

Biogeography

Ecological niche models as a tool for estimating the distribution of plant communities

Modelos de nicho ecológico como herramienta para estimar la distribución de comunidades vegetales

Mayra Flores-Tolentino, Enrique Ortiz, José Luis Villaseñor*

Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, Apartado postal 70-233, 04510 México City, Mexico

*Corresponding author: vrios@ib.unam.mx (J.L. Villaseñor)

Received: 2 October 2018; accepted: 26 April 2019

Abstract

Ecological niche models allow inferences about the distribution of species and communities, facilitating the identification of priority areas for conservation, mainly in threatened environments such as the seasonally dry tropical forest (SDTF). The objectives of this work were to delimit SDTF using mostly restricted and endemic vascular plant species, in addition to evaluating and comparing the performance of the model with 4 additional proposals of the distribution of SDTF in Michoacán. To delimit the distribution of SDTF in Michoacán, 76 individual ecological niche models were constructed. Then, the individual models were assembled to obtain the biome distribution model, and the resulting map was compared with the other SDTF proposals for Michoacán. The model best supported by the observed data and balanced in the percentage of omission and commission errors was our model, and the model most like ours in terms of the predicted area, was the one proposed by INEGI (2003). The use of widely distributed species in the definition of communities results in models with greater overestimation. It is important to adapt the available information and knowledge about the object of study, to properly integrate them into the different algorithms that allow us to obtain an approximation of what happens with the species or communities.

Keywords: Seasonally dry tropical forest; Endemic species; Maxent; Michoacán; Mexico

Resumen

Los modelos de nicho ecológico permiten hacer inferencias sobre la distribución de especies y comunidades, permitiendo identificar áreas prioritarias para la conservación, principalmente en ambientes amenazados como el bosque tropical estacionalmente seco (SDTF, por sus siglas en inglés). Los objetivos de este trabajo fueron delimitar el SDTF utilizando especies de plantas vasculares endémicas y características del bioma, además de evaluar y comparar el modelo con 4 propuestas sobre la distribución del SDTF en Michoacán. Para delimitar su distribución, se realizaron 76 modelos de nicho ecológico. Los modelos individuales fueron ensamblados para obtener el modelo de distribución del bioma y el mapa resultante fue comparado con las propuestas del SDTF para Michoacán. El mejor modelo respaldado

por los datos observados y equilibrado en el porcentaje de errores de omisión y comisión fue nuestro modelo y el más similar a éste fue el propuesto por el INEGI (2003). El uso de especies de amplia distribución en la definición de comunidades resulta en modelos con mayor sobreestimación. Es importante adecuar la información disponible y el conocimiento sobre el objeto de estudio, para integrar los adecuadamente a los diferentes algoritmos que permitan obtener un acercamiento sobre lo que ocurre con las especies o comunidades.

Palabras clave: Bosque tropical estacionalmente seco; Especies endémicas; Maxent; Michoacán, México

Introduction

Environmental deterioration demands better documentation of the planet's biodiversity. This has required the development of tools that facilitate data management and decision-making, such as geographic information systems (Ferrier, 2002; Guisan & Theurillat, 2000; Guisan & Thuiller, 2005), which allow inferences about the distribution of a species or a group of species to be made more quickly and easily (Elith et al., 2006; Peterson & Soberón, 2012). Among the multiple uses of geographic information systems (GIS), there are studies in which they use these tools to model plant communities (Franklin, 1995). Predicting the distribution of a plant community using GIS has been important in studies of conservation of priority areas, of ecological restoration and to evaluate the effect of climate change (Bojórquez-Tapia et al., 1995; Franklin, 1995; Zimmermann, 2000).

The geographical limits of a plant community are not always easy to define (Méndez-Toribio et al., 2014). In studies that model communities, several strategies have been proposed for identifying the spatial patterns that define them (Elith et al., 2006; Ferrier, 2002; Ferrier & Guisan, 2006; Franklin, 1995). One of the most used methodologies in the modeling of the distribution of plant communities is using species assembly, which consists of the combination of individual potential distribution models to generate a map of the distribution of a community. This combination provides information on the distribution of species richness in certain areas within the study community. In addition, it allows the identification of spatial patterns followed by the species that make up this community (Ferrier & Guisan, 2006; Franklin, 1995; Guisan & Zimmermann, 2000).

The different methodologies and algorithms used to model the potential distribution of species yield different results (Araújo & Guisan, 2006; Elith et al., 2006; Pearson et al., 2006). For this reason, alternative statistical tests have been implemented to evaluate the performance of the models, such as the binomial test, the partial ROC (receiver operating characteristics), the AUC (area under the curve), etc. (Fielding & Bell, 1997; Hirzel et al., 2006; Liu et al., 2016; Oreskes et al., 1994; Peterson et al., 2008). These statistics use a set of data independent from the data used

to generate the model to evaluate the performance and constancy of a model. Model evaluation allows researchers to discern which model is the most appropriate and best supported by the observed data (Johnson & Omland, 2004), guaranteeing the correct interpretation and quality of the results (Guisan & Thuiller, 2005; Mateo et al., 2011; Muñoz & Felicísimo, 2004). Despite the importance of these tests, such validation with independent data (some set of data not used as training or test) is rarely used in species modeling studies.

The seasonally dry tropical forest (SDTF), also known as tropical dry forest, is considered one of the most diverse and distinctive biomes worldwide due to the large number of endemic species it harbors (Olson et al., 2000). In Mexico, SDTF contains more than 6,000 species of vascular plants, of which more than 40% are endemic (Rzedowski, 1978; Villaseñor & Ortiz, 2014). National states with the highest number of endemic or characteristic (mostly restricted to) SDTF species are Oaxaca (1,396), Guerrero (1,251), Jalisco (1,237), and Michoacán (1,053; Villaseñor, 2016). In Michoacán, SDTF is located mainly in the Balsas Depression, a physiographic and floristic province that also contains the largest area of SDTF in the country (Rzedowski, 1978). The SDTF in the state of Michoacán contains 18% of the total vascular flora reported for the state (Villaseñor, 2016).

For several decades, the SDTF has faced serious conservation problems due to urban expansion and the constant exploitation of plant and animal species. As such, its original distribution area has been drastically reduced, causing the loss of and constant threat to a high number of species, mainly endemic. Due to this problem, it is considered among the most threatened tropical biomes in the world (DRYFLOR et al., 2016; Challenger & Soberón, 2008; Trejo & Dirzo, 2000).

This study has the purpose of evaluating different proposals of the potential distribution of the SDTF in Mexico, specifically in the state of Michoacán. In Mexico, SDTF includes low deciduous forest, deciduous tropical forest, dry forest, and thorny forest (Miranda & Hernández, 1963; Rzedowski, 1978; Villaseñor & Ortiz, 2014) and is distributed mainly in the Pacific slope from southern Sonora and southwestern Chihuahua to Chiapas and on the

gulf slope from Tamaulipas to the Yucatán Peninsula. It is found at altitudes ranging from 0 m to 1,900 m (Villaseñor & Ortiz, 2014). An important determining characteristic of SDF distribution is temperature, particularly the extreme minimum temperature (Challenger & Soberón, 2008).

Michoacán, with 1,053 species of native plants mostly distributed in the SDF, ranks fourth in richness at the state level. In addition, several floristic studies have documented a large part of the flora and its biodiversity in general (e.g., Cué-Bár, Villaseñor, Arredondo et al., 2006; Cornejo-Tenorio & Ibarra-Manríquez, 2017; Leavenworth, 1946; Rodríguez & Espinosa, 1995, 1996), information that allows inferences to be made about the distribution of species among different biomes. Given the number of species present in the SDF and the information available (databases with enough information to develop ecological niche models), it is important to evaluate the hypothesis that flowering plants (Angiosperms) that are endemic to Mexico, whose distribution is restricted to the SDF, are good predictors of its geographical distribution.

The delimitation of the SDF in Mexico has been addressed several times using GIS that consider different inputs, including the combination of plant and animal species (Prieto-Torres & Rojas-Soto, 2016), empirical knowledge (Rzedowski, 1990), vegetation censuses (INEGI, 2003) or particular plant species (Villaseñor & Ortiz, 2014). However, to date there has been no evaluation of these proposals to define which model is more statistically accurate in delimiting the biome.

The objective of this work is to propose a delimitation of the SDF in the state of Michoacán based on the assembly of ecological niche models of species that are endemic to or whose distribution is restricted to the biome. Once this potential distribution of the SDF has been defined, it is compared with other published proposals that delimit the biome, such as Rzedowski (1990), INEGI (2003), Villaseñor and Ortiz (2014) and Prieto-Torres and Soto-Rojas (2016).

Materials and methods

The endemic or characteristic (mostly restricted to) species of the seasonally dry tropical forest (Villaseñor, 2016, unpublished data) in Michoacán were selected based on the number of records from 2 sources of information: the National Biotic Information System (SNIB-REMIB) at the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (Conabio, 2018), and the digital repository of the National Herbarium of Mexico (MEXU-UNIBIO), at the Instituto de Biología, Universidad Nacional Autónoma de México (UNAM). The selected records were stored in a database which was later refined following the

recommendations of Castillo et al. (2014) and Chapman (2005). Namely, 1) duplicate records were eliminated, 2) records that did not have coordinates were georeferenced, 3) records that could not be georeferenced were deleted, and 4) species with < 5 records were excluded (Pearson et al., 2007). The final database consisted of 76 species. Finally, for each species, an exploration of the data was carried out by generating their biogeographic track in ArcMap using the tool “minimum spanning tree tools”/EMST in order to detect sampling points of a species that deviate from the geographical pattern followed by most of the records. Any disjoint points detected were carefully evaluated, both taxonomically and geographically, for inclusion or elimination in the analyses (Fig. 1).

We considered 58 environmental variables, with a resolution of 30 seconds (~1 km²), obtained from different sources. Twenty-six variables are climatic (Hijmans et al., 2005; Cruz-Cárdenas, López-Mata, Villaseñor et al., 2014), 9 topographic, 9 edaphological and 14 include remote sensing data (Cruz-Cárdenas, López-Mata, Ortiz-Solorio et al., 2014). The 58 layers were cut using the polygon of the state of Michoacán as a mask; the clipping of the variables was carried out in ArcMap 10.0 (ESRI, 2010). Restriction of the variables to the limits of the state of Michoacán was done in order to represent more precisely the accessible area (M in the BAM diagram) that a certain species can occupy (Soberón & Peterson, 2005; Soberón et al., 2017), thus improving model calibration (Barve et al., 2011).

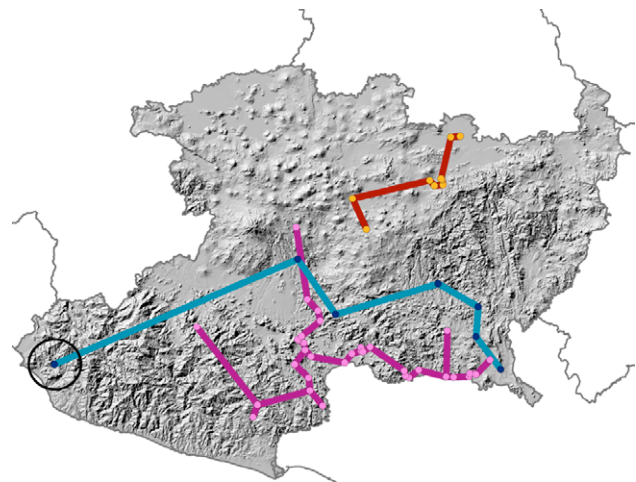


Figure 1. Biogeographic tracks of 3 species endemic or characteristic of the seasonally dry tropical forest. Notice that species with a blue line shows a site (black circle) which departs significantly from the pattern followed by the track in the study area (Michoacán).

The use of too many environmental predictors can cause overfitting of the models, leading to errors in the prediction of species distributions (Peterson & Nakazawa, 2008). To reduce the number of variables, a Principal Component Analysis (PCA) was carried out (Dormann et al., 2007). For this analysis, the values of the 58 variables were extracted from each of the presence records; these values were used to run a PCA in R program (R Development Core Team, 2017). The components that accumulated 80% of the variance, and within each component, the variables with above 95% of the load values were used for further analyses. The PCA made it possible to detect the components with the greatest contribution to the models and eliminate collinearity among the variables (Cruz-Cárdenas, López-Mata, Villaseñor et al., 2014).

An ecological niche model was constructed individually for each of the selected endemic/characteristic SDTF species. Niche models were generated using Maxent's algorithm ver. 3.3.3e (Phillips et al., 2006), which uses only presence data to model the most uniform distribution of a given species throughout the study area, restricted by the environmental specifications of the presence data. The extrapolation of these values to a given area results in a probability distribution map ranging from 0 to 1 (Phillips & Dudík, 2008). Maxent has proven to be a reliable algorithm and generates good results even with small sample sizes (< 10) (Pearson et al., 2007; Phillips et al., 2006).

The inputs used to run the ecological niche models were the records of the presence of each species and the environmental variables with the greatest contribution in the PCA. The program was used with the default configuration except for the do clamping and extrapolate option, to avoid extrapolation of the extreme values of the variables (Elith et al., 2011). It was also configured to use 25% of the data to validate the model and 75% to run the analysis (Phillips et al., 2006). Individual models were evaluated using independent threshold statistics, such as the Area under the curve (AUC), although it has been strongly criticized for the similar weight attributed to omission and commission errors (Lobo et al., 2008; Peterson et al., 2008). Therefore, in addition the models were also evaluated using the partial ROC (Peterson et al., 2008) which was calculated with the tool for partial ROC V. 1.0 program (Barve, 2008), applying 1,000 iterations and considering 5% error. The significance of the AUC ratio was estimated with bootstrapping (1,000 replications) using 50% of training localities; such AUC ratios were calculated with the Z test of the R program (<http://www.r-project.org/>).

The Maxent logistic outputs from each of the 76 species-level ecological niche model were processed in

ArcMap ver.10.0, to obtain the potential distribution of the SDTF. The probabilistic models obtained from Maxent were converted to binary layers (0-1) using the maximum training sensitivity plus specificity threshold, which has been shown to be the most accurate in predicting the distribution of species (Jiménez-Valverde & Lobo, 2007; Liu et al., 2013; Weber, 2011). The resulting binary models were summed using the Algebra map tool implemented in ArcMap (D'Amen et al., 2015, 2017; Henderson et al., 2014). The layer with the sum of the models was reclassified using 76 thresholds, each representing the superposition from 1 to 76 of the individual models. At each threshold errors of omission were calculated considering all the records of the modeled species plus an additional number of data documenting the presence of the SDTF in Michoacán. An error of omission was considered when a point of presence of the species was not predicted by the threshold (Fielding & Bell, 1997). The threshold with the lowest error rate of omission was considered as the potential distribution of the SDTF in Michoacán.

The comparison and evaluation of the different models of the distribution of the SDTF was carried out using several statistical tests. These tests also allowed to evaluate the over- and under-estimation of the proposals of the SDTF distribution. The following proposals were compared to the model generated in this study: Rzedowski (1990), INEGI (2003), Villaseñor and Ortiz (2014), and Prieto-Torres and Soto-Rojas (2016). The first test was a comparison of the geographic overlap between our model and each of the previous models, calculated by adding each pair of models using the "Algebra map tool package" of ArcMap. In a second test, the similarity of the models was evaluated using the Kappa estimator. For this analysis, our model was considered the closest to the potential distribution of the SDTF (reference model). The analysis was performed with the Kappa Analysis Tools extension in ArcView 3.x (Jenness & Wynne, 2005, 2007; Pontius & Millones, 2011). This estimator measures the overall similarity of the model, considering a rank from 0-1, in which values < 0.40 are considered bad models, < 0.60 regular models, > 0.80 good and 1 excellent (Landis & Koch, 1977).

A third evaluation test consisted of quantifying errors of omission, using all the presence records of the selected species. The presence data points were superimposed on each of the 4 models to identify where the models failed to predict them. A fourth evaluation analysis was a binomial test using 183 additional records of species of vascular plants characteristic of the SDTF that were independent of the points used to generate the model. This test gives an estimate of the amount of omission errors of each proposal. Finally, 2,320 occurrence records of species (Conabio, 2018)

reported in vegetation types adjacent to the SDTF were used to obtain the model with the greatest uncertainty in the classification of the biome. These vegetation types were tropical sub-deciduous forest (373 records), *Quercus* forest (1,806), humid mountain forest (118) and xerophytic scrub (23). This final test allowed the detection of over- and under-estimation by the models and the rate of commission errors.

Results

We documented 253 species recorded as endemic to or characteristic of (mostly restricted to) the SDTF in Mexico and occurring in Michoacán; in total they included 2,781 records. Of these species, 76 species had more than 5 records in the state, totaling 1,051 records that were used for modeling of the distribution of SDTF in Michoacán (Table 1).

The principal components analysis applied to eliminate the collinearity of the variables indicated that the first 8 components explain 82% of the environmental variation, which included 30 of the 58 variables initially

proposed (Table 2). Those 30 variables were used to run the ecological niche models for each species. A detailed explanation of the selected variables is provided by Cruz-Cárdenas, López-Mata, Villaseñor et al. (2014) and López-Mata et al. (2012).

An individual ecological niche model was generated for each of the 76 focal SDTF species. Most models had high AUC values (92%: 0.900-0.999, considered very good models). The remaining 8% of the models' AUC values were considered good (0.835-0.0.887). The ecological niche models obtained were better than a randomly generated model (Baldwin, 2009; Peterson et al., 2011). Performance estimation of the models was also favorable using the partial ROC. Likewise, 100% of models were statistically significant, with AUC ratios ranging from 1.052 to 1.987 ($p < 0.001$; Table 1). The assembly of the binary models was considered to represent the potential distribution of the SDTF in Michoacán; the threshold limit of the SDTF was considered with the pixels where 5 or more species coincided, which was the cut-off threshold that produced the lowest percentage of errors of omission (17%).

Table 1

Species used for the delimitation of the seasonally dry tropical forest in the state of Michoacán, Mexico (N = 76). For each species, the number of records used for training and test is indicated; likewise, the performance value of the model estimated through the AUC and Partial ROC, and the omission of the binary model (expressed as the number of records omitted) are indicated.

Species	Training/ testing	AUC	Partial ROC	Omission
Acanthaceae				
<i>Aphelandra lineariloba</i> Leonard	6/1	0.938	1.333	1
<i>Dicliptera haenkeana</i> Nees	7/2	0.942	1.285	1
<i>Tetramerium langlassei</i> Happ	5/1	0.835	1.193	0
<i>Tetramerium rubrum</i> Happ	6/1	0.913	1.152	1
Apocynaceae				
<i>Fernaldia asperoglottis</i> Woodson	11/3	0.912	1.132	1
<i>Marsdenia callosa</i> Juárez-Jaimes & W.D. Stevens	5/1	0.977	1.509	0
<i>Prestonia contorta</i> (M. Martens & Galeotti) Hemsl.	5/1	0.901	1.127	0
<i>Thevetia pinifolia</i> (Standl. & Steyerf.) J.K. Williams	5/1	0.982	1.148	1
Aristolochiaceae				
<i>Aristolochia mutabilis</i> Pfeifer	4/1	0.956	1.733	0
Asteraceae				
<i>Bidens mexicana</i> Sherff	4/1	0.963	1.385	1
<i>Cosmos pacificus</i> Melchert var. <i>pacificus</i>	6/1	0.950	1.297	1
<i>Cymophora accedens</i> (S.F. Blake) B.L. Turner & A.M. Powell	4/1	0.998	1.637	1
<i>Dendroviguiera puruana</i> (Paray) E.E. Schill. & Panero	8/2	0.959	1.052	1
<i>Guardiola pappifera</i> Paul G. Wilson	3/1	0.993	1.518	1
<i>Melampodium dicoelocarpum</i> B.L. Rob.	9/3	0.917	1.366	1
<i>Melampodium nutans</i> Stuessy	9/2	0.962	1.153	1

Table 1
Continued

Species	Training/ testing	AUC	Partial ROC	Omission
<i>Pectis decemcarinata</i> McVaugh	12/3	0.980	1.181	0
<i>Pectis linifolia</i> L. var. <i>hirtella</i> S.F. Blake	9/3	0.887	1.335	1
Boraginaceae				
<i>Cordia globulifera</i> I.M. Johnst.	9/2	0.905	1.271	0
Burseraceae				
<i>Bursera confusa</i> (Rose) Engl.	8/2	0.849	1.156	2
<i>Bursera coyucensis</i> Bullock	27/9	0.986	1.121	2
<i>Bursera crenata</i> Paul G. Wilson	38/12	0.970	1.052	4
<i>Bursera discolor</i> Rzed.	19/6	0.958	1.088	1
<i>Bursera excelsa</i> (Kunth) Engl. var. <i>acutidens</i> (Sprague & L. Riley) McVaugh & Rzed.	3/1	0.998	1.987	0
<i>Bursera fagaroides</i> (Kunth) Engl. var. <i>purpusii</i> Brandegeee) McVaugh & Rzed.	9/2	0.901	1.183	0
<i>Bursera infernalis</i> Guevara & Rzed.	21/7	0.988	1.079	3
<i>Bursera kerberi</i> Engl.	23/7	0.971	1.066	2
<i>Bursera occulta</i> McVaugh & Rzed.	3/1	0.996	1.907	0
<i>Bursera paradoxa</i> Guevara & Rzed.	18/6	0.982	1.138	2
<i>Bursera sarcopoda</i> Paul G. Wilson	10/3	0.905	1.401	1
<i>Bursera sarukhanii</i> Guevara & Rzed.	26/8	0.976	1.052	1
<i>Bursera toledoana</i> Rzed. & Calderón	6/2	0.984	1.332	0
<i>Bursera trifoliolata</i> Bullock	12/3	0.970	1.157	2
<i>Bursera trimera</i> Bullock	24/8	0.973	1.094	2
Cactaceae				
<i>Backebergia militaris</i> (Audot) Bravo ex Sánchez-Mej.	10/3	0.955	1.058	2
<i>Pachycereus tepamo</i> Gama & S. Arias	6/1	0.999	1.619	0
<i>Stenocereus chrysocarpus</i> Sánchez-Mej.	7/2	0.956	1.328	1
Celastraceae				
<i>Crossopetalum managuatillo</i> (Loes.) Lundell	13/4	0.932	1.060	3
Convolvulaceae				
<i>Calycobolus nutans</i> (Moc. & Sessé ex Choisy) D.F. Austin	6/1	0.900	1.525	0
<i>Ipomoea puniculata</i> Benth.	7/2	0.977	1.203	2
Cucurbitaceae				
<i>Cucurbita argyrosperma</i> K. Koch var. <i>argyrosperma</i>	6/1	0.947	1.519	0
<i>Rytidostylis longisepala</i> (Cogn.) C. Jeffrey	8/2	0.878	1.087	2
Cyperaceae				
<i>Carex arsenei</i> Kük.	5/1	0.991	1.288	0
Euphorbiaceae				
<i>Euphorbia linguiformis</i> McVaugh	5/1	0.987	1.294	0
<i>Euphorbia umbellulata</i> Engelm. ex Boiss.	7/2	0.953	1.348	0
<i>Jatropha galvanii</i> J. Jiménez Ram. & J.L. Contr.	4/1	0.982	1.899	0
<i>Jatropha stephanii</i> J. Jiménez Ram. & Mart. Gord.	11/3	0.994	1.161	3

Table 1
Continued

Species	Training/ testing	AUC	Partial ROC	Omission
<i>Manihot tomatophylla</i> Standl.	14/4	0.928	1.121	2
Fabaceae				
<i>Acaciella igualensis</i> Britton & Rose	4/1	0.997	1.614	1
<i>Aeschynomene hintonii</i> Sandwith	4/1	0.990	1.351	1
<i>Aeschynomene paucifoliolata</i> Micheli	5/1	0.925	1.252	0
<i>Desmanthus interior</i> (Britton & Rose) Bullock	8/2	0.973	1.193	1
<i>Lonchocarpus balsensis</i> M. Sousa & J.C. Soto	13/4	0.946	1.104	0
<i>Lonchocarpus obovatus</i> Benth.	4/1	0.951	1.479	0
<i>Lonchocarpus schubertiae</i> M. Sousa	7/2	0.975	1.341	0
<i>Mimosa egregia</i> Sandwith	7/2	0.974	1.596	0
<i>Mimosa rhododactyla</i> B.L. Rob.	5/1	0.981	1.486	1
<i>Mimosa rosei</i> B.L. Rob.	16/5	0.969	1.127	1
<i>Mimosa tricephala</i> Schltldl. & Cham. var. <i>nelsonii</i> (B.L. Rob.) Chehaibar & R. Grether	5/1	0.878	1.204	1
Lythraceae				
<i>Cuphea lobophora</i> Koehne var. <i>elongate</i> S.A. Graham	7/2	0.946	1.343	0
Malpighiaceae				
<i>Galphimia multicaulis</i> A. Juss.	9/3	0.928	1.482	0
<i>Galphimia paniculata</i> Bartl.	5/1	0.902	1.440	0
Malvaceae				
<i>Gossypium lobatum</i> Gentry	18/5	0.985	1.135	0
<i>Gossypium schwendimanii</i> Fryxell & S.D. Koch	6/2	0.999	1.336	2
<i>Gossypium trilobum</i> (Sessé & Moc. ex DC.) Skovst.	5/1	0.988	1.950	0
<i>Pavonia oxyphylla</i> (DC.) Fryxell var. <i>melanommata</i> (B.L. Rob. & Seaton) Fryxell	6/1	0.876	1.478	0
<i>Sida fastuosa</i> Fryxell & S.D. Koch	6/1	0.952	1.543	0
<i>Waltheria pringlei</i> Rose & Standl.	15/5	0.933	1.077	1
Nyctaginaceae				
<i>Salpianthus aequalis</i> Standl.	6/2	0.900	1.264	1
Passifloraceae				
<i>Passiflora juliana</i> J.M. MacDougal	3/1	0.990	1.869	1
<i>Passiflora viridiflora</i> Cav.	8/2	0.990	1.067	2
Primulaceae				
<i>Bonellia pringlei</i> (Bartlett) B. Ståhl & Källersjö	4/1	0.921	1.295	1
Ranunculaceae				
<i>Delphinium subscandens</i> Ewan	8/2	0.926	1.170	1
Rhamnaceae				
<i>Karwinskia johnstonii</i> Ric. Fernández	11/3	0.977	1.558	0
Rubiaceae				
<i>Simira mexicana</i> (Bullock) Steyerem.	9/2	0.971	1.215	1
Santalaceae				
<i>Phoradendron dolichocarpum</i> Kuijt	6/1	0.990	1.595	1

Table 2

Variables used in the modeling of 76 species characteristics of the seasonally tropical dry forest in the state of Michoacán, Mexico.

Type	Variable	
Climatic	bio02 (average daytime variation)	
	bio03 (isothermality)	
	bio04 (seasonality of temperature)	
	bio05 (maximum temperature of the warmest month)	
	bio06 (minimum temperature of the coldest month)	
	bio07 (annual variation in temperature)	
	bio11 (average of the coldest quarter temperature)	
	bio12 (annual rainfall)	
	bio13 (precipitation of the wettest month)	
	bio15 (seasonality of precipitation)	
	bio18 (precipitation of the warmest quarter)	
	bio19 (precipitation of the coldest quarter)	
	evahumed (real evapotranspiration of wet months)	
	evasecos (real evapotranspiration of dry months)	
	ppsecos (precipitation of dry months)	
	Topographic	aspect (orientation 0° to 90°)
mexdem (digital model of elevation)		
mexslope (slope)		
tri (terrain roughness index)		
twi (topographic moisture index)		
Edaphic Cruz-Cárdenas, López-Mata, Ortiz-Solorio et al. (2014)		mexca (calcium)
	mexce (electric conductivity)	
	mexco (organic carbon)	
	mexk (potassium)	
	mexmg (magnesium)	
	mexras (sodium absorption radius)	
	*MODIS	modisdic (normalized vegetation index December)
		modisfeb (normalized vegetation index February)
		modismar (normalized index of vegetation March)
		modisabr (normalized vegetation index April)

* Variables obtained with remote perception data (MODIS web): Moderate Resolution Imaging Spectroradiometer; December, February, March, and April 2009.

The model of STDF distribution in the state of Michoacán generated here estimated a surface area of 22,483.3 km², representing 38.7% of the state's territory (Fig. 2). According to the model, SDTF is distributed mainly in the regions of the Balsas Depression and part of the Mexican Pacific Coast (Lázaro Cárdenas and Aguililla municipalities). Figure 2 shows the potential distribution of the SDTF in Michoacán and the number of species considered for its delimitation.

The superposition (assembly) of the 76 models generated a detailed map of species richness along the distribution area of the SDTF (Fig. 3). The portions that concentrate more than 50% (29-43 species) of the species analyzed are found mostly in the southern and southeastern parts of the distribution, with an archipelago of small patches running to the northeast, near the state of Jalisco. The richest zones identified are included in 3 physiographic subprovinces (INEGI, 2014): the southern part of the Balsas Depression

(the westernmost richest area), the Infiernillo Dam in Tierra Caliente (Southern Cordillera subprovince) near the border with the state of Guerrero, and the Tepalcatepec Depression at the western edge of the SDTF distribution, near the border with the state of Jalisco (Fig. 3). From these nuclei of high concentration of species, a gradient is formed that decreases towards the limits of the SDTF with other biomes, such as the temperate forest.

Figure 3 shows the results of the comparison using the spatial coincidence, of the model proposed here (22,483.3 km²) with those proposed by Rzedowski (1990; 33,238 km²), INEGI (2003; 22,049 km²), Villaseñor and Ortiz (2014; 28,010 km²), and Prieto-Torres and Soto-Rojas (2016; 37,717 km²). The model with the largest coinciding area was that proposed by Prieto-Torres and Rojas-Soto (95%), and the model with the smallest coincidence was that of Villaseñor and Ortiz (78.4%). The models proposed by Prieto-Torres and Rojas-Soto (2016) and Rzedowski (1990) predicted larger areas, including many areas not included in our model (non-coincidence area), with 71.7% and 64.4% of overpredicted surface respectively (Fig. 4). Model evaluation metrics showed that the INEGI's (2003) proposal is the most similar to ours, with a Kappa value of 0.54, considered a model fairly similar (Landis & Koch, 1977). At the other extreme, the model with the lowest similarity with respect to ours was that of Prieto-Torres and Soto-Rojas (2016), with a Kappa value of 0.34 (Table 3); in this case Kappa values refer to the similarity with respect to our model, not to the actual distribution of STDF.

The model comparison based on omission errors using the 1,051 records of the 76 modeled species, indicated that proposal of Prieto-Torres and Soto-Rojas (2016) had the lowest omission rate (9.1%, 96 out of 1,051 records), followed by ours (16.8%, 177 records). The proposal with the highest rate of omission errors was that of Villaseñor and Ortiz (2014), with 19.7% (207 records, Table 3).

The binomial test carried out with 183 records that were not used to generate the models revealed that the 5 proposals resulted in probability values greater than 0.5 and were therefore better than a randomly generated model. The Prieto-Torres and Soto-Rojas (2016) model had the highest prediction probability (0.81, p -value < 2.20×10^{-16}) and the model with the least predictive success was the one proposed by Villaseñor and Ortiz (2014) (0.55, p -value < 0.001). Our model had a prediction probability of 0.74 (p -value < 2.20×10^{-16}). All models had high probabilities of classifying occurrences better than a randomly generated model (Table 3).

Table 4 shows how well the different models discriminated between SDTF and points of different vegetation types neighboring the biome. The models with

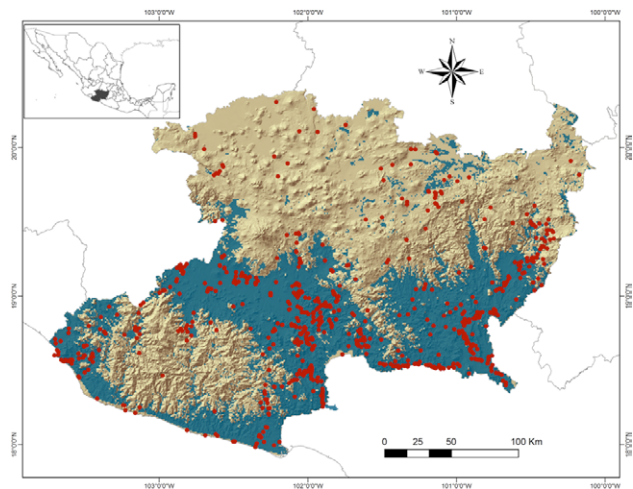


Figure 2. Potential distribution of the seasonally dry tropical forest (blue area) in the state of Michoacán obtained with the assembly of 76 individual species ecological niche models. The red circles show the occurrence points used in the modeling of the species considered.

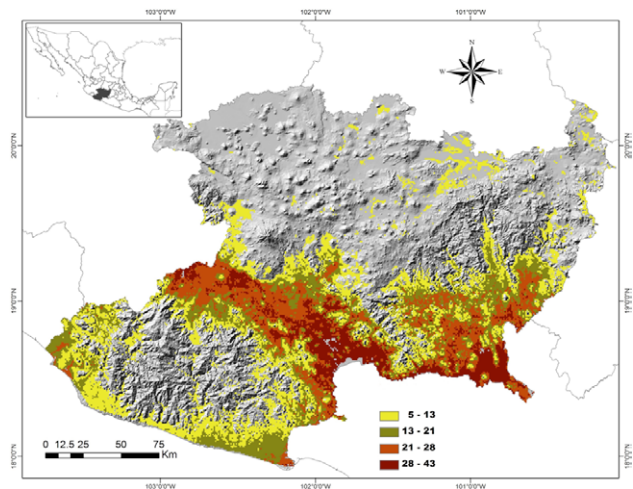


Figure 3. Number of species characteristic of the seasonally dry tropical forest (Table 1). The red color shows the areas of highest concentration of species (> 28), while the areas of lower concentration are shown in yellow.

the highest points' misclassification were those proposed by Rzedowski (1990) and Prieto-Torres and Soto-Rojas (2016), 65.7% and 63.6%, respectively. Our model presented intermediate confusion (36%), compared to the other models. The model that presented the lowest commission error was INEGI (2003) with 14.8% confusion. The type of vegetation with the highest number of points erroneously classified as SDTF was *Quercus* forest.

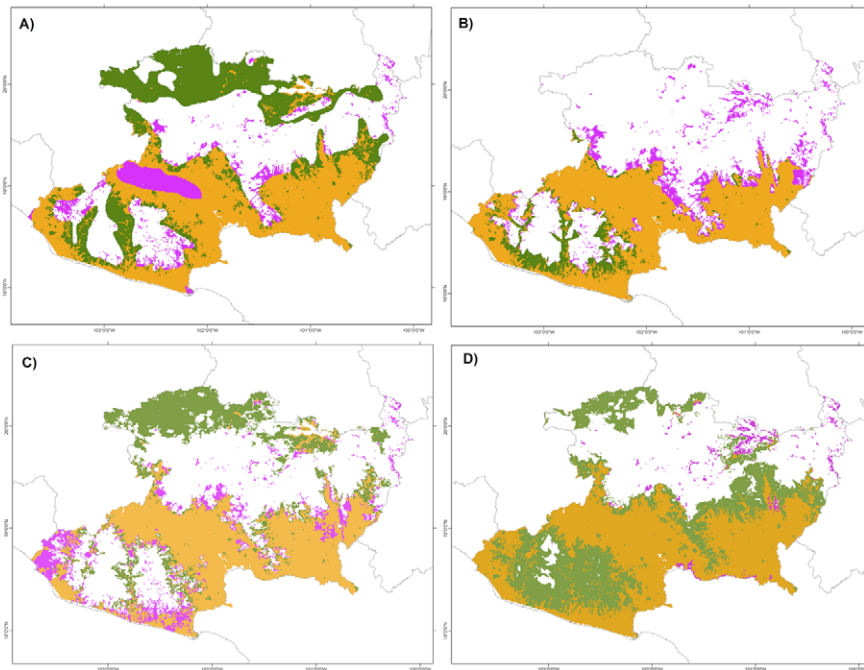


Figure 4. Comparison of the different proposals of distribution of the seasonally dry tropical forest in Michoacán with the potential distribution obtained in this study: A) Rzedowski (1990) model, B) INEGI (2003), C) Villaseñor and Ortiz (2014), D) Prieto-Torres and Soto-Rojas (2016). In all figures the orange color indicates the coincident areas, the lilac color indicates areas predicted by our model non-coincident with the proposed model, and the green areas the exclusive areas of each model not recorded by our model.

Table 3

Percentages of the statistical tests using 183 occurrence points of species characteristic of the seasonally dry tropical forest in Michoacán. These records were not used for models generation of the 76 species. The percentage of omission errors of each model is also indicated, considering in addition the total of points used to run the models.

Models	Kappa	Binomial test	p -value < 0.5	% Omission
Rzedowski	0.35	0.62	8.80×10^{-07}	18.36
INEGI	0.54	0.58	0.0001004	17.98
Villaseñor and Ortiz	0.41	0.55	0.001501	19.70
Prieto-Torres and Soto-Rojas	0.34	0.81	2.20×10^{-16}	9.13
This paper	1.00	0.74	2.20×10^{-16}	16.84

Table 4

Collecting points classified as part of the biome by different proposals of distribution of the seasonally dry tropical forest. The number of erroneous points assigned to the biome per type of vegetation and the percentage of confusion considering the total points erroneously classified are indicated. BTSC = Tropical sub-deciduous forest; BQ = *Quercus* forest, BMM = humid mountain forest, MXE = xerophytic scrub.

Models	BTSC (373)	BQ (1,806)	BMM (118)	MXE (23)	Total	% confusion
Rzedowski	275	1,214	21	14	1,524	65.7
INEGI	368	19	1	7	345	14.9
Villaseñor and Ortiz	156	636	2	13	807	34.8
Prieto-Torres and Soto-Rojas	370	1,087	9	9	1,475	63.6
This paper	291	535	1	7	834	35.9

Discussion

The assembly of individual ecological niche models turned out to be a useful technique to find the environmentally suitable distribution of the SDTF in Michoacán. Such assembly of models is a widely used technique that identifies spatial patterns of species that characterize a community (Clark et al., 2014; D'Amen et al., 2015; Ferrier & Guisan, 2006; Franklin, 1995; Guisan & Zimmermann, 2000; Jiménez-Alfaro et al., 2018). A biome like SDTF is characterized by having well defined climatic conditions (Challenger & Soberón, 2008; Gurevitch et al., 2002); the species endemic or characteristic to it, due to their specific requirements, are good indicators of environmental conditions of the biome. One disadvantage attributed to this technique is that the community-level prediction can accumulate the errors from each individual model (Pottier et al., 2013). However, the assembly of models presents an important advantage in that the species that make up the community are identified, and therefore its composition is known, which for conservation issues is of great importance (D'Amen et al., 2015).

The use of species that are endemic to and characteristic of the SDTF in Michoacán to delimit its distribution resulted in a statistically well-supported model. This is mainly because the species considered are characterized by their specific environmental conditions; the smaller geographical ranges characterizing such species, offer complete scenarios about their distribution, allowing to capture more realistically the greater part of their environmental niche (Lomba et al., 2010). Lomba et al. (2010) discussed that the use species that are rare or have a restricted distribution, as is the case in species that are endemic or exclusive to a habitat, provides better projections across scales. These models with good predictive capacity lead to more accurate predictions about the dynamics of these species and the communities where they live.

In the geographical overlap test, although our model had a high degree of coincidence with that proposed by Prieto-Torres and Rojas-Soto (2016), the Prieto-Torres and Rojas-Soto model had a larger overestimated area. These authors included in their model species not restricted to the SDTF biome, such as *Enterolobium cyclocarpum* (Jacq.) Griseb. and *Lysiloma divaricatum* (Jacq.) J.F. Macbr., which are also recorded in contiguous temperate forests (Gopar-Merino & Velázquez, 2016). The use of widely distributed species can lead to multiple vegetation types being lumped together as one (Clark et al., 2014). The overestimation caused by the use of species with wide distribution increases the probability that when comparing an adjusted model characteristic of a biome, with a model

in which species shows broad environmental conditions, results have greater coincidence in area. On the other hand, the overestimation observed in the proposals of Rzedowski (1990), INEGI (2003) and Villaseñor and Ortiz (2014), mostly observed in the southern and northwestern parts of the state, may be the result of errors due mainly to severely fragmented habitats (Mas et al., 2017), e.g. difficulty in distinguishing SDTF from secondary vegetation derived from SDTF or subtropical scrubs derived from *Quercus* forests (Challenger & Soberón, 2008; Trejo & Dirzo, 2000).

The Kappa statistics suggested that the model most like ours was the INEGI (2003) model. This statistic estimates classification errors when using one model to predict the other, so each of these models had a moderate capacity to predict the other, which involves both identifying the location of the SDTF as well as the omission and commission errors. On the other hand, the low value of Kappa for the model of Prieto-Torres and Soto-Rojas (2016), may be the result of high rates of commission errors with respect to our model, that is, a high percentage of area predicted by their model but not by ours. The high rates of commission errors translate into overestimation (Anderson et al., 2003), which in this case may be due to the use of widely distributed species to generate the model.

The omission error test favors the proposal of Prieto-Torres and Soto-Rojas (2016), inevitable with an overestimated model. The model had a larger total predicted area, increasing the probability of more points falling within the area proposed by this model, compared to tighter models such as the Rzedowski model (1990) and our model. The low rates of omission errors of Prieto-Torres and Soto-Rojas (2016) model with respect to ours, may also be due to the less restrictive threshold they used ("fixed omission value 5") to obtain the binary models they used to determine the SDTF distribution. This lax omission value (0.05) contrasts with the more restricted values with which we convert the models to binary (minimum 0.09, maximum 0.7, average 0.4), which caused a higher rate of omission errors in our STDF model. Another test that also favors the proposal of Prieto-Torres and Soto-Rojas (2016) is the binomial test, also surely because the greater predicted surface increases the probability that more points fall within the area predicted by the model, and therefore more points are correctly classified when they should be considered incorrectly placed.

The commission errors test favors the INEGI (2003) proposal; this delimitation confuses the SDTF with other plant communities less frequently than the other models. In contrast, Rzedowski (1990) and Prieto-Torres and Soto-Rojas (2016) more frequently determine localities of other biomes as part of the SDTF (Table 4). It is likely

that including more generalist and widely distributed species resulted in a higher percentage of commission errors (64%), when compared with localities of neighbor biomes (Table 4). For example, *Cochlospermum vitifolium* (Willd.) Spreng., *Enterolobium cyclocarpum* (Jacq.) Griseb., *Haematoxylum brasiletto* H. Karst., *Ipomoea wolcottiana* Rose, and *Lysiloma watsonii* Rose, used in model's generation by these authors, are widely distributed in 3 or more biomes other than SDTF. These species' models include environmental conditions from other biomes, mistakenly transferring these characteristics to the analysis and leading to an overestimated model.

Broadly distributed species are difficult to model because both their ecological and environmental characteristics affect the accuracy of the model compared to species of restricted distribution (Hernández et al., 2006; Segurado & Araujo, 2004; Thuiller et al., 2004). A high rate of commission errors can have negative consequences for conservation because they would be defining areas as potentially important for a species or a group of species when that species or group of species is not actually present (Kramer-Schadt et al., 2013). Accordingly, the appropriate selection of the species to be considered to model a plant community with specific characteristics can help to obtain more accurate conclusions (Aitken et al., 2007; Lomba et al., 2010; Mateo et al., 2011).

An additional problem with the comparison of proposals on the geographical definition of the SDTF is that the techniques applied are difficult to replicate. Sometimes the proposals have been elaborated based on the opinion and experience of a group of experts (for example Rzedowski, 1990), or require high investment and too much time to conduct vegetation censuses to define it, which are not always convincing due to non-uniformity in data collection (for example INEGI, 2003). Although the model of Prieto-Torres and Rojas-Soto (2016) was favored by the highest number of validation tests, it does not mean that it is the model that best represents the distribution of the SDTF in Michoacán. These tests favored this proposal because it was biased by the overestimation of the model by using species of wide distribution. The overestimation led to more successfully predicted points in the omission and binomial errors tests, but it also resulted in greater confusion and that records of species from other contiguous biomes were classified as SDTF. This made it the second-least adequate proposal (following Rzedowski's) of the potential geographical distribution of SDTF in Michoacán.

Only 8.5% (2,856 km²) of the area with the highest concentration of species richness (> 28 species) is within a protected natural area (the Zicuirán-Infiernillo Natural Reserve). Given that SDTF is considered one of the most important and endangered biomes, especially due to its high

degree of endemism (Cué-Bär, Villaseñor, Morrone et al., 2006; DRYFLOR et al., 2016; Olson et al., 2000), results discussed here can be used to make proposals for ANPs in the areas with the highest concentration of endemic species. There are currently no protected areas in the Balsas Depression and the Southern Coastal Range, where there was a high concentration of characteristic species.

Modeling the potential distribution of species at the community level allows for effective forecasting of factors that threaten biodiversity such as climate change, as well as providing knowledge on the functioning of the ecosystem (D'Amen et al., 2015). That is why it is important to make a rigorous selection of the species integrated to the algorithms for modeling the distribution of communities. The consideration of endemic and characteristic species of a plant community, described by its specific requirements (for example the SDTF), provide results closer to the observed data than those found when widely distributed species are considered. More precise models will surely allow for more accurate conservation proposals of potentially important areas due to their species richness.

Acknowledgments

Conabio and the Instituto de Biología, UNAM, are grateful for access to the information stored in the SNIB-REMIB and UNIBIO databases, respectively, which formed a fundamental part of the analysis presented here. Lynna Kiere reviewed the English translation and made important suggestions. We appreciate the relevant observations made by two anonymous reviewers and associate editor.

References

- Aitken, M., Roberts, D. W., & Shultz, L. M. (2007). Modelling distributions of rare plants in the great basin, Western North America. *Western North American Naturalist*, 67, 26–38. [https://doi.org/10.3398/1527-0904\(2007\)67\[26:MDORPI\]2.0.CO;2](https://doi.org/10.3398/1527-0904(2007)67[26:MDORPI]2.0.CO;2)
- Anderson, R. P., Lew, D., & Peterson, A. T. (2003). Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling*, 162, 211–232. [https://doi.org/10.1016/S0304-3800\(02\)00349-6](https://doi.org/10.1016/S0304-3800(02)00349-6)
- Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33, 1677–1688. <https://doi.org/10.1111/j.1365-2699.2006.01584.x>
- Baldwin, R. A. (2009). Use of maximum entropy modeling in Wildlife Research. *Entropy*, 11, 854–866. <https://doi.org/10.3390/e11040854>
- Barve, N. (2008). Tool for Partial-ROC (Biodiversity Institute. Lawrence). Ver 1.0.
- Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T. et al. (2011). The crucial role of

- the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, 222, 1810–1819. <https://doi.org/10.1016/j.ecolmodel.2011.02.011>
- Bojorquez-Tapia, L., Azura, I., Ezcurra, E., & Flores-Villela, O. (1995). Identifying conservation priorities in Mexico through geographic information systems and modeling. *Ecological Applications*, 5, 215–231. <https://doi.org/10.2307/1942065>
- Castillo, M., Michán, L., & Martínez, A. L. (2014). La biocuración en biodiversidad: proceso, aciertos, errores, soluciones y perspectivas. *Acta Botanica Mexicana*, 108, 81–103. <https://doi.org/10.21829/abm108.2014.199>
- Challenger, A., & Soberón, J. (2008). Los ecosistemas terrestres. In J. Sarukhán (Ed.), *Capital natural de México. Vol. I. Conocimiento actual de la biodiversidad* (pp. 87–108). México City: Conabio.
- Chapman, A. D. (2005). *Principles and methods of data cleaning*. Report for the Global Biodiversity Information Facility 2004. Copenhagen: GBIF.
- Clark, J. S., Gelfand, A. E., Woodall, C. W., & Zhu, K. (2014). More than the sum of the parts: forest climate response from joint species distribution models. *Ecological Applications*, 24, 990–999. <https://doi.org/10.1890/13-1015.1>
- Conabio (Comisión Nacional para el Conocimiento y Uso de la Biodiversidad). (2018). *Sistema Nacional de Información sobre Biodiversidad*. Registros de ejemplares. Ciudad de México: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.
- Cornejo-Tenorio, G., & Ibarra-Manríquez, G. (2017). Flora of the core areas in the Monarch Butterfly Biosphere Reserve, Mexico: composition, geographical affinities and beta diversity. *Botanical Sciences*, 95, 103–129. <https://doi.org/10.17129/botsci.803>
- Cruz-Cárdenas, G., López-Mata, L., Ortiz-Solorio, C. A., Villaseñor, J. L., Ortiz, E., Silva, J. T. et al. (2014). Interpolation of Mexican soil properties at a scale of 1: 1,000,000. *Geoderma*, 213, 29–35. <https://doi.org/10.1016/j.geoderma.2013.07.014>
- Cruz-Cárdenas, G., López-Mata, L., Villaseñor, J. L., & Ortiz, E. (2014). Potential species distribution modeling and the use of principal component analysis as predictor variables. *Revista Mexicana de Biodiversidad*, 85, 189–199. <https://doi.org/10.7550/rmb.36723>
- Cué-Bär, E., Villaseñor, J. L., Arredondo, L., Cornejo, G., & Ibarra-Manríquez, G. (2006). La flora arbórea de Michoacán, México. *Boletín de la Sociedad Botánica de México*, 78, 47–48. <https://doi.org/10.17129/botsci.1721>
- Cué-Bär, E. M., Villaseñor, J. L., Morrone, J. J., & Ibarra-Manríquez, G. (2006). Identifying priority areas for conservation in Mexican tropical deciduous forest based on tree species. *Interciencia*, 31, 712–719.
- D'Amen, M., Dubuis, A., Fernandes, R. F., Pottier, J., Pellissier, L., & Guisan, A. (2015). Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models. *Journal of Biogeography*, 42, 1255–1266. <https://doi.org/10.1111/jbi.12485>
- D'amen, M., Rahbek, C., Zimmermann, N. E., & Guisan, A. (2017). Spatial predictions at the community level: from current approaches to future frameworks. *Biological Reviews*, 92, 169–187. <https://doi.org/10.1111/brv.12222>
- Dormann, C. F., McPherson, J. M., Araujo, M. B., Bivand, R., Bolliger, J., Carl, G. et al. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30, 609–628. <https://doi.org/10.1111/j.2007.0906-7590.05171.x>
- DRYFLOR, Banda, K., Delgado-Salinas, A., Dexter, K. G., Linares-Palomino, R., Oliveira-Filho, A. et al. (2016). Plant diversity patterns in neotropical dry forests and their conservation implications. *Science*, 353, 1383–1387. <https://doi.org/10.1126/science.aaf5080>
- Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A. et al. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129–151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>
- Elith, J., Phillips, S., Hastie, T., Dudík, M., Chee, Y., & Yates, C. A. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distribution*, 17, 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>
- ESRI (Environmental Systems Research Institute). (2010). ArcGIS 10.0. Redlands, California.
- Ferrier, S. (2002). Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? *Systematic Biology*, 51, 331–363. <https://doi.org/10.1080/10635150252899806>
- Ferrier, S., & Guisan, A. (2006). Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, 43, 393–404. <https://doi.org/10.1111/j.1365-2664.2006.01149.x>
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24, 38–49. <https://doi.org/10.1017/S0376892997000088>
- Franklin, J. (1995). Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography*, 19, 474–499. <https://doi.org/10.1177/030913339501900403>
- Gopar-Merino, L. F., & Velázquez, A. (2016). Componentes del paisaje como predictores de cubiertas de vegetación: estudio de caso del estado de Michoacán, México. *Investigaciones Geográficas, Boletín del Instituto de Geografía, UNAM*, 90, 75–88. <https://doi.org/10.14350/rig.46688>
- Guisan, A., & Theurillat, J. P. (2000). Equilibrium modeling of alpine plant distribution: how far can we go?. *Phytocoenologia*, 30, 353–384. <https://doi.org/10.1127/phyto/30/2000/353>
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8, 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147–186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)
- Gurevitch, J., Scheiner, S. M., & Fox, G. A. (2002). *The ecology of plants*. Sunderland, Massachusetts: Sinauer.
- Henderson, E. B., Ohmann, J. L., Gregory, M. J., Roberts, H.

- M., & Zald, H. (2014). Species distribution modelling for plant communities: stacked single species or multivariate modelling approaches?. *Applied Vegetation Science*, 17, 516–527. <https://doi.org/10.1111/avsc.12085>
- Hernández, P. A., Graham, C. H., Master, L. L., & Albert, D. L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, 29, 773–785. <https://doi.org/10.1111/j.0906-7590.2006.04700.x>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–1978. <https://doi.org/10.1002/joc.1276>
- Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., & Guisan, A. (2006). Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, 199, 142–152. <https://doi.org/10.1016/j.ecolmodel.2006.05.017>
- INEGI (Instituto Nacional de Estadística y Geografía). (2003). *Conjunto de datos vectoriales de la carta de vegetación primaria*. Escala 1:1,000,000 (NIVEL I y II). Retrieved on September 5th, 2017 from: <http://www.inegi.org.mx/geo/contenidos/>
- INEGI (Instituto Nacional de Estadística y Geografía). (2014). *Fisiografía. Datos vectoriales*. Escala 1:1000000. Retrieved on August 10th, 2018 from: <http://www.inegi.org.mx>
- Jenness, J., & Wynne, J. J. (2005). Cohen's Kappa and classification table metrics 2.0: an ArcView 3x extension for accuracy assessment of spatially explicit models: U.S. Geological Survey Open-File Report OF 2005-1363.
- Jenness, J., & Wynne, J. J. (2007). Kappa Analysis (Kappa_stats.avx) extension for ArcView. 3.x. Jenness Enterprises.
- Jiménez-Alfaro, B., Suárez-Seoane, S., Chytrý, M., Hennekens, S. M., Willner, W., Hájek, M. et al. (2018). Modelling the distribution and compositional variation of plant communities at the continental scale. *Diversity and Distributions*, 24, 978–990. <https://doi.org/10.1111/ddi.12736>
- Jiménez-Valverde, A., & Lobo, J. M. (2007). Threshold criteria for conversion of probability of species presence to either-or presence-absence. *Acta Oecologica*, 31, 361–369. <https://doi.org/10.1016/j.actao.2007.02.001>
- Johnson, J. B., & Omland, K. S. (2004). Model selection in ecology and evolution. *Trends in Ecology & Evolution*, 19, 101–108. <https://doi.org/10.1016/j.tree.2003.10.013>
- Kramer-Schadt, S., Niedballa, J., Pilgrim, J. D., Schroder, B., Lindenborn, J., Reinfelder, V. et al. (2013). The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions*, 19, 1366–1379. <https://doi.org/10.1111/ddi.12096>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Leavenworth, W. C. (1946). A preliminary study of the vegetation of the region between Cerro Tancitaro and the Rio Tepalcatepec, Michoacán, Mexico. *American Midland Naturalist*, 36, 137–206.
- Liu, C., Newell, G., & White, M. (2016). On the selection of thresholds for predicting species occurrence with presence-only data. *Ecology and Evolution*, 6, 337–348. <https://doi.org/10.1002/ece3.1878>
- Liu, C., White, M., & Newell, G. (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, 40, 778–789. <https://doi.org/10.1111/jbi.12058>
- 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. <https://doi.org/10.1111/j.1466-8238.2007.00358.x>
- Lomba, A., Pellissier, L., Randin, C., Vicente J., Moreira, F., Honrado, J. et al. (2010). Overcoming the rare species modelling paradox: a novel hierarchical framework applied to an Iberian endemic plant. *Biological Conservation*, 143, 2647–2657. <https://doi.org/10.1016/j.biocon.2010.07.007>
- López-Mata, L., Villaseñor, J. L., Cruz-Cárdenas, G., Ortiz, E., & Ortiz-Solorio, C. (2012). Predictores ambientales de la riqueza de especies de plantas del bosque húmedo de montaña de México. *Botanical Sciences*, 90, 27–36. <https://doi.org/10.17129/botsci.383>
- Mas, J. F., Lemoine-Rodríguez, R., González, R., López-Sánchez, J., Piña-Garduño, A., & Herrera-Flores, E. (2017). Evaluación de las tasas de deforestación en Michoacán a escala detallada mediante un método híbrido de clasificación de imágenes SPOT. *Madera y Bosques*, 23, 1–13. <https://doi.org/10.21829/myb.2017.2321472>
- Mateo, R. G., Felicísimo, A. M., & Muñoz, J. (2011). Modelos de distribución de especies: una revisión sintética. *Revista Chilena de Historia Natural*, 84, 217–240.
- Méndez-Toribio, M., Martínez-Cruz, J., Cortés-Flores, J., Rendón-Sandoval, F. J., & Ibarra-Manríquez, G. (2014). Composición, estructura y diversidad de la comunidad arbórea del bosque tropical caducifolio en Tziritzicuaró, Depresión del Balsas, Michoacán, México. *Revista Mexicana de Biodiversidad*, 85, 1117–1128. <https://doi.org/10.7550/rmb.43457>
- Miranda, F., & Hernández, X. E. (1963). Los tipos de vegetación de México y su clasificación. *Boletín de la Sociedad Botánica de México*, 28, 29–179. <http://dx.doi.org/10.17129/botsci.1084>
- Muñoz, J., & Felicísimo, A. M. (2004). Comparison of statistical methods commonly used in predictive modeling. *Journal of Vegetation Science*, 15, 285–292. <https://doi.org/10.1111/j.1654-1103.2004.tb02263.x>
- Olson, D. M., Dinerstein, E. Abell, R., Allnutt, T., Carpenter, C., McClenachan, L. et al. (2000). *The global 200: a representation approach to conserving the Earth's distinctive ecoregions*. Washington D.C.: Conservation Science Program, World Wildlife Fund-US. <https://doi.org/10.1046/j.1523-1739.1998.012003502.x>
- Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263, 641–646. <https://doi.org/10.1126/science.263.5147.641>
- Pearson, R. G., Raxworthy, C. J., Nakamura, M., & Peterson, A. T. (2007). Predicting species distributions from small

- numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, 34, 102–117. <https://doi.org/10.1111/j.1365-2699.2006.01594.x>
- Pearson, R. G., Thuiller, W., Araújo, M. B., Martínez-Meyer, E., Brotons, L., McClean, C. et al. (2006). Model-based uncertainty in species range prediction. *Journal of Biogeography*, 33, 1704–1711. <https://doi.org/10.1111/j.1365-2699.2006.01460.x>
- Peterson, A. T., & Nakazawa, Y. (2008). Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography*, 17, 135–144. <https://doi.org/10.1111/j.1466-8238.2007.00347.x>
- Peterson, A. T., Papeş, M., & Soberón, J. (2008). Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, 213, 63–72. <https://doi.org/10.1016/j.ecolmodel.2007.11.008>
- Peterson, A. T., & Soberón, J. (2012). Species distribution modeling and ecological niche modeling: getting the concepts right. *Natureza & Conservação*, 10, 102–107. <http://dx.doi.org/10.4322/natcon.2012.019>
- Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M. et al. (2011). *Ecological niches and geographical distributions*. Monographs in Population Biology 49. Princeton: Princeton University Press.
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31, 161–175. <https://doi.org/10.1111/j.0906-7590.2008.5203.x>
- Pontius, Jr, R. G., & Millones, M. (2011). Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32, 4407–4429. <https://doi.org/10.1080/01431161.2011.552923>
- Pottier, J., Dubuis, A., Pellissier, L., Maiorano, L., Rossier, L., Randin, C. F. et al. (2013). The accuracy of plant assemblage prediction from species distribution models varies along environmental gradients. *Global Ecology and Biogeography*, 22, 52–63. <https://doi.org/10.1111/j.1466-8238.2012.00790.x>
- Prieto-Torres, D. A., & Rojas-Soto, O. R. (2016). Reconstructing the Mexican tropical dry forests via an autoecological niche approach: reconsidering the ecosystem boundaries. *Plos One*, 11, e0150932. <https://doi.org/10.1371/journal.pone.0150932>
- R Development Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.r-project.org/>
- Rodríguez, J. L. S., & Espinosa, G. J. (1995). *Listado florístico del estado de Michoacán*. Sección I (Gymnospermae; Angiospermae: Acanthaceae-Commelinaceae). Flora del Bajío y de regiones adyacentes. Fascículo Complementario VI. Pátzcuaro, Michoacán: Instituto de Ecología, A.C./ Centro Regional del Bajío.
- Rodríguez, J. L. S., & Espinosa, G. J. (1996). *Listado florístico del estado de Michoacán*. Sección V (Angiospermae: Najadaceae-Zygophyllaceae). Flora del Bajío y de regiones adyacentes. Fascículo complementario XV. Pátzcuaro, Michoacán: Instituto de Ecología, A.C./ Centro Regional del Bajío.
- Rzedowski, J. (1978). *Vegetación de México*. Mexico D.F.: Limusa.
- Rzedowski, J. (1990). *Vegetación potencial*. Atlas nacional de México. Mexico D.F.: Instituto de Geografía, UNAM.
- Segurado, P., & Araujo, M. B. (2004). An evaluation of methods for modelling species distributions. *Journal of Biogeography*, 31, 1555–1568. <https://doi.org/10.1111/j.1365-2699.2004.01076.x>
- Soberón, J., Osorio-Olvera, L., & Townsend, P. (2017). Diferencias conceptuales entre modelación de nichos y modelación de áreas de distribución. *Revista Mexicana de Biodiversidad*, 88, 237–441. <http://dx.doi.org/10.22201/ib.20078706e.2017.2.1987>
- Soberón, J. M., & Peterson, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2, 1–10. <https://doi.org/10.17161/bi.v2i0.4>
- Thuiller, W., Lavorel, S., Midgley, G., Levergne, S., & Rebelo, T. (2004). Relating plant traits and species distributions along bioclimatic gradients for 88 *Leucadendron* taxa. *Ecology*, 85, 1688–1699. <https://doi.org/10.1890/03-0148>
- Trejo, I., & Dirzo, R. (2000). Deforestation of seasonally dry tropical forest: a national and local analysis in Mexico. *Biological Conservation*, 94, 133–142. [https://doi.org/10.1016/S0006-3207\(99\)00188-3](https://doi.org/10.1016/S0006-3207(99)00188-3)
- Villaseñor, J. L. (2016). Checklist of the native vascular plants of Mexico. *Revista Mexicana de Biodiversidad*, 87, 559–902. <https://doi.org/10.1016/j.rmb.2016.06.017>
- Villaseñor, J. L., & Ortiz, E. (2014). Biodiversidad de las plantas con flores (División Magnoliophyta) en México. *Revista Mexicana de Biodiversidad*, 85(Supl.), S134–S142. <https://doi.org/10.7550/rmb.31987>
- Weber, T. C. (2011). Maximum entropy modeling of mature hardwood forest distribution in four U.S. states. *Forest Ecology and Management*, 261, 779–788. <https://doi.org/10.1016/j.foreco.2010.12.009>
- Zimmerman, B. J. (2000). *Attaining self-regulation. A social cognitive perspectives*. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). London: Academic Press. <https://doi.org/10.1016/B978-012109890-2/50031-7>