Species Distribution Modelling for Conservation and Restoration

Sofie Costin

July 21, 2020

Species Distribution Modelling for Conservation and Restoration
The history, strengths, and limitations of species distribution modelling in ecological applications

Sofie Costin

Abstract

The use of species distribution models to guide conservation efforts has increased in recent years due to the increased accessibility and cost-effectiveness when compared with other methods. Many ecologists support the use of species distribution modelling and recognize the benefits of modelling before undertaking complex and expensive conservation efforts such as species translocations and pest management schemes, however, challenges remain in ensuring that models are as accurate as possible, and the model type used suits the data available.

1 Introduction

It has long been understood that an association exists between species and their environment (von Humboldt & Bonpland, 1807). Climate and geographical factors have been often used to explain the distribution of plant and animal species over the world, however it is only recently that attempts to quantify these relationships have been made (Guisan & Zimmermann, 2000). Predictive ecological models, otherwise known as niche models or species distribution models (SDMs) have become a widely used tool for the planning of conservation strategies such as pest management schemes and species translocations in recent years (McPherson, Jetz, & Rogers, 2004; Pearson, 2007; Stockwell & Peterson, 2002; Thuiller, Lafourcade, Engler, & Araújo, 2009). In short, SDMs assess the relationship between environmental conditions and species occurrences, and then estimate the spatial distribution of habitats suited to the study species outside of the species' current range (Kearney & Porter, 2009; Pearson, 2007).

While the use of SDMs can reduce the time and cost associated with conservation research, and conservation managers are relying increasingly on SDMs to inform their conservation strategies (Stockwell & Peterson, 2002), SDMs are by no means a one-stop solution to all conservation issues. Variables such as species range size (McPherson et al., 2004), sample size (Elith et al., 2006; Stockwell & Peterson, 2002), and sample (Bean, Stafford, & Brashares, 2012) all impact on the reliability of the model projections, and a range of model types have been developed with these limitations in mind (Elith & Graham, 2009; Elith et al., 2006). Therefore, it is imperative that conservation managers using SDMs to guide their actions a) ensure they are using the modelling method most suited to the data they have available, and b) assess the accuracy of models before acting upon the results (Elith & Graham, 2009).

2 History of species distribution modelling methods

SDMs are widely used across marine, freshwater, and terrestrial applications (Elith & Leathwick, 2009). Commonly, SDMs are used to predict the spread of pest organisms and to identify suitable habitats for the translocations of species outside the species' current range (Sutherst, 2014).

With advances in technology, including improved statistical techniques and geographical information systems (GIS), has come the implementation of new and improved modelling methods (Elith & Graham, 2009; Guisan & Zimmermann, 2000). The complexity of these models has increased over time, from the matching of simple environmental variables, to fitting non-linear relationships between species presence and environmental, ecological and climate variables (Elith & Graham, 2009). Multiple regression and generalized multiple regression are popular methods of modelling species distributions, and neural networks, locally weighted approaches and environmental envelope models are also commonly used (Guisan & Zimmermann, 2000). Additionally, weighted ensemble models (e.g. BIOMOD2) created from a combination of these aforementioned processes have been developed to increase model accuracy (Thuiller et al., 2009). Machine learning techniques are on the rise and it is likely that new, more accurate, methods with a wider range of applications will continue to be developed (Elith & Graham, 2009).

3 Strengths and limitations of species distribution modelling

SDMs can provide a useful, reliable, cost effective method of estimating species distribution (Elith & Leathwick, 2009). The selection of a suitable modelling method is key to ensuring that the results from the model provide an accurate representation of likely real-world phenomena (Elith & Leathwick, 2009; Guisan & Zimmermann, 2000). Discrepancies in predictive performance between modelling techniques are large, making it difficult for researchers without extensive modelling experience to choose the most reliable model for their needs (Hao, Elith, Guillera-Arroita, & Lahoz- Monfort, 2019; Thuiller et al., 2009). Additionally, small sample size and sampling bias can substantially reduce model accuracy (Bean et al., 2012; Stockwell & Peterson, 2002). This is particularly troublesome when attempting to model suitable habitats for a rare species (Lomba et al., 2010).

It is also essential that model results are tested for goodness of fit and accuracy before results are acted on (Thuiller et al., 2009). Model assessment techniques such as bootstrap and cross-validation and receiver-operating characteristic (ROC) plots have been implemented within modelling processes to test the accuracy of model outcomes (Guisan & Zimmermann, 2000).

Recently developed ensemble forecasting methods (such as the BIOMOD2 platform) can provide a solution to these limitations (Thuiller et al., 2009). Ensemble models combine the predictions of many different modelling methods and weight the models according to cross-validation, calibration, goodness of fit and accuracy (Hao et al., 2019; Thuiller et al., 2009).

While ensemble modelling may appear to be the answer to all SDM limitations, the ability of ensemble models to increase the performance of SDMs is debated in recent literature (Hao et al., 2019; Hao, Elith, Lahoz-Monfort, & Guillera-Arroita, 2020). Studies have shown that combining and weighting SDMs in an ensemble approach does not necessarily improve the reliability of the model results, and on occasion, machine learning methods on their own can provide more reliable and accurate results (Hao et al., 2019; Hao et al., 2020).

This is likely where the model selection process comes in to play – while ensemble models can produce results slightly less reliable results than using a specific modelling method (Hao et al., 2019; Hao et al., 2020), extensive research is required before an ecologist, conservationist, or environmental scientist will be able to readily select the specific method for their needs. To increase the complexity in selecting the individual models most suitable for a specific dataset and intended application, the criteria and recommendations that inform model choice are often incomplete and are dispersed throughout the literature (Elith & Graham, 2009). Therefore, the use of ensemble modelling methods may still produce the most accurate and reliable models given the limitations of individual researchers.

4 Conclusions

While species distribution models can provide a reliable, cost effective way of estimating the possible distribution of a species outside of its current range, consideration needs to be given that the modelling method chosen suits the dataset and intended application (Elith & Leathwick, 2009; Guisan & Zimmermann, 2000). The use of SDMs to guide conservation actions presents a range of limitations, however many of these limitations are addressed with the use of ensemble modelling methods (Thuiller et al., 2009). Resent research has shown that ensemble models may not be more reliable than specific SDM methods (Hao et al., 2019; Hao et al., 2020), but given limited researcher understanding and selection criteria information in the literature, ensemble models still provide a reasonably accurate and reliable method of estimating species distribution (Elith & Graham, 2009). Additionally, with an increase in the development of machine learning techniques comes new and improved methods of modelling the distributions of species (Elith & Graham, 2009), bringing with them a new understanding of the factors that drive the distribution of both threatened and pest species alike.

5 References

William T Bean, Robert Stafford, and Justin S Brashares. 2012. The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. *Ecography, 35*(3): 250-258.

Jane Elith, and Catherine H Graham. 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography, 32*(1): 66-77.

Jane Elith, Catherine H. Graham, Robert P. Anderson, Miroslav Dudík, Simon Ferrier, Antoine Guisan, . . . Anthony Lehmann. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography, 29*(2): 129-151.

Jane Elith, and John R Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual review of ecology, evolution, and systematics, 40*: 677-697.

Antoine Guisan, and Niklaus E Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological modelling, 135*(2-3): 147-186.

Tianxiao Hao, Jane Elith, Gurutzeta Guillera - Arroita, and José J Lahoz - Monfort. 2019. A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. *Diversity and Distributions, 25*(5): 839-852.

Tianxiao Hao, Jane Elith, José J Lahoz - Monfort, and Gurutzeta Guillera - Arroita. 2020. Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography*.

Michael Kearney, and Warren Porter. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology letters, 12*(4): 334-350.

A Lomba, L Pellissier, C Randin, J Vicente, F Moreira, J Honrado, and A Guisan. 2010. Overcoming the rare species modelling paradox: a novel hierarchical framework applied to an Iberian endemic plant. *Biological conservation, 143*(11): 2647-2657.

Jana M McPherson, Walter Jetz, and David J Rogers. 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *Journal of applied ecology, 41*(5): 811-823.

Richard G Pearson. 2007. Species' distribution modelling for conservation educators and practitioners. *Synthesis. American Museum of Natural History, 50*: 54-89.

David RB Stockwell, and A Townsend Peterson. 2002. Effects of sample size on accuracy of species distribution models. *Ecological modelling, 148*(1): 1-13.

Robert W Sutherst. 2014. Pest species distribution modelling: origins and lessons from history. *Biological Invasions, 16*(2): 239-256.

Wilfried Thuiller, Bruno Lafourcade, Robin Engler, and Miguel B Araújo. 2009. BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography, 32*(3): 369-373.

Alexander von Humboldt, and Aimé Bonpland. (1807). *Essai sur la géographie des plantes*.