



## Using one vs. many, sensitivity and uncertainty analyses of species distribution models with focus on conservation area networks



Leticia M. Ochoa-Ochoa<sup>a,\*</sup>, Oscar A. Flores-Villela<sup>a</sup>, Juan E. Bezaury-Creel<sup>b,1</sup>

<sup>a</sup> Museo de Zoología “Alfonso L. Herrera”, Facultad de Ciencias, UNAM. Circuito exterior, Ciudad Universitaria, Del. Coyoacán., México, D.F. CP. 04510, Mexico

<sup>b</sup> The Nature Conservancy México, Ricardo Palmerín #110, Colonia Guadalupe Inn., México, D.F., CP. 01020, Mexico

### ARTICLE INFO

#### Article history:

Received 16 May 2015

Received in revised form

21 September 2015

Accepted 27 October 2015

#### Keywords:

Amphibians

Commission

Conservation planning

Environmental niche modelling

Mexico

Omission

### ABSTRACT

Species Distribution Models (SDM) are currently common currency as *proxies* of species distribution range, and using consensus among different algorithms is becoming the latest tendency. This information is frequently used to estimate conservation status or for conservation planning. Nonetheless, different algorithms have huge variation in the outcomes. Usually experts determine whether or not a model is *accurate*, often followed by a trimming process. However, this accuracy estimation cannot be reproduced. Using Mexican endemic amphibians we evaluate the performance of nine modelling algorithms (Artificial Neural Networks, Classification Tree Analysis, Flexible Discriminant Analysis, Generalised Boosting Model, Generalised Linear Models, Multiple Adaptive Regression Splines, MaxEnt, RandomForest, Surface Range Envelope), their strict geographic consensus, locality records and simple convex-hull areas through comparison of: (1) their presence/absence within Mexico's governmental protected areas, (2) range sizes projected, and (3) differences in estimated richness by all methods. We conducted all good practices *prior* modelling but removed the trimming factor *after* modelling to make the process *repeatable*. Presence-absence threshold was determined through the use of the receiver-operating characteristic (ROC). Presence within conservation network of strict consensus and locality records was similar which indicates an over-fitting of the former, the rest of the algorithms performed similarly, with exception of Surface Range Envelope. Richness patterns varied greatly among algorithms. Distribution borders were the areas with higher sensitivity. MaxEnt obtained the highest performance in omission but consensus performed best in correctly predicting species ranges. Closer interaction between curators and modelers would increase SDMs accuracy, which would improve conservation planning effectiveness.

© 2015 Published by Elsevier B.V.

### 1. Introduction

Humans, like any other species, can affect their environment. But since the Neolithic Revolution around 10,000 BC, human populations have increasingly modified the ecosystems they live in. Anthropogenic activities have been having widespread and diverse effects on biodiversity, particularly after the Industrial Revolution of the 19th Century with the over-use of all kinds of fossil fuels—coal, oil, and their derivatives—that modify climate. Expansions of domesticated ungulates (cattle, sheep, and goats) and the agricultural frontier due to technological advances have transformed the world's ecosystems into fragmented semi-natural

landscapes (Dirzo and Raven, 2003; Dornelas, 2010; Ehrlich and Pringle, 2008). Moreover, the increase in use of chemicals on farmlands (pesticides and agrochemicals), the production of solid wastes, overfishing, the introduction of invasive non-native species, as well as the ground, water and air pollution have significantly modified almost every place on Earth, putting several species on a threatened status (e.g., Butchart et al., 2010; Wake and Vredenburg, 2008).

Over the last two decades, amphibians have come to be regarded as one of the most threatened taxonomic groups (Stuart et al., 2004). Land use change has been identified as the major threat for amphibian diversity loss (Alford and Richards, 1999; e.g., Alford et al., 2007; Curado et al., 2011; Cushman, 2006; Gallant et al., 2007; Johnson et al., 2011). Moreover, because they are rather small poikilotherms animals depending on water or humidity for reproduction, they show several of the most labile traits towards environmental changes (Pearson et al., 2014). In addition, Mexico is considered as one of the fourth-top countries with the highest amount of threatened species (Frías-Alvarez et al., 2010;

\* Corresponding author at: Current address: Center of Macroecology, Evolution and Climate, University of Copenhagen, Denmark, Universitetsparken 15, 2100 Copenhagen, Denmark.

E-mail address: [leticia.ochoa@snm.ku.dk](mailto:leticia.ochoa@snm.ku.dk) (L.M. Ochoa-Ochoa).

<sup>1</sup> Tel.+ 52 55 56 22 48 32

Stuart et al., 2004). It has 375 species of amphibians, of which 252 are endemic species, 80% are micro-endemics, i.e., species with very restricted distribution range of less than 10,000 km<sup>2</sup> (Frías-Alvarez et al., 2010; Johnson et al., 2001; Ochoa-Ochoa et al., 2009, 2014; Ochoa-Ochoa and Flores-Villela, 2006). Endemics and/or highly rare species are expected on grounds of parsimony to be more prone to extinction due to habitat loss and other drivers (Brook et al., 2008; Payne and Finnegan, 2007; Sodhi et al., 2008).

Biodiversity conservation requires managing an integrated landscape matrix that includes areas allocated to both production and protection (Margules and Pressey, 2000). A key strategy for protecting biodiversity from human related pressures has been the establishment and maintenance of protected areas as refuges for species, ecosystems and the natural processes upon which their existence depend (e.g., Chape et al., 2005; Ochoa-Ochoa et al., 2009; Rodrigues et al., 2004a). In the past, the selection of places to be protected was to some extent a fortuitous and capricious process. Nowadays, science based approaches such as systematic conservation planning can involve different steps, where species distribution ranges become the basic information for any conservation plan (Butchart et al., 2012; Margules and Pressey, 2000; Rodrigues et al., 2004b).

Nonetheless most places on Earth currently still present Linnean and Wallacean shortfalls (Whittaker et al., 2005). This means that we do not know the full list of occurring species for the vast majority of sites (Linnean shortfall) nor where the species are distributed (Wallacean shortfall). In order to overcome these deficiencies, species distribution modelling (SDM) based on niche modelling using primarily environmental layers has become the bread and butter for most scientists involved in conservation planning (e.g., Koleff et al., 2008; Richardson and Whittaker, 2010).

Species distribution modelling is a booming field and, as a consequence, a wide range of algorithms to model species distribution has been developed. It has been reported that different algorithms perform differently, not only in terms of ‘accuracy’ with respect to a specific statistical measure (i.e., ROC, partial ROC, TSS, etc.) but in terms of ‘distributional performance’, size, form and continuity of the resultant area of distribution (Elith et al., 2006; Pearson et al., 2006; Peterson et al., 2011). Ultimately, it has been left to the experts to decide when an algorithm is modelling ‘accurately’ a species distribution. This situation should not be a problem for widely spread and well-known species, but unfortunately this type of assessment can be problematic for rare or poorly known species.

This type analysis becomes especially relevant, since the use of species distribution modelling tools, based on climatic niche modelling, are rapidly gaining critical importance in resilient conservation area design, in other words projections for the future, including climate change (e.g., Garcia et al., 2014b). Due to the static nature of protected areas, they must include future niches for species derived from global climate change, in order to remain relevant in the future. Thus climatic niche species distribution modelling is a critical tool for the design of effective protected area networks.

Here we explored nine different algorithms (GLM, CTA, ANN, SRE, GBM, randomForest, FDA, MARS and MaxEnt) including two of the three most common used for species distribution modelling (i.e., MaxEnt and randomForest). In this study we did not include the Genetic Algorithm for Rule-set Production (Stockwell and Peters, 1999). Our aim is to compare systematically and quantitatively the performance of these algorithms in Mexican governmental protected area networks without the trimming process, which is difficult to standardise in order to replicate. A trimming process consists in cut (like cookie-cutter) the projection of the SDM based on some sort of ecogeographical aspect, i.e., vegetation types, altitude, ecoregions, etc., or based on the expert knowledge (e.g., Anderson and Martínez-Meyer, 2004;

Anderson et al., 2002; Jiménez-Valverde et al., 2010; Peterson et al., 2011; Velásquez-Tibatá et al., 2013). To accomplish this goal, we evaluated the proportion of species conserved in Mexican protected area networks (i.e., evaluate spatial projection) based on: (1) locality records, (2) convex Hulls of each species, (3) each of the nine algorithms, and (4) the strict geographic consensus. We also (5) compared the size of the geographic projections of the algorithms and finally (6) we assessed the differences in richness from the cells containing locality records and the estimated richness of all algorithms including the consensus.

We used Mexican endemic amphibians as an example group for two main reasons: we wanted to model complete distributions, and because the existence of a high percentage of endemic amphibians (Frías-Alvarez et al., 2010; Parra-Olea et al., 2014), represents a large and robust sample.

## 2. Materials and methods

### 2.1. Database

To achieve the proposed goals, we used the most comprehensive database of locality records for the endemic amphibians of Mexico. This database includes records of more than 47 collections. Most of these were obtained directly from the museums. This database was refined and updated both geographically and taxonomically. Geographical refinement consisted of a visual spatial inspection of species by specie locality records with ArcMap 10 and Access. During this process, the records were compared against the known distribution addressed in literature. If a record was geographically distant from the ‘known distribution range’, then it was removed from the database. We decided to use this drastic method to reduce uncertainty in subsequent analyses. It is worth mentioning that when we had access to a specimen corresponding to a record geographically distant, all data -including animal determination- was verified. In some cases, when both specimen determination and locality georeference were right, we kept the original record. In other cases, where we could correct the record either by actualizing determination or locality georeferenciation, records were mostly kept; but there were some cases that had to be eliminated. Taxonomic revision consisted of nomenclature updated according to various sources, mainly following taxonomy proposed by Frost (2014). The database contains 77,812 locality records of 252 species. From this database, we selected unique records based on climatic variables grain/cell size 0.083° (approx. 1 km × 1 km). This was done to avoid bias during the modelling process towards areas with high amount of records. Using the complete database of unique-gridcell records and based on correlation we selected ten of the 19-bioclimatic variables of Worldclim (Hijmans et al., 2005; <http://www.worldclim.org>) that were the less covariant based on the records of the whole database. We used the same set of variables to model all species, being: (i) isothermality ((mean diurnal range/temperature annual range) × 100); (ii) temperature seasonality (standard deviation × 100); (iii) temperature annual range (max temperature of warmest month—min temperature of coldest month); (iv) mean temperature of driest quarter; (v) mean temperature of warmest quarter; (vi) annual precipitation; (vii) mean precipitation of driest month; (viii) mean precipitation of driest quarter; (ix) mean precipitation of warmest quarter; and (x) mean precipitation of coldest quarter.

### 2.2. Accessible area

Before the actual modelling, we selected a specific extent of the area of distribution modelling or the accessible area (Barve et al., 2011; “M”, Soberón and Nakamura, 2009), because it has

Download English Version:

<https://daneshyari.com/en/article/6296299>

Download Persian Version:

<https://daneshyari.com/article/6296299>

[Daneshyari.com](https://daneshyari.com)