review. Modified from Liberati et al. 2009.
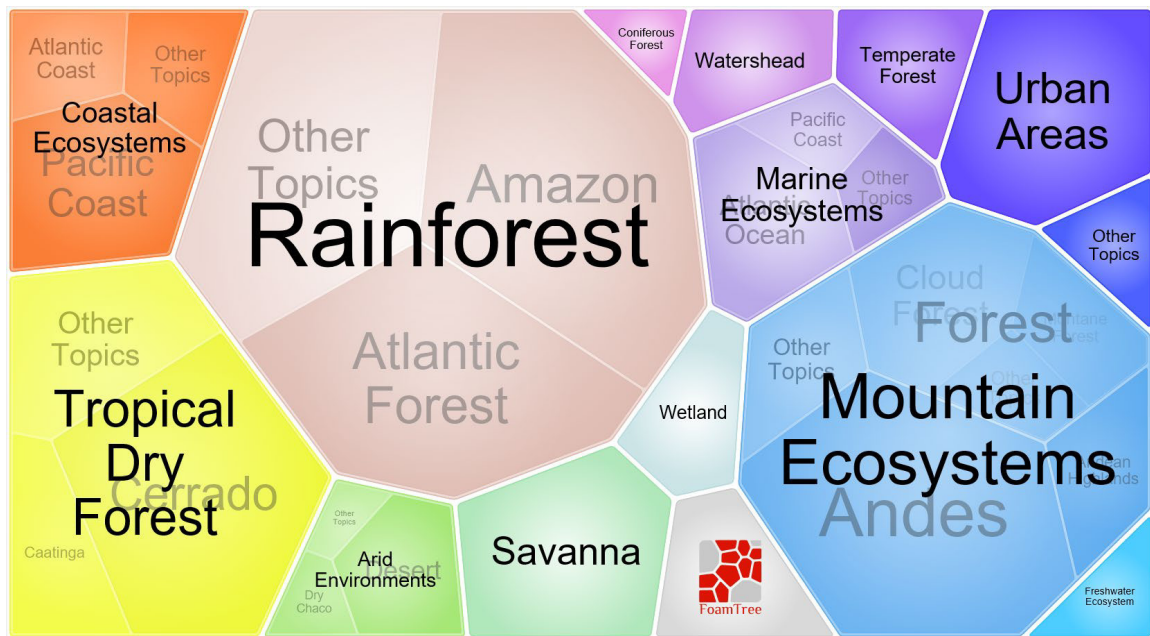
Appendix 3. Co-ocurrence matrix of 556 Latin American institutions that interacts with 591 other institutions outside Latin American countries

Appendix 4. Carrot search FoamTree visualization of geographic regions in which the SDM were conducted in documents in which there is a collaboration of Latin American institutions.

Appendix 5. Carrot search FoamTree visualization of ecosystem types in which the SDM were conducted in documents in which there is a collaboration of Latin American institutions.

Appendix 6. Number of papers published per each method used to model species distribution in 673 articles published with coautorship of some Latin American institution.

MODELING METHOD	1995-2000	2001-2005	2006-2010	2011-2015	2016-2018	TOTAL	Key reference
Profile and distance based methods							
BIOCLIM	0	2	7	15	7	31	Nix & Busby (1986)
Chebyshev Distance	0	0	0	0	1	1	Carpenter et al. (1993)
CLIMEX	0	0	0	0	1	1	Sutherst (2007)
DOMAIN	0	1	5	6	5	17	Carpenter et al. (1993)
ECOCROP	0	0	0	3	0	3	Ramírez-Villegas (2013)
Envelope Score	0	0	0	2	2	4	Piñeiro et al. (2007)
Euclidean Distance	0	0	1	3	1	5	Munguia et al. (2012)
Floramap	0	2	1	0	0	3	Jones et al. (2003)
Mahalanobis Distance	0	2	3	9	3	17	Farber & Kadmon (2003)
Multivariable kernel density estimation	0	0	0	0	1	1	Blonder et al. (2014)
SPEDInstabR	0	0	0	0	1	1	Guisande et al. (2017)
Surface Range Envelope	0	0	0	0	1	1	Jiguet et al. (2011)

Statistical methods							
Bayesian logistic regression	0	1	0	0	1	2	Gelman et al. (2013)
Bayesian networks	0	0	0	0	1	1	Mello et al. (2013)
Canonical quadratic ordination (CQO)	0	0	1	0	0	1	Yee (2004)
Climate Space Model	0	0	0	1	0	1	Robertson et al. (2001)
Ecological Niche Factor Analysis (ENFA)	0	0	0	2	8	10	Hirzel et al. (2002)
Factorial Discriminant Analysis (FDA)	0	0	0	1	0	1	Marmion et al. (2009)
Flexible Discriminant Analysis	0	0	0	0	1	1	Hastie et al. (1994)
Generalized Additive Models (GAM)	0	1	4	8	8	21	Hastie & Tibshirani (2017)
Generalized Linear Models (GLM)	0	0	7	19	17	43	Guisan et al. (2002)
Multivariate Adaptive Regression Splines (MARS)	0	0	0	2	2	4	Elith & Leathwick (2008)
Ordered weighted averaging	0	0	1	0	2	3	Jiang & Eastman (2000)
Weights of evidence	0	0	1	0	0	1	Raines et al. (2000)

Machine Learning Methods							
Artificial Neural Networks	0	0	3	1	3	7	Ripley (1996)
Automatically induced rule-based fuzzy models	0	0	1	0	0	1	Vinterbo et al. 2005
BioEnsembles	0	0	0	1	0	1	Rangel et al. (2009)
BIOMOD	0	0	2	1	2	5	Thuiller et al. (2009)
Boosted Regression Trees	0	0	2	4	10	16	Elith et al. (2008)
Classification and regression trees (CART)	0	0	2	1	2	5	Breiman et al (1984)
Environmental Distance	0	0	0	1	0	1	de Souza Muñoz (2011)
Genetic Algorithm for Rule -Set Prediction (GARP)	1	23	47	34	21	126	Stockwell & Peters (1999)
MaxEnt	0	0	38	206	216	460	Phillips et al. (2006)
Niche Mosaic	0	0	0	1	0	1	de Souza Muñoz (2011)
Random Forests	0	0	3	10	20	33	Prasad et al. (2006)
Support Vector Machines	0	0	2	3	8	13	Guo et al. (2005)

References from Appendix 6:

Blonder, B., Lamanna, C., Violle, C., & Enquist, B. J. (2014). The n-dimensional hypervolume. *Global Ecology and Biogeography*, 23, 595–609.

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and Regression Trees*. Chapman & Hall/CRC, Boca Raton, FL, USA

Carpenter, G., Gillson, A. N., & Winter, J. (1993). DOMAIN: A flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation,* 2, 667–680.

de Souza Muñoz, M. E., De Giovanni, R., de Siqueira, M. F., Sutton, T., Brewer, P., Pereira, R. S., ... & Canhos, V. P. (2011). openModeller: a generic approach to species' potential distribution modelling. *GeoInformatica*, 15, 111-135.

Elith, J., & Leathwick, J. R. (2007). Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. *Diversity and Distributions,* 13, 265–275.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77, 802-813.

Farber O., & Kadmon, R. (2003). Assessment of alternative approa-ches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling,* 160, 115–130.

Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. Chapman and Hall/CRC.

Guisan, A., Edwards, Jr. T. C., & Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecological Modelling,* 157, 89–100.

Guisande, C., García-Roselló, E., Heine, J., González-Dacosta, J., Vilas, L. G., Pérez, B. J. G., & Lobo, J. M. (2017). SPEDInstabR: an algorithm based on a fluctuation index for selecting predictors in species distribution modeling. *Ecological informatics*, 37, 18-23.

Guo, Q. H., Kelly, M., & Graham, C. H. (2005) Support vector machines for predicting distribution of sudden oak death in California. *Ecological Modelling,* 182, 75−90.

Hastie T. J., & Tibshirani, R. (1990). *Generalized Additive Models*. Chapman and Hall, London.

Hastie, T., Tibshirani, R., & Buja, A. (1994). Flexible discriminant analysis by optimal scoring. *Journal of the American Statistical Association,* 89, 1255–1270.

Hirzel, A. H., Hausser, J., Chessel, D., & Perrin, N. (2002). Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data?. *Ecology*, 83, 2027-2036.

Jiang, H. & Eastman, J. R. (2000). Application of fuzzy measures in multi-criteria evaluation in GIS. *International Journal of Geographical Information Science,* 14, 173-184.

Jiguet, F., Barbet-Massin, M., & Chevallier, D. (2011). Predictive distribution models applied to satellite tracks: modelling the western African winter range of European migrant Black Storks Ciconia nigra. *Journal of Ornithology, 152, 111-118.*

Jones, P. G., Gladkow, A., & Jones, A. L. (2003). *FloraMap: A computer tool for predicting the distribution of plants and other organisms in the wild, version 1.02*. Centro Internacional de Agricultura Tropical (CIAT).

Marmion, M., Luoto, M., Heikkinen, R. K., & Thuiller, W. (2009) The performance of state-of-the-art modelling techniques depends on geographical distribution of species. *Ecological Modelling,* 220, 3512−3520.

Mello, M.P., Risso, J., Atzberger, C., Aplin, P., Pebesma, E., Vieira, C.A.O., Rudorff, B.F.T. (2013). Bayesian networks for raster data (BayNeRD): Plausible reasoning from observations. *Remote Sensing,* 5, 5999–6025.

Munguía, M., Rahbek, C., Rangel, T. F., Diniz-Filho, J. A. F., & Araújo, M. B. (2012). Equilibrium of global amphibian species distributions with climate. *PLoS One*, 7(4), e34420.

Nix, H. A., & Busby, J. (1986). *BIOCLIM, a bioclimatic analysis and prediction system*. Division of Water and Land Resources: Canberra.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* 190, 231–259.

Piñeiro, R., Aguilar, J. F., Munt, D. D. & Feliner, G. N. (2007) Ecology matters: Atlantic-Mediterranean disjunction in the sand-dune shrub Armeria pungens (Plumbaginaceae). *Molecular Ecology,* 16, 2155-2171.

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006) Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems,* 9, 181−199.

Raines, G. L., Bonham-Carter, G. F., & Kemp, L. (2000). *Predictive probabilistic modeling using ArcView GIS*. ArcUser April-June: 45-48

Ramirez-Villegas, J., Jarvis, A., & Läderach, P. (2013). Empirical approaches for assessing impacts of climate change on agriculture: The EcoCrop model and a case study with grain sorghum. *Agricultural and Forest Meteorology*, 170, 67-78.

Rangel, T. F. L. V. B., Diniz-Filho, J. A. F., & Araújo, M. B. (2009). *BIOENSEMBLES 1.0. Software for Computer Intensive Ensemble Forecasting of Species Distributions Under Climate Change*. Goias, Madrid: Evora.

Ripley, B.D. (1996). *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge, UK.

Robertson, M.P., Caithness, N., Villet, M. H. (2001) A PCA-based modelling technique for predicting environmental suitability for organisms from presence records. *Diversity and Distributions*, 7, 15-27.

Stockwell, D.R.B., & Peters, D. P. (1999). The GARP modelling system: Problems and solutions to automated spatial prediction. *International Journal of Geographical Information Systems,* 13, 143–158.

Sutherst, R. W., maywald, G. F., & kriticos, D. J. (2007). *CLIMEX Version 3: User's Guide*. Melbourne, Australia: Hearne Scientific Software.

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography*, 32, 369-373.

Vinterbo, S.A., Kim, E.-Y., Ohno-Machado, L. (2005). Small, fuzzy and interpretable gene expression based classifiers. *Bioinformatics,* 21, 1964–1970.

Yee, T. W. (2004). A new technique for maximum-likelihood canonical Gaussian ordination. *Ecological Monographs,* 74, 685-701.

Appendix 7. Matrix of 1000 papers on SDM in which Latin American institutions had collaborated. From this database, the modeling methods used on each document to model species distributions, were obtained by reading the abstract section, allowing us to categorize 687 papers (673 papers with modeling methods identified are represented in yellow cells in Colum E "Modeling method"; and 14 were review papers that didn´t conduct any SDM modeling, identified in blue cells in the same column).

SEE EXCEL FILE INCLUDED SEPARATELY

Worksheet can be downloaded from this web link:

https://docs.google.com/spreadsheets/d/1syrKp8PMV_ckT3wLyXuGdFb2XSmH22EWztix5Yn8fLM/edit?usp=sharing