HOW WELL DO SPECIES DISTRIBUTION MODELS PREDICT OCCURRENCES IN EXOTIC RANGES?

Dat Nguyen
Department of Biology
McGill University
Montreal, Quebec, Canada

January 2021

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

TABLE OF CONTENTS

GENERAL ABSTRACT

Species distribution models (SDMs) are one of the most popular numerical tools in ecology, and have been used predictively in biological invasion management. However, the degree to which SDMs are predictive under new contexts – their transferability – is largely unknown with the majority of studies being restricted to few species with equivocal results. This thesis focused on the evaluation of the spatial transferability of SDMs applied to biological invasions, and what factors may contribute to their predictive success. I examined the discriminatory ability of SDMs fitted using two popular modelling approaches for 648 species within their native ranges, and extrapolated to 1918 invaded ranges. I assessed various general factors of transferability, as well as factors related specifically to biological invasions. Despite strong model performance within the native range, more than half of the SDMs had poor performance when validated using their exotic range distributions. Model performance varied between taxonomic classes and invaded continents, and increased with the performance of the model and the environmental breadth of the species in the native range, as well as the number of occurrences and year of first record in the exotic range. Models performed worse with greater environmental breadth in the exotic range, gross domestic product in the native range, geographic distance between ranges and when extrapolating to islands. While poor transferability was observed on average, SDM performances differed depending on characteristics of the species, ranges and models themselves. Some factors were contrary to expectations, where SDMs performed better with higher habitat generalism and for more recent invasions. The latter is counter to assumptions of equilibrium, and suggests that SDMs may be useful for management early in the invasions process. This thesis contributes to the growing literature examining SDM transferability, and provides further insights on their potential uses as a management tool for biological invasions.

RESUMÉ

Les modèles de répartition des espèces (species distribution models; SDMs) sont l'un des outils numériques les plus populaires en écologie et sont régulièrement utilisés de façon prédictive pour la gestion d'invasions biologiques. Cependant, le niveau de prévisibilité des SDMs dans de nouveaux contextes — c'est-à-dire leur transférabilité — demeure largement inconnu, la majorité des études étant limitées à quelques espèces avec des résultats équivoques. Cette thèse se concentre sur l'évaluation de la transférabilité spatiale des SDMs appliqués aux invasions biologiques et les facteurs qui contribuent à leur prévisibilité. J'ai examiné la capacité discriminatoire des SDMs en ajustant deux approches de modélisation populaires pour 648 espèces dans leurs aires de répartition indigènes afin d'extrapoler les modèles aux 1918 aires de répartition envahies. J'ai évalué divers facteurs de transférabilité généraux, ainsi que des facteurs liés spécifiquement aux invasions biologiques. Malgré de bonnes performances envers leur application sur les aires de répartition indigènes, la plupart des SDMs ont offert des performances médiocres lors de leur validation sur les aires de répartition exotiques. La performance des modèles en répartition exotique a varié en fonction des classes taxonomiques utilisées et des continents envahis. Celle-ci a augmenté avec la performance du modèle et l'étendue environnementale des espèces dans leur aire de répartition indigène, ainsi qu'avec le nombre d'occurrences et l'année de la première observation dans les aires de répartition exotiques. La performance des modèles extrapolés a diminué avec la présence de plus grandes étendues environnementales dans l'aire de répartition exotique, de plus gros produits intérieurs bruts dans l'aire de répartition indigène, et de plus longues distances géographiques entre les aires de répartition, ainsi qu'avec leur application aux îles exotiques. Bien qu'une faible transférabilité ait été observée en moyenne, la performance des SDMs différait selon les

caractéristiques des espèces, les aires de répartition impliquées et les modèles eux-mêmes. Certains résultats vont à l’encontre de nos attentes ; notamment, les SDMs extrapolés ont offert de meilleures performances dans les habitats plus généralisés et pour les invasions plus récentes. Ce dernier résultat n’appuie pas les hypothèses de l’équilibre existantes et suggère que les SDM peuvent être utiles pour la gestion dès le début du processus d’invasion. Cette thèse contribue à la littérature de plus en plus vaste examinant la transférabilité des SDMs et fournit des aperçus supplémentaires sur leur potentiel comme outil de gestion des invasions biologiques.

ACKNOWLEDGEMENTS

I would first like to thank my supervisor Professor Brian Leung for his guidance throughout both my undergraduate and graduate studies. This project would not have been possible without Brian's support and commitment to his students.

I am grateful to the members of my supervisory committee, Professor Virginie Millien (Redpath Museum, McGill University) and Professor Laura Pollock (Department of Biology, McGill University) for providing their suggestions and insight during our meetings.

I would like to thank the following institutions for their financial support of this project: Natural Sciences and Engineering Research Council (NSERC), McGill University, and the McGill Department of Biology.

I thank all current and past members of the Leung lab (Abbie Gail Jones, Andrew Sellers, Charlotte Steeves, Emma Hudgins, Jorge Morales, Henry Morgane, Kasia Johnson, Lidia Della Venezia, Natalie Richards, Sarah Chamberland-Fontaine, Shriram Varadarajan) for their feedback, advice and support throughout my time at the lab. This thesis would not have been possible without my fellow lab members fostering a productive and supportive work environment.

Finally, I would like to thank my family and friends for their love and support throughout my project. A special thank you to Emily Horvat and Abbie Gail Jones for their help with the French translations of the Abstract. This thesis would not have been possible without their inspiration and emotional support.

CONTRIBUTION OF AUTHORS

All stages of the thesis – project conceptualization, data collection, modelling and analyses – were developed primarily by the candidate Dat Nguyen with guidance from the supervisor Professor Brian Leung (McGill University). Additional insights and suggestions were provided by the candidate's supervisory committee: Professor Virginie Millien (McGill University) and Professor Laura Pollock (McGill University). The thesis was written by the candidate, with editorial contributions made by Brian Leung.

LIST OF TABLES

Table 1. Full list of variables and acronyms used to predict the transferability of SDMs, with a description of how the factors were generated. Rationale for why each factor may influence SDM transferability is provided for each factor, with sources. Variables marked with an asterisk (*) denote factors generated for both the native and exotic ranges.

Table 2. Estimated model parameters, with standard error, *t*-value and *P*-values of the fixed effects for the linear mixed model fitted using the full dataset for GAMs and MaxEnt with taxonomic class as a categorical predictor.

Table 3. Mean performance of the species distribution models divided between taxonomic groups, measured as area under the receiver operating characteristic curve (AUC), in both the native (fitting) range and the extrapolated (validation) range.

SUPPLEMENTARY TABLES

Table S1. Description of the data sources used in the analyses, for both the species distribution modelling and the predictors of transferability.

Table S2. List of the 648 species included in the analysis along with their taxonomic class and the total number of occurrences classified in their native and exotic ranges.

Table S3. Full list of GBIF DOIs for downloading occurrence records used to sample target-group background sites, with classes and countries as ISO3 corresponding to each DOI link, semi-colon separated.

Table S4. Pairwise correlation table between all continuous and binary predictors included in the transferability model for GAMs.

Table S5. Summary table for models calculated using species mobility (SESS) as a binary predictor of SDM performance, in place of taxonomic class (CLASS). Parameter estimates, standard error, t-statistic and *P*-values are provided.

Table S6. Predictor estimates for the transferability model with GAMs and MaxEnt, including taxonomic class and continent as categorical predictors. Mean and standard deviation are provided for the scaled factors. Predictors marked with an asterisk (*) were log-transformed and predictors marked with two asterisks (**) were logit-transformed before scaling.

Table S7. Mean predictive performance of GAM and MaxEnt species distribution models for each country in the study, with standard deviation and number of invading species.

LIST OF FIGURES

Figure 1. Flow chart outlining the analyses performed in the study.

Figure 2. Density histogram of the performance of GAMs (a) and MaxEnt (b) measured as area under the receiver operating characteristic curve (AUC) when predicting the native range, or fitting dataset (blue), using 10-fold cross-validation and when extrapolating to the exotic range, or validation dataset (red).

Figure 3. Dot-whisker plot of the parameter estimates for the binary and continuous predictors included in the linear mixed model for GAM (black) and MaxEnt (orange), with taxonomic class

and continent as categorical predictors. Whiskers represent the 95 percent confidence intervals of the parameter estimates.

Figure 4. Map of the mean discriminatory performance (area under the receiver operating characteristic curve; AUC) of the extrapolated models using (a) GAM and (b) MaxEnt. Regions marked with an asterisk (*) represent areas that were invaded by fewer than three species.

SUPPLEMENTARY FIGURES

Figure S1. Change in area under the receiver operating characteristic curve (AUC) and Continuous Boyce Index (CBI) values at decreasing proportion of sampled occurrences for the simulated species-environment. Points represent the mean values at the given proportion of sampled presences, with the red line representing the mean value at 95 percent sampled presences and error bars representing the standard deviation at each fraction.

Figure S2. Comparison of native range model performance using 10-fold random cross-validation and 5-fold spatial-block cross-validation, measured as the mean area under the receiver operating characteristic curve (AUC).

Figure S3. Example of PCA-based environmental breadth and coverage as a predictor for the prickly pear (*Opuntia ficus-indica*). The grey points represent the 10,000 global reference site points used to generate the PCA environment space for all species. The blue and black points represent the occurrences and sampled background sites in geographic and environmental space, respectively.

GENERAL INTRODUCTION

Rapid global change currently threatens biodiversity, which has led to new challenges to understand and manage these drivers of extinction. The current biodiversity extinction crisis has been attributed to various global drivers, including biological invasions (Mack et al., 2000), climate change and land use change (Millennium Ecosystem Assessment, 2005). There has been a growing need to develop tools to better understand the ramifications of global change on species, and predict future responses for conservation. This thesis focuses on biological invasions, and on estimating the distributions of species in their exotic ranges.

Biological invasions

Biological invasions, in which species are introduced, establish and spread throughout a new region, have led to severe ecological damages and extirpations of vulnerable populations (Doherty et al., 2016; Mack et al., 2000; Vitousek et al., 1996), as well as severe economic impacts with billions of dollars lost annually in the United States alone (Pimentel et al., 2005). Modern invasions have been primarily driven by the movement of humans, particularly through global trade between biogeographical regions (Meyerson & Mooney, 2007). In particular, the rise in international trade has led to increases in the frequency of introduction events as well as the number of introduced individuals (propagule pressure; Lockwood et al., 2005), which is a primary factor determining invasions. The rate of new invasions has been increasing over time globally where a lack of saturation has been observed (Seebens et al., 2017), demonstrating an increasingly urgent concern for conservation.

The management of biological invasions may occur at any stage in the invasion process. Prevention is an important component of invasive species management, and therefore understanding the likelihood of species establishment and potential subsequent spread in a new region is crucial to estimate the potential risk posed by each species and to prioritize management efforts (Leung et al., 2012). Once the species has established within a region, identifying the most suitable areas may help prioritize efforts to control or contain the species (Gormley et al., 2011). Early detection and rapid response are necessary management strategies, before invading species become widespread and too challenging to control or contain (Lodge et al., 2006). To this effect, species distribution models (SDMs) have become one of the most promising tools to predict potentially invaded or currently invaded ranges.

Species distribution models

The growth of large datasets has given rise to diverse ecological models that may be used to inform management decisions. Species distribution models (SDMs) are an example of these numerical tools, relating species occurrence information to the environmental factors of their distribution through a quantitative or rules-based framework (Franklin, 2010). Many other names have been used to describe the species distribution modelling process, including ecological niche models (ENMs), habitat models, climate envelopes, and range mapping (Elith & Leathwick, 2009).

The ecological niche of a species is a vital concept in ecology, and is defined as the set of biotic and abiotic conditions required for a species to persist, where the fundamental niche of a species consists of all the suitable conditions for their survival, and the realized niche consists of the ecological space that the species actually occupies (Hutchinson, 1957). While SDMs may attempt to capture the ecological requirements of species, the fundamental niche of a species is unknown in practice (Phillips et al., 2006). SDMs measure the observed distributions, which may be constrained by dispersal limitations or historical factors, or other factors and thus reflect the realized niche (Lobo et al., 2010).

SDMs may be implemented in a number of different ways, such as the response variable of the model, the modelling approach itself and the predictors included. The most common response variables used in SDMs for single species are presence-only (in which the absences of the species are not known; Yates et al., 2018). While presence-absence and abundance data contain the more information content and may yield better model performance, the use of presence-only data has become more common due to their wide availability through large databases (e.g. GBIF, 2020) and the practical challenges associated with obtaining rigorously sampled presence-absence or abundance data (Barbet-Massin et al., 2012).

Different approaches have emerged to best employ the use of presence-only data. Few modelling procedures exclusively use presences (e.g. BIOCLIM; Busby, 1991), where most approaches employ the use of pseudo-absences, or background information in place of true absences (Barbet-Massin et al., 2012). Background sites have been interpreted as a sample of available environments in the study site for a given species, or as an indicator of their unsuitable conditions (Capinha et al., 2011). One consequence of using SDMs fitted using presence-only data is that they can only provide relative likelihoods of occurrences, as background sites do not reflect true absences (which is unknown), and the number of sampled background sites influence the absolute probabilities of the model predictions (Pearce & Boyce, 2006). However, the relative suitability values from presence-only SDMs can yield important information, allowing for “discrimination” between invaded and background sites (rather than “calibration” related to absolute probabilities of occurrence; Pearce & Boyce, 2006). Indeed, presence-only SDMS may discriminate between high and low risk areas which may be used to inform invasion management priorities (e.g. Crafton, 2014).

SDMs may use various predictors to describe species distributions. SDMs may be implemented using indirect, correlative predictors (which comprise abiotic or biotic predictors), or as mechanistic models based on experimental data to link physiological responses directly to their environment (Evans et al., 2015; Yates et al., 2018). While mechanistic models may provide direct linkages between species and their environment (Evans et al., 2015), the required experimental data is often resource intensive, making such models currently impractical (Elith and Leathwick, 2009; Yates et al., 2018), particularly for biological invasions where rapid assessment is often needed. Consequently, correlative models have become the most common form of SDM, particularly using widely available, high resolution climatic data (e.g. Fick & Hijmans, 2017) that may be easily implemented as predictors of species distributions.

SDMs have been applied for invasive species across several taxa, including marine (e.g. Crafton, 2014) and terrestrial (e.g. Lozier & Mills, 2011) ecosystems. SDMs may be used, for example, to identify areas of high risk for species in transit where they may establish and spread (Leung et al., 2012). Where species have established, SDMs are particularly useful to identify where the species may be located or may spread, to inform management priorities for control (e.g. Gormley et al., 2011). The application of SDMs may also be used to link global drivers of biodiversity loss, for example by predicting the shift in invasive species suitability over time due to climate change (e.g. Ahmad et al., 2019). Indeed, the benefits of SDMs to better inform conservation efforts for biological invasions are clear, where risk assessments that allow the prioritization of species or sites may be made through these spatial predictions (Elith, 2017).

Broadly, SDMs have become widely popular as an approach to answer large-scale ecological questions. SDMs can be used for explanation, in order to gain insight on the causal drivers of species distributions and their niche requirements (e.g. Graham et al., 2004; Heegaard, 1997; Leathwick & Austin, 2001). Over time, the primary application of distribution modelling has shifted to prediction as a result of the increasingly large quantity of available data and the development of more advanced models (Elith & Leathwick, 2009).

SDMs transferability for biological invasions

In the context of biological invasions, SDMs fitted on the native range of a species have been used to predict their suitability or distribution in new ranges (e.g. Gama et al., 2016; Padalia et al., 2014). Unlike SDMs used for explanation, these predictive models are extrapolated to exotic ranges, where many factors may differ, and the reliability of the model may not hold. Therefore, understanding the degree to which SDMs may be predictive outside of the data used to fit the model, termed transferability, is necessary if SDMs are to be useful for forecasting the distributions of biological invasions.

Despite the widespread use of SDMs, recent publications have criticized the use of SDMs when making predictions in new environments, either temporally or spatially, suggesting that these models may fail outside of the data used to calibrate the models. The majority of SDMs are not rigorously examined before being applied where predictive models are typically assessed using the range used to generate the model itself rather than on an independent dataset and consequently may not reflect its actual predictive ability (Vanreusel et al., 2007).

Recent reviews have outlined the challenges of developing ecological models for prediction, highlighting various factors that may influence transferability (Sequeira et al., 2018; Werkowska et al., 2017; Yates et al., 2018). Transferability may be influenced by the quality of the data, including low sample sizes, where sampling biases and low positional accuracy may consequently lead to poor SDM predictions (Yates et al., 2018). The choice of modelling approach influences the complexity of the models themselves and consequently their transferability, where overfitted models may fit to noise in the data, and underfitted models may fail to effectively capture complex ecological relationship (Werkowska et al., 2017). Identifying relevant predictors (and the number of predictors) may also influence transferability, in order to effectively describe a species distribution (Petitpierre et al., 2017). Whether specific taxa may be more or less transferable has been of interest (Yates et al., 2018), which may contain traits that make them more or less transferable (Urban et al., 2016). For example, ecological models have been shown to perform poorly for habitat generalist species (Brotons et al., 2004; Brotons et al., 2007; Zhang et al., 2015), as these species may not be restricted by the environmental predictors, as well as fail to differentiate between the presence and background environment. Shifting species distributions over time may violate the assumption that species are in equilibrium with their environment (Araújo & Peterson, 2012; Guisan & Thuiller, 2005). This assumption is often violated in the early stages of biological invasion, which may lead to the miscalibration of SDMs when characterizing the species-environment relationship (Václavík & Meentemeyer, 2011). Population density may influence the spread and distribution of species, and consequently model transferability to new ranges (Yates et al., 2018). While propagule pressure is a major predictor of invasive species establishment, it may also influence the geographic spread of established species, increase the genetic variation of the population and allow for multiple introductions in different locations (Leung et al., 2012; Lockwood et al., 2005). Therefore, propagule pressure and its surrogate measures such as gross domestic product (GDP; Sardain et al., 2019) may correlate with human-mediated spread and the degree to which species occupy all of the suitable environments within invaded ranges.

One of the primary challenges for developing models for transferability is the potential failure of SDMs when extrapolating to non-analogous conditions (Sequeira et al., 2018; Yates et al., 2018). SDMs fitted on a truncated curve may fail to predict outside of the conditions used to fit the model, for which species responses are unknown (Thuiller et al., 2004). Alternatively, SDMs may fail in cases where the realized niche has shifted (e.g. release from antagonizing biotic interactions; Tingley et al., 2014), leading to an apparent change in their species-environment relationship. Yet, the majority of SDMs do not include important biological factors such as physiological and evolutionary traits, species interactions and dispersal, instead only considering the abiotic environment that the species occupies (Urban et al., 2016). Niche shifts have been observed across a variety of taxa for biological invasions, including plants (Broennimann et al., 2007; Early & Sax, 2014), insects (Hill et al., 2017), marine fish (Parravicini et al., 2015), birds (Cardador & Blackburn, 2020; Strubbe et al., 2012), reptiles and amphibians (Li et al., 2014; Tingley et al., 2014).

While there has been an increasing number of studies that have evaluated the spatial transferability of SDMs, the majority of these assessments have been restricted to few species or taxonomic groups and with varying methodologies leading to equivocal results. For example, Petitpierre et al. found that SDMs could reliably predict the exotic ranges for plants, noting that few species in their invaded range existed outside of their native climatic niche (2012). Other studies have reported the opposite trend, finding that species occurred outside of the climatic conditions recorded in their native range (e.g. Beaumont et al., 2009; Goncalves et al., 2014). This lack of consensus has demonstrated a general limited understanding of model transferability, which is crucial for the application of SDMs for prediction. Therefore, a broad systematic assessment of SDM transferability, with the identification of factors that drive model success, would be timely and useful.

This thesis seeks to evaluate the transferability of SDMs when extrapolated across spatial landscapes, and identify the potential factors associated with model failure. I constructed SDMs for 648 terrestrial species across 8 taxonomic classes using occurrence records from their native range and extrapolated these models to 1918 characterized invaded ranges as an independent form of validation to assess the loss in transferability. Using global occurrence data to provide the most extensive assessment of SDM transferability for invasive species to date, my goal was to examine under what conditions SDMs may succeed based on hypotheses surrounding transferability in general, as well as factors related specifically to biological invasions.

REFERENCES

Ahmad, R., Khuroo, A. A., Charles, B., Hamid, M., Rashid, I., & Aravind, N. A. (2019). Global distribution modelling, invasion risk assessment and niche dynamics of Leucanthemum vulgare (Ox-eye Daisy) under climate change. *Scientific reports*, 9(1), 1-15.

Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, 93(7), 1527-1539.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many?. *Methods in ecology and evolution*, 3(2), 327-338.

Beaumont, L. J., Gallagher, R. V., Thuiller, W., Downey, P. O., Leishman, M. R., & Hughes, L. (2009). Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions*, 15(3), 409-420.

Broennimann, O., Treier, U. A., Müller-Schärer, H., Thuiller, W., Peterson, A. T., & Guisan, A. (2007). Evidence of climatic niche shift during biological invasion. *Ecology letters*, 10(8), 701-709.

Brotons, L., Thuiller, W., Araújo, M. B., & Hirzel, A. H. (2004). Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, 27(4), 437-448.

Brotons, L., Herrando, S., & Pla, M. (2007). Updating bird species distribution at large spatial scales: applications of habitat modelling to data from long-term monitoring programs. *Diversity and distributions*, 13(3), 276-288.

Busby, J. (1991). BIOCLIM-a bioclimate analysis and prediction system. *Plant protection quarterly*, 61, 8-9.

Capinha, C., Leung, B., & Anastácio, P. (2011). Predicting worldwide invasiveness for four major problematic decapods: an evaluation of using different calibration sets. *Ecography*, 34(3), 448-459.

Cardador, L., & Blackburn, T. M. (2020). A global assessment of human influence on niche shifts and risk predictions of bird invasions. *Global Ecology and Biogeography*, 29(11), 1956-1966.

Crafton, R. E. (2014). Modeling invasion risk for coastal marine species utilizing environmental and transport vector data. *Hydrobiologia*, 746(1), 349-362.

Doherty, T. S., Glen, A. S., Nimmo, D. G., Ritchie, E. G., & Dickman, C. R. (2016). Invasive predators and global biodiversity loss. *Proceedings of the National Academy of Sciences*, 113(40), 11261-11265.

Early, R., & Sax, D. F. (2014). Climatic niche shifts between species' native and naturalized ranges raise concern for ecological forecasts during invasions and climate change. *Global ecology and biogeography*, 23(12), 1356-1365.

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

Elith, J. (2017). Predicting distributions of invasive species. *Invasive species: Risk assessment and management*, 10(9781139019606.006).

Evans, T. G., Diamond, S. E., & Kelly, M. W. (2015). Mechanistic species distribution modelling as a link between physiology and conservation. *Conservation physiology*, 3(1), cov056.

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(12), 4302-4315.

Franklin, J. (2010). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press.

Gama, M., Crespo, D., Dolbeth, M., & Anastácio, P. (2016). Predicting global habitat suitability for Corbicula fluminea using species distribution models: The importance of different environmental datasets. *Ecological Modelling*, 319, 163-169.

GBIF (2020). Global Biodiversity Information Facility. Available from https://www.gbif.org/

Goncalves, E., Herrera, I., Duarte, M., Bustamante, R. O., Lampo, M., Velasquez, G., ... & García-Rangel, S. (2014). Global invasion of *Lantana camara*: has the climatic niche been conserved across continents?. *PLoS One*, 9(10), e111468.

Gormley, A. M., Forsyth, D. M., Griffioen, P., Lindeman, M., Ramsey, D. S., Scroggie, M. P., & Woodford, L. (2011). Using presence-only and presence–absence data to estimate the current and potential distributions of established invasive species. *Journal of Applied Ecology*, 48(1), 25-34.

Graham, C. H., Ron, S. R., Santos, J. C., Schneider, C. J., & Moritz, C. (2004). Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution*, 58(8), 1781-1793.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology letters*, 8(9), 993-1009.

Heegaard, E. (1997). Ecology of Andreaea in western Norway. *Journal of Bryology*, 19(3), 527-636.

Hill, M. P., Gallardo, B., & Terblanche, J. S. (2017). A global assessment of climatic niche shifts and human influence in insect invasions. *Global ecology and biogeography*, 26(6), 679-689.

Hutchinson, G.E. (1957) Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22, 415-427.

Leathwick, J. R., & Austin, M. P. (2001). Competitive interactions between tree species in New Zealand's old-growth indigenous forests. *Ecology*, 82(9), 2560-2573.

Leung, B., Roura-Pascual, N., Bacher, S., Heikkilä, J., Brotons, L., Burgman, M. A., ... & Sol, D. (2012). TEASIng apart alien species risk assessments: a framework for best practices. *Ecology Letters*, 15(12), 1475-1493.

Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, 33(1), 103-114.

Lockwood, J. L., Cassey, P., & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends in ecology & evolution*, 20(5), 223-228.

Lodge, D. M., Williams, S., MacIsaac, H. J., Hayes, K. R., Leung, B., Reichard, S., ... & Carlton, J. T. (2006). Biological invasions: recommendations for US policy and management. *Ecological applications*, 16(6), 2035-2054.

Lozier, J. D., & Mills, N. J. (2011). Predicting the potential invasive range of light brown apple moth (*Epiphyas postvittana*) using biologically informed and correlative species distribution models. *Biological Invasions*, 13(10), 2409.

Li, Y., Liu, X., Li, X., Petitpierre, B., & Guisan, A. (2014). Residence time, expansion toward the equator in the invaded range and native range size matter to climatic niche shifts in non-native species. *Global Ecology and Biogeography*, 23(10), 1094-1104.

Mack, R. N., Simberloff, D., Mark Lonsdale, W., Evans, H., Clout, M., & Bazzaz, F. A. (2000). Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological applications*, 10(3), 689-710.

Meyerson, L. A., & Mooney, H. A. (2007). Invasive alien species in an era of globalization. *Frontiers in Ecology and the Environment*, 5(4), 199-208.

Millennium Ecosystem Assessment (2005). *Ecosystems and human well-being: biodiversity synthesis*. World Resources Institute, Washington, DC.

Padalia, H., Srivastava, V., & Kushwaha, S. P. S. (2014). Modeling potential invasion range of alien invasive species, Hyptis suaveolens (L.) Poit. in India: Comparison of MaxEnt and GARP. *Ecological informatics*, 22, 36-43.

Parravicini, V., Azzurro, E., Kulbicki, M., & Belmaker, J. (2015). Niche shift can impair the ability to predict invasion risk in the marine realm: an illustration using Mediterranean fish invaders. *Ecology Letters*, 18(3), 246-253.

Pearce, J. L., & Boyce, M. S. (2006). Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, 43(3), 405-412.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C., & Guisan, A. (2012). Climatic niche shifts are rare among terrestrial plant invaders. *Science*, 335(6074), 1344-1348.

Petitpierre, B., Broennimann, O., Kueffer, C., Daehler, C., & Guisan, A. (2017). Selecting predictors to maximize the transferability of species distribution models: lessons from cross-continental plant invasions. *Global Ecology and Biogeography*, 26(3), 275-287.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological modelling*, 190(3-4), 231-259.

Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological economics*, 52(3), 273-288.

Sardain, A., Sardain, E., & Leung, B. (2019). Global forecasts of shipping traffic and biological invasions to 2050. *Nature Sustainability*, 2(4), 274-282.

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., ... & Bacher, S. (2017). No saturation in the accumulation of alien species worldwide. *Nature communications*, 8(1), 1-9.

Sequeira, A. M., Bouchet, P. J., Yates, K. L., Mengersen, K., & Caley, M. J. (2018). Transferring biodiversity models for conservation: opportunities and challenges. *Methods in Ecology and Evolution*, 9(5), 1250-1264.

Strubbe, D., Broennimann, O., Chiron, F., & Matthysen, E. (2012). Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography*, 22(8), 962-970.

Thuiller, W., Brotons, L., Araújo, M. B., & Lavorel, S. (2004). Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, 27(2), 165-172.

Tingley, R., Vallinoto, M., Sequeira, F., & Kearney, M. R. (2014). Realized niche shift during a global biological invasion. *Proceedings of the National Academy of Sciences*, 111(28), 10233-10238.

Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J. B., Pe'er, G., Singer, A., ... & Gonzalez, A. (2016). Improving the forecast for biodiversity under climate change. *Science*, 353(6304), aad8466.

Václavík, T., & Meentemeyer, R. K. (2011). Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions*, 18(1), 73-83.

Vanreusel, W., Maes, D., & Van Dyck, H. (2007). Transferability of species distribution models: a functional habitat approach for two regionally threatened butterflies. *Conservation biology*, 21(1), 201-212.

Vitousek, P., D'Antonio, C., Loope, L., & Westbrooks, R. (1996). Biological invasions as global environmental change. *American scientist*, 84, 468.

Werkowska, W., Márquez, A. L., Real, R., & Acevedo, P. (2017). A practical overview of transferability in species distribution modeling. *Environmental reviews*, 25(1), 127-133.

Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., ... & Dormann, C. F. (2018). Outstanding challenges in the transferability of ecological models. Trends in ecology & evolution, 33(10), 790-802.

Zhang, L., Liu, S., Sun, P., Wang, T., Wang, G., Zhang, X., & Wang, L. (2015). Consensus forecasting of species distributions: The effects of niche model performance and niche properties. *PLoS One*, 10(3), e0120056.

CHAPTER ONE

ABSTRACT

Aim

Species distribution models (SDMs) are widely used predictive tools to forecast potential biological invasions. However, the reliability of SDMs extrapolated to exotic ranges remains understudied, with most analyses restricted to few species with equivocal results. We examine the spatial transferability of SDMs for 648 non-indigenous species extrapolated across new ranges, and identify what factors may lead to predictive failure.

Location

Global.

Time period

Current.

Major taxa studied

648 terrestrial species; 8 taxonomic classes.

Methods

We performed a large-scale assessment of the transferability of SDMs using two modelling approaches: generalized additive models (GAMs) and MaxEnt. We fitted SDMs on the native ranges of species and extrapolated them to 1918 exotic ranges. We examined the effect of several general factors, and factors related to biological invasions, on spatial transferability.

Results

Despite both modelling approaches performing well when predicting the fitting range of the species (GAM mean AUC = 0.809; standard deviation (s) = 0.076; MaxEnt mean AUC = 0.830; $s = 0.070$), we observed low spatial transferability on average when extrapolating to their invaded ranges (GAM mean AUC = 0.686; $s = 0.124$; MaxEnt mean AUC = 0.689; $s = 0.122$). Transferability differed between taxonomic classes and invaded continents, and was positively influenced by the performance of the model and environmental breadth in the native range, number of occurrences in the exotic range, and the year of first record. Models performed worse with greater environmental breadth in the exotic range, gross domestic product of the native range, geographic distance between ranges and when extrapolating to islands.

Main conclusions

Our study showed poor spatial transferability on average, with SDM performance differing based on characteristics of the data, species and ranges. Contrary to expectations, specialized species (in native ranges) had the lowest predictability in exotic ranges, and more recent invasions showed stronger discriminatory ability. The latter suggests equilibrium conditions may not favor higher predictability in exotic ranges, and SDMs could be useful for management early in an invasion. We discuss the possible mechanisms of these findings.

Keywords: *GBIF; global; invasive species; presence-only data; species distribution model; transferability*

INTRODUCTION

Developing models to predict species distributions within exotic ranges has become increasingly important for the management of biological invasions (Elith, 2017) due to the severe effects of invaders on biodiversity loss (Bellard et al., 2016) and economic damages (Bradshaw et al., 2016). Given the increasing availability of large databases and sophisticated algorithms (Elith & Leathwick, 2009), species distribution models (SDMs) have become the tool of choice for prediction (e.g. Lozier & Mills, 2011). SDMs are statistical tools that relate species occurrence data to their environment and other spatial characteristics (Franklin, 2009). SDMs have been shown to perform well when characterizing the distribution of the species within its fitting range when utilizing informative predictors and rigorously sampled survey data (Elith & Leathwick, 2009). However, the reliability of SDMs when extrapolating to novel contexts must be viewed with caution (Yates et al., 2018).

The challenges of developing of SDMs as predictive tools have been summarised in recent publications (Sequeira et al., 2018; Werkowska et al., 2017; Yates et al., 2018), noting that the transferability of SDMs across spatial landscapes is largely unknown. Most direct analyses of transferability to exotic ranges have been restricted to single or few species or taxonomic groups (e.g. Beaumont et al., 2009; Goncalves et al., 2014; Peterson et al., 2003; Sofaer et al., 2018). Moreover, there has been a lack of consensus on whether such models are transferable despite the popularity of ecological models as predictive tools (e.g. Beaumont et al., 2009; Goncalves et al., 2014; Morán-Ordóñez et al., 2017; Petitpierre et al., 2012). Thus, a broader systematic analysis is needed, with a focus on SDMs built on native ranges and extrapolated to exotic ranges, as well as the potential factors affecting their transferability.

Several factors may influence model transferability generally, across time and space. One proposed issue is whether SDMs may be more or less transferable depending on the taxonomic group of the species (Yates et al., 2018), given relevant differences in biological traits that influence species distributions (Urban et al., 2016). For example, the movement and behaviour of species may affect whether occurrences may be accurately recorded, as well as their ability to spread to suitable habitats. Habitat-generalist species may lead to poorer transferability as they may not be constrained by the environmental factors considered, or comprise multiple subpopulations that differ environmentally (Brotons et al., 2007), affecting whether the fitted SDMs can discriminate between presences and background data (Brotons et al., 2004; Zhang et al., 2015). Other general factors have been highlighted, such as data quality (Yates et al., 2018), model complexity (Werkowska et al., 2017), statistical approach (Yates et al., 2018), stochasticity and small sample sizes (Yates et al., 2018), and the choices of predictors (Petitpierre et al., 2017).

Beyond these general issues, some factors may be particularly relevant for biological invasions. A key element of invasions is that species may still be spreading within their introduced range (Václavík & Meentemeyer, 2011), violating the underlying assumption of SDMs that species are in equilibrium with their environment (Araújo & Peterson, 2012; Guisan & Thuiller, 2005). In turn, spread is influenced by propagule pressure (the number of individuals introduced and/or rate of introduction events; Lockwood et al., 2005), where higher propagule pressure leads to a greater opportunity to encounter suitable habitats (i.e. they are less limited by "exposure"). Further, lack of exposure should be more important earlier in the invasion process, and distributions should become increasingly driven by habitat suitability as a species continues to spread (Leung et al., 2012; Runquist et al., 2019). Thus, we hypothesized that species with higher propagule pressure and those that had invaded earlier would show stronger transferability.

Previous works have identified non-analogous environments as a key factor affecting transferability, where models that are fitted on truncated curves may fail to predict species responses to new conditions (Yates et al., 2018). This is particularly relevant as invasive species are introduced to disjoint ranges by definition. Thus, we predicted a positive relationship between spatial transferability and environmental similarity between native and exotic ranges. Likewise, ranges may differ over longer geographic distances, which could correlate to other unmeasured environmental, biotic or human factors. Finally, we also explored the effect of islands on transferability. Islands differ from mainland areas in terms of their higher levels of endemism, lower species richness and restricted land areas (Kier et al., 2009), suggesting non-analogous conditions. Moreover, impacts often differ between island and mainland invasions (Spatz et al., 2017), suggesting ecological differences at play. Thus, we predicted lower transferability when predicting island invasions.

The objectives of this study were (1) to evaluate the capacity of SDMs to predict species distribution when extrapolated to exotic ranges (i.e., disjointed and outside of the range used for calibration), and (2) to identify factors that may influence model transferability. To do so, we analyzed the transferability of SDMs constructed on the native range of 648 terrestrial species across 8 taxonomic classes and extrapolated to their exotic ranges, providing the largest systematic analysis of native to exotic range transferability to date.

METHODS

To assess the transferability of SDMs, we fitted models for species on their native range and extrapolated them to their exotic ranges (see Figure 1 for full flowchart of methodology). Background reference sites were sampled using the ‘target-group background’ approach (TGB, Phillips et al., 2009) within each range to account for biases associated with presence-only data. After extrapolating to the exotic range of the species, we assessed the transferability of SDMs in terms of their discriminatory power and examined several potential predictors of model success based on hypotheses associated with the models, species or characteristics of the ranges themselves (Table 1). All analyses were performed in R (R Core Team, 2019). Descriptions of the datasets used in the model fitting, validation and predictors of transferability are provided in Supporting Information Table S1.

Environmental data

We included environmental predictors that are available globally and frequently used in SDM literature. We used the 19 bioclimatic variables from WorldClim version 2 at 2.5-arcmin resolution (approximately 5 km grid size; Fick & Hijmans, 2017). Variables were derived from monthly temperature and precipitation values relating to its mean, variation and extremes (Fick & Hijmans, 2017). All bioclimatic variables were standardized to a mean of zero and standard deviation of one. Prior to model fitting, we accounted for multicollinearity by removing highly correlated variables across all cells globally using a threshold pairwise correlation coefficient value of $|r| > 0.7$. After excluding collinear variables, the reduced set was used in all SDMs: annual mean temperature (bio1), mean diurnal range (mean of monthly maximum and minimum temperatures; bio2), temperature annual range (bio7), annual precipitation (bio12), precipitation of the driest month (bio14), precipitation seasonality (coefficient of variation; bio15) and precipitation of the coldest quarter (bio19).

We included two variables in addition to the seven bioclimatic predictors: elevation and normalized difference vegetation index (NDVI) at 2.5-arcmin and 3-arcmin resolution, respectively. Elevation has been shown to drive many ecological processes affecting species distributions (Wang et al., 2017), while NDVI, which is an index of vegetation cover ("greenness"), has been used as surrogate for habitat quality (Pettorelli et al., 2011) and land cover changes (Lunetta et al., 2006). Both predictors have been widely used in SDMs (Morán-Ordóñez et al., 2017; Wen et al., 2015) and globally available. Elevation data was downloaded from WorldClim (Fick & Hijmans, 2012), and MODIS NDVI data was downloaded through the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC; Acker & Leptoukh, 2007). Neither elevation nor NDVI were highly correlated ($|r| < 0.7$) with the climate predictors.

Species data

We obtained species occurrence records from the Global Biodiversity Information Facility (GBIF, 2020; see Supporting Information Text S1 for the list of DOI links), which were subsequently gridded to the environmental data (at 2.5-arcmin resolution). Grid cells containing multiple occurrences were counted as a single presence. We filtered out records with listed geospatial issues within GBIF and coordinate uncertainty greater than 5 kilometers (the resolution of our environmental data). We applied standardized cleaning using the

‘CoordinateCleaner’ package (Zizka et al., 2019), which removed any records with equal or zero/zero coordinates, found in urban areas, near biodiversity institutions, outside of their listed country, or at the centroids of countries and its subdivisions. Finally, we removed all occurrence records that did not have associated environmental data, or were dated before the year 1970 to match the temporal resolution of WorldClim.

Species occurrences were classified as either native or exotic at the country level using distributional information from the CABI Invasive Species Compendium (CABI, 2020) and the IUCN Global Invasive Species Database (GISD, 2015). We defined “regions” and “ranges” as physical locations and the contiguous geographic extent of the species, respectively. Regions were generated at the sub-national level (state/province/equivalent), using the Global Administrative Areas database (2018). Exotic ranges were defined as the set of contiguous regions (that is, regions that share a land border with each other) containing at least one occurrence point. Therefore, an SDM fitted on the native range for a species may be extrapolated to one or more exotic ranges, each comprising one or more connected regions. We excluded records that were found in countries where the species was considered both native and exotic, or not classified. Ranges (or species) with fewer than 20 occurrences were excluded to avoid overfitting. The species included in the study, along with relevant taxonomic information are presented in Supporting Information Table S2.

Background environment selection

Presence-only occurrence records represent the most common form of distributional data. We compared the environments our occurrence records to 10,000 sampled background sites in both the native and exotic ranges (Barbet-Massin et al., 2012; Capinha et al., 2011). While the number of background points included in the model affects the absolute probabilities of prediction, the likelihood of presences could still be interpreted in a relative manner, termed discrimination (Pearce & Boyce, 2006). We used the 'target-group background' approach to select our background sites (TGB; Phillips et al., 2009) by randomly sampling GBIF records belonging to the same taxonomic class as the species of interest, within the same range and independently of species presences (see Supporting Information Table S3 for the GBIF DOIs). This accounted for observation biases associated with presence-only data by mimicking the sampling approach of the occurrence records. We applied the same data cleaning procedure to our target-group backgrounds sites as the occurrence records. For computational reasons, we downloaded occurrence records for 300 randomly sampled species within class-region combinations that possessed greater than 10,000,000 occurrences in total, which were then sub-sampled to 10,000 points per range. Additionally, we used the maximum number of available occurrences where ranges had fewer than 10,000 available background sites. However, species that did not possess at least 4000 potential background sites were excluded.

Modelling species distributions

We applied two modelling approaches as the framework of the SDMs in this study: generalized additive models (GAMs; Hastie & Tibshirani, 1990) and MaxEnt (Phillips et al., 2006). GAMs were implemented in this study using the 'mgcv' package in R (Wood, 2017). Specifically, for GAMs, non-linear relations (the 'smooth' terms in the model) are defined by segments specified by the number of knots, each of which may have their own polynomial function form.

$$z_{ij} = b_0 + s(x_{1,i}) + s(x_{2,i}) + \ldots + s(x_{m,i}) \quad (1)$$

$$y_{ij} = \frac{1}{1+e^{-z_{ij}}} \quad (2)$$

Where y_{ij} was the binary presence or background at site i and species j, and $x_{1,i} \ldots x_{m,i}$ were the m continuous climatic predictors with smoothing terms s. We used a maximum of 5 knots to allow each smoothing function to remain flexible, but computationally efficient. To prevent overfitting, we allowed terms to be excluded by setting the 'select' argument to true within the 'gam' function. Additionally, we removed variables showing concurvity (the non-linear extension of multicollinearity where the smooth term for one independent variable can be approximated by a linear combination of the other smooth terms in the model; Buja et al., 1989; Morlini, 2006). We applied the 'concurvity' function from the package 'mgcv' (Wood, 2017), and removed predictors with the highest 'worst' case concurvity. We refit the GAM, iterating this procedure until all variables possessed estimated concurvity values less than 0.8.

MaxEnt is a machine learning method designed specifically for presence-background modelling (Phillips et al., 2006; Phillips & Dudík 2008). Similar to GAMs, MaxEnt may be subject to similar issues of overfitting given its flexible modelling procedures. To reduce possible overfitting, we limited the model complexity to only include linear, quadratic and product features (Merow et al., 2014). The SDMs were fitted using 'maxnet' package in R (version 3.4.0; Phillips, 2017), with the default arguments for the rest of the settings.

As the metric of model performance, we used the area under the receiver operating characteristic curve (AUC; Hanley & McNeil, 1982). An AUC value of 1 indicates perfect discrimination, while a value of 0.5 indicates discrimination no better than chance. We compared AUC in the fitted range against AUC in the exotic ranges, to determine transferability, and conversely, loss of discriminatory power. As presence-only models discriminate between presences and background sites rather than absences, the interpretation of AUC is less clear than with presence-absence (Phillips et al., 2006). Nonetheless, we considered AUC values greater than or equal to 0.7 to have useful model performance (Morán-Ordóñez et al., 2017; Swets, 1988). While several authors have criticized the use of AUC due to the equal weighting of omission and commission errors (Jiménez-Valverde, 2012; Lobo et al., 2007), AUC remains one of the most commonly used metrics of discrimination for SDMs, and is suited for the relative behaviour of presence-background models (Phillips et al., 2006). We also considered the Continuous Boyce Index (CBI; Hirzel et al., 2006), but using simulation analyses we found that CBI was sensitive to the proportion of unsampled presences, while AUC was robust (Supporting Information Text S2; Figure S1). Therefore, we only report the AUC in the main text (but provide the CBI results in the Supporting Information Text S2).

Predictors of SDM transferability

We examined several factors based on characteristics of the SDMs and species that may be related to model transferability in general, as well as characteristics related specifically to biological invasions (Table 1). As a general predictor of transferability, we included the discriminatory performance within the native range, based on mean AUC from 10-fold cross-validation. Additionally, we compared the mean AUC generated using 10-fold random cross-validation to spatial-block cross-validation (Supporting Information Text S3). Both techniques were highly correlated with each other and had little overall difference in results, with a small decline in performance with spatial-block cross-validation (Supporting Information Text S3; Figure S2). Therefore, we reported the results using random cross-validation. We considered the number of occurrences used to fit the model, as models with few sample points may lead to the species-environment relationships being poorly captured (Wisz et al., 2008). Stochasticity may also lead to poor predictive performance, particularly at low numbers (Yates et al., 2018), and thus we included the number of occurrences in the exotic range. To examine generalities within large geographic areas, we included the invaded continent as a categorical predictor. Multi-continental invasions were treated as a unique level within the factor.

We considered differences in transferability between taxonomic classes. The taxonomic class of each species was obtained using the 'taxize' package in R (Chamberlain & Szöcs, 2013). Relevant species traits, such as their dispersal ability (Miller & Holloway, 2015; Syphard & Franklin, 2010), may directly influence species distributions and consequently the performance of SDMs in a predictive context. We included the mobility of species in their adult life stage as a binary predictor, where sessile species were assigned a value of 1 and species capable of movement were given a value of 0. The levels of mobility were highly collinear with the taxonomic class of the species and thus were assessed separately.

We examined the effects of ecological generalism on transferability (Brotons et al., 2004; Zhang et al., 2015) using two environment-based predictors: environmental breadth and environmental coverage (Supporting Information Text S4). We quantified the environmental breadth and coverage of each native species distribution by projecting the occurrence cells, as well as the background environment cells into a two-dimensional environment space using principal component analysis (PCA; Pearson, 1901). The environmental breadth of a species was defined as the area of the environment space encompassing the occurrences in the native range, representing their specialization (Supporting Information Text S4). Environmental coverage was calculated by dividing the environmental breadth of the projected occurrences by the environmental breadth of all cells in the range (i.e. the occurrence and background cells combined; Figure S3). We incorporated environmental coverage to estimate how much of the available environment the species occupied, where a high environmental coverage may result in an SDM that cannot distinguish between occurrences and background sites. We examined the environmental breadth and coverage of species for both their native and exotic ranges.

We considered invasion specific factors in addition to the general predictors of transferability (Table 1). We expected species to be farther from equilibrium, with lower exposure (i.e. propagule pressure) and earlier in an invasion, resulting in distributions being driven by other factors besides environmental constraints (Leung et al., 2012, Runquist et al., 2019). We used gross domestic product (GDP) as a proxy of propagule pressure, which is strongly correlated with trade-related transport (Sardain et al., 2019). The GDP of the range was calculated using the sum of all cells within its regions using a gridded GDP dataset (at 5-arcmin resolution) provided by Kummu et al. (2018). We also included the year of first record within an invaded range as a surrogate for time since invasion, using Seebens et al. (2017) first records database, and supplemented with the earliest recorded GBIF presence associated with that species and range (GBIF, 2020).

As transferability may be influenced by non-analogous conditions when extrapolating (Yates et al., 2018), we examined the dissimilarity between the native and exotic ranges using geographical and environmental distances as well as islands versus mainlands. We calculated geographic distances as the Haversine distance (in kilometres) between the centroids of the native and exotic ranges with background environment points included. We estimated the similarity between ranges using a multivariate environmental similarity surfaces (MESS; Elith et al., 2010). MES values measure the similarity of a given point to a set of reference points for each environmental predictor, providing a continuous measure with positive values indicating cells that are environmentally similar and negative values indicating cells that are dissimilar (Elith et al., 2010). We calculated environmental similarity as the proportion of negative MES values in the exotic range with the native range as reference sites and the environmental predictors used to fit the SDMs, using the ‘MESS’ function from the ‘modEvA’ package (Barbosa et al., 2013). For our analysis of islands, ranges were classified as island or mainland based on whether all of its regions were contained within the global shoreline vector and islands database by Sayre et al. (2018).

We included all predictors into a “transferability model” to examine the potential factors that influence the predictive success of SDMs in the exotic range. We used the AUC value in the exotic range as the response variable. The transferability model was generated as a linear mixed-effects model (LMM; Breslow & Clayton, 1993) using the ‘lmer’ function from the ‘lme4’ package (Bates et al., 2015). We included species as a random effect factor, as species could invade multiple disjoint ranges. Nominal variables use one level of the factor as the reference, which we set as *Mammalia* and North America for taxonomic class and continent, respectively. Continuous variables were scaled to a mean of zero and standard deviation of one to allow comparability between fitted model parameters. Both native and exotic AUC values were logit-transformed, while the number of occurrences, environmental breadth and GDP were log-transformed (Table 1). None of the continuous or binary predictors were highly correlated with each other ($|r| < 0.7$; Supporting Information Table S4). We excluded species-locations that did not have available data for the factors, resulting in 648 species and 1918 species-location combinations. The vast majority of species in the analyses were plants (Supporting Information Table S2), comprising 591 of the 648 species and 4 of the 8 classes (*Liliopsida*, *Magnoliopsida*, *Pinopsida*, and *Polypodiopsida*). Other classes were examined, consisting of 13 mammals (*Mammalia*), 21 birds (*Aves*), 19 insects (*Insecta*) and 4 reptiles (*Reptilia*). We reported the R^2 following Nakagawa et al. (2017) for the mixed-effects models to provide the conditional (the variance explained by both the fixed and random factors) and marginal (the variance explained solely by the fixed factors) R^2 values.

RESULTS

Species distribution modelling performance

Both SDM approaches, GAM and MaxEnt, were able to accurately discriminate between presences and background sites in the range used for fitting when tested using 10-fold cross-validation (Figure 2). SDMs fitted using GAMs had a mean AUC of 0.809 ($N = 648$; standard deviation (s) = 0.076) when predicting the native range. SDMs fitted using MaxEnt performed better within the fitting (native) range with a mean AUC of 0.830 ($N = 648$; $s = 0.070$). However, when extrapolated to one or more exotic ranges, a significant drop in model performance was observed for both GAM and MaxEnt, with both model approaches performing similarly (Figure 2). The extrapolated GAMs had a mean AUC of 0.686 ($N = 1918$; $s = 0.124$), while MaxEnt models had a mean AUC of 0.689 ($N = 1918$; $s = 0.122$). Roughly half of model extrapolations were poor, where 1078 (56.2%) and 1076 (56.1%) predicted exotic ranges had AUC values below 0.7 for GAMs and MaxEnt, respectively. 229 (11.9%) and 228 (11.9%) predictions had AUC values between 0.7 and 0.75, 227 (11.8%) and 222 (11.6%) predictions had AUC values between 0.75 and 0.8, and 384 (20.0%) and 392 (20.4%) predictions had AUC values greater than 0.8, for GAM and MaxEnt respectively.

Predictors of model performance

The transferability model, fitted using a linear mixed model had a marginal R^2 value of 0.224, and a conditional R^2 value of 0.323 for GAMs. The transferability model fitted with MaxEnt had a marginal R^2 value 0.236 and conditional R^2 value 0.291. Therefore, while the predictors included in the model explained some of the variation associated with model transferability, a considerable amount of unexplained variation remained. Of the 17 fixed effects used to characterize model transferability, 11 predictors were significant for the transferability model fitted for GAMs, and 9 were significant for the transferability model fitted for MaxEnt (Table 2; Supporting Information Table S5, Table S6; Figure 3).

We found that higher discriminatory ability (AUC) in the native range was positively related to AUC in the exotic range, as expected (Table 2; Figure 3). In contrast, the number of occurrences used to fit the model did not significantly affect transferability, while the number of occurrence points in the invaded range was positively related for GAMs but not significant with MaxEnt models. Transferability varied geographically, with significant differences between continents (Table 2; Figure 4). SDMs had worse discrimination in Asia, South America and Africa. The United States, Australia and Mexico, which accounted for most of the recorded invasions in the dataset, performed better than the average with mean AUC values greater than 0.7 (but less than 0.75; Supporting Information Table S7).

The transferability of the SDMs significantly differed between taxonomic classes for both GAMs and MaxEnt (Table 2; Table 3). Birds and mammals had the lowest discriminatory performance when predicting the invaded range. Birds had a mean AUC of 0.638 (s = 117; N =

58) and 0.661 (*s* = 0.114) for GAMs and MaxEnt, respectively, while Mammals had a mean AUC of 0.639 (*s* = 0.112; *N* = 30) and 0.641 (*s* = 0.095). Insects showed better performance, with mean AUC values of 0.673 (*s* = 0.117; *N* = 34) and 0.676 (*s* = 0.117). Plants, which comprised 591 of the 648 species in the analyses, had an overall mean of 0.689 (*s* = 0.124; *N* = 1790) and 0.690 (*s* = 0.122) across the four plant classes. Reptiles performed well with mean AUC values of 0.816 (*s* = 0.155) and 0.738 (*s* = 0.187), however only 4 species were examined, with 6 extrapolated species-locations. The mobility of the species was assessed separately from class as they were highly collinear, and was not significant (Supporting Information Table S5). Contrary to our hypothesis, SDMs for species with larger environmental breadths in their native range performed better when extrapolating to new ranges (Table 2; Figure 3). Native environmental coverage, however, did not significantly affect transferability. Conversely, SDMs performed worse when species had a large environmental breadth and environmental coverage in the exotic range.

The GDP of the exotic range did not significantly affect model transferability (Table 2; Figure 3), while models fitted on native ranges with larger GDPs tended to perform worse, contrary to our hypotheses. Also unexpectedly, the year of first record significantly affected transferability, with newer species sightings leading to higher transferability in GAMs but not MaxEnt models. Indeed, a *post-hoc* analysis examining relatively recent invasions (after the year 1970) in North America (which showed the highest transferability; Figure 4) resulted in a respectable AUC for both GAMs (mean AUC = 0.799, *s* = 0.129, *N* = 33) and MaxEnt (mean AUC = 0.792, *s* = 0.134), suggesting that strong transferability may be possible in some circumstances.

Both GAMs and MaxEnt models tended to perform worse when extrapolated across larger geographic distances, as predicted (Table 2; Figure 3). In contrast, the degree of environmental similarity did not significantly correlate to the transferability of SDMs. On average however, only 11.7% and 15.6% of sites across all exotic ranges showed negative MES values for GAMs and MaxEnt, respectively, suggesting high environmental similarity between ranges, and potentially insufficient differences to detect an effect. Finally, SDMs performed substantially worse when attempting to predict the distribution of species on islands, as predicted. Island invasions accounted for 387 out of the 1918 species-locations, with 383 of these invasions by species originating from mainlands. Species originating from islands, however, did not significantly affect model transferability, but these only consisted of 13 out of 648 species and could simply reflect low power.

DISCUSSION

While many studies have used SDMs to forecast the distributions of exotic species (e.g. Shrestha et al., 2014; Vicente et al., 2013), recent studies on transferability have cautioned the many pitfalls associated with extrapolating SDMs to novel conditions (Sequeira et al., 2018; Werkowska et al., 2017; Yates et al., 2018). Indeed, we found that the performance of the SDMs in the exotic ranges showed a marked decline in discriminatory ability, with less than half of the model predictions being considered acceptable (based on a 0.7 AUC threshold) despite a relatively strong discriminatory performance in the native range (Morán-Ordóñez et al., 2017; Swets, 1988). On the other hand, most successfully invaded ranges were similar to their native ranges with few sites considered different on average (between 11 and 15 percent), based on MES scores. This suggests predictability of invasions based broadly on environmental conditions, at least at a regional level. Additionally, from a more optimistic perspective, while half the species failed in their SDM transferability, the other half maintained some discriminatory ability (AUC > 0.7). Whether this is sufficient to justify using SDMs built in native ranges to forecast species distributions in exotic ranges is equivocal. This highlights the necessity to identify whether transferability of such ecological models may be improved, and whether we can identify when models may be reliable (Yates et al., 2018). Ideally, such methods would be based on widely available data.

Predictors of transferability

As expected, SDMs that were better able to predict their own range (native AUC) were more likely to succeed when extrapolating to new ranges (exotic AUC). The number of occurrences included in model fitting did not influence spatial transferability, while models performed better when predicting ranges with more occurrences. The latter supports our hypothesis that a larger number of occurrences in the exotic range may result in increased strength of validation, as low occurrence numbers may be more affected by stochasticity in which species do not occur in suitable habitats by chance (Yates et al., 2018).

Discrimination in the exotic ranges differed between taxonomic classes, suggesting that SDMs may be more reliable for some taxonomic groups over others, and may be linked by shared traits (Regos et al., 2019). In contrast, the mobility of the species did not significantly affect model transferability, however other traits more specific to the taxonomic groups may influence transferability and should be assessed in the future. For instance, model performance has been shown to differ between bird species based on traits related to biogeographic origin, migratory status, and habitat preference (Regos et al., 2019). Other potential traits include commonness, body size, dispersal ability and trophic position (Dobrowski et al., 2011; Soininen & Luoto, 2014; Urban et al., 2016; Wogan, 2016).

Transferability also differed between invaded continents, suggesting that generalities can be made on whether models may be more reliable when extrapolating to certain areas of the world. For example, geographical biases have been observed in data availability (Yesson et al., 2007) and research focus (Pyšek et al., 2008), where some areas such as Africa may be poorly sampled resulting in worse transferability, while North America, with high sampling intensity, showed the strongest performance.

Several of our findings differed from our *a priori* expectations, suggesting other processes were in play and requiring new hypotheses. While exotic environmental breadth and coverage negatively influenced model transferability (as predicted; Brotons et al., 2004; Brotons et al., 2007; Zhang et al., 2015), environmental breadth in the native range was positively related to transferability (contrary to our hypothesis). As a potential explanation, the positive relationship between native environmental breadth and SDM transferability may be due to a censoring effect, wherein we only can analyze successful establishments. Smaller niche breadths may result in better discrimination in the native range in general, but most of these specialists may fail to establish at all in any exotic range (i.e., they would never enter the analysis). Species with narrow environmental breadths that did establish may then represent species restricted by other processes (e.g. biotic factors; Urban et al., 2016), that are released when invading a new region and may therefore occur outside of its realized niche and be poorly predicted. This *post hoc* hypothesis, however, requires further testing.

Also contrary to our initial hypotheses, we found that higher exposure led to poorer model performance. We had expected that as species had greater opportunity to spread, exposure should become less a limiting factor, thereby allowing environmental determinants to become more important in shaping species distributions (Leung et al., 2012, Runquist et al., 2019). First, using GDP as a proxy of propagule pressure (GDP should be positively related to the amount of trade-related movement, and thus to propagule pressure; Sardain et al., 2019) we found a negative relationship between GDP in the native range and SDM transferability, suggesting that higher exposure may actually lead to poorer predictions. Second, species that had more recent recorded arrivals also showed better spatial transferability, contradicting the prediction that newer invaders would be less predictable (as these species may still be spreading and farther from equilibrium with its environment; Václavík & Meentemeyer, 2011). One possibility is that older invaders have evolved over time to diverge from the native range population (Mooney & Cleland, 2001). However, the majority of invasions included in the study were observed after the year 1900, and therefore evolution would have had to be rapid. Alternatively, species may first establish in the most suitable habitats, as these should offer the highest probabilities of survival. Instead of being stochastic events, driven primarily by exposure, early phases of invasion could be driven by the interaction between propagule pressure and environmental suitability. With increased exposure, species at higher population densities may begin to occupy less suitable sites due to processes such as competition, resulting in poorer transferability (McLoughlin et al., 2010). From a management perspective, this could be important as it suggests that SDMs may be able to predict "recent" invasions, despite being far from their equilibrium distribution.

Recent publications have highlighted the risk of extrapolating to non-analogous conditions as models fitted on a truncated range of environmental data may fail to predict species response beyond these limits (Sequeira et al., 2018; Yates et al., 2018). Consistent with this, SDMs exhibited poorer discriminatory performance when extrapolated to more distance ranges. Additionally, SDMs performed worse when extrapolating to islands, where environments often differ (Vitousek, 1990), although other non-ecological factors could also be at play (for example, poorer data quality on islands; Ficetola et al., 2014). While the environmental dissimilarity between ranges did not influence model success, few sites were dissimilar between native and exotic ranges for the climatic variables examined. Nonetheless, given the low discriminatory performance on average in exotic ranges, the similarity between measured environmental conditions was insufficient to confer high transferability.

Limitations and future directions

The SDMs in this study were generated using presence-only data, which is the most common approach due to the wide availability of occurrence records (Elith and Leathwick, 2009). SDMs fitted using presence-only data do not provide absolute probabilities of occurrence, but rather only relative probabilities (Pearce & Boyce, 2006). Despite survey data representing a more rigorous sampling design, such data is less available for most species (Barbet-Massin et al., 2012). Thus, presence-only SDMs will likely remain an essential part of ecological modelling.

Although we found only low to moderate transferability, we acknowledge that we primarily focused on climate-based variables. Other drivers of species distributions have been suggested such as land use, ecosystem functional attributes and biotic factors (Arenas-Castro et al., 2018; Regos et al., 2019; Urban et al., 2016). Not surprisingly though, the majority of ecological models have been based on climatic predictors, given the challenges of implementing mechanistic predictors based on experimental data (Werkowska et al., 2017) and the greater availability of high resolution global climate data (e.g. Fick & Hijmans, 2017).

We identified several factors that influenced spatial transferability, with the fixed effects explaining 22.4% of the variance associated with model performance. Of course, only a subset of possible predictors of transferability were assessed in this study, and given the relatively low average transferability of SDMs, identifying additional predictors to better understand the reliability is crucial (Yates et al., 2018). For example, functional trait-based approaches in ecology are becoming increasingly popular though data remains scarce, particularly for many rarer species (Majekova et al., 2016). Beyond the choice of predictors included in the model, SDM transferability may also be influenced by the quality of the dataset itself, which has seldom been examined as a source of failure (Datta et al., 2020). Other factors have been proposed when contrasting ranges as well, such as changes in environmental correlation structures (Werkowska et al., 2017) and biotic interactions (Urban et al., 2016). Much room exists for improvement in identifying what might make SDMs more transferable. If such factors were simultaneously based on widely available data, they could have a substantial effect in improving the usage of SDMs in exotic ranges.

Conclusion

Many studies have emphasized the need for proper validation of predictive models when applying them to new ecological contexts. In this study, we provided the most extensive assessment of spatial transferability to date, providing vital insights into the future use of SDMs for predicting biological invasions. Caution is warranted, as half of the SDMs exhibited poor discriminatory performance in the exotic range. However, SDM performance was heterogeneous, and may be more reliable in areas with more data and for more recent invasions (e.g., invasions in North America after 1970). Identifying predictors of model transferability and improving the reliability of SDMs for biological invasions should be a priority.

REFERENCES

Acker, J. G., & Leptoukh, G. (2007). Online analysis enhances use of NASA earth science data. *Eos, Transactions American Geophysical Union*, 88(2), 14-17.

Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, 93(7), 1527-1539.

Arenas-Castro, S., Gonçalves, J., Alves, P., Alcaraz-Segura, D., & Honrado, J. P. (2018). Assessing the multi-scale predictive ability of ecosystem functional attributes for species distribution modelling. *PLoS One*, 13(6), e0199292.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many?. *Methods in ecology and evolution*, 3(2), 327-338.

Barbosa, A. M., Real, R., Muñoz, A. R., & Brown, J. A. (2013). New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Diversity and Distributions*, 19(10), 1333–1338.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.

Beaumont, L. J., Gallagher, R. V., Thuiller, W., Downey, P. O., Leishman, M. R., & Hughes, L. (2009). Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions*, 15(3), 409-420.

Bellard, C., Cassey, P., & Blackburn, T. M. (2016). Alien species as a driver of recent extinctions. *Biology letters*, 12(2), 20150623.

Bradshaw, C. J., Leroy, B., Bellard, C., Roiz, D., Albert, C., Fournier, A., ... & Courchamp, F. (2016). Massive yet grossly underestimated global costs of invasive insects. *Nature communications*, 7(1), 1-8.

Breslow, N. E., & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88(421), 9-25.

Brotons, L., Thuiller, W., Araújo, M. B., & Hirzel, A. H. (2004). Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, 27(4), 437-448.

Brotons, L., Herrando, S., & Pla, M. (2007). Updating bird species distribution at large spatial scales: applications of habitat modelling to data from long-term monitoring programs. *Diversity and distributions*, 13(3), 276-288.

Buja, A., Hastie, T., & Tibshirani, R. (1989). Linear smoothers and additive models. *The Annals of Statistics*, 17(2), 453-510.

CABI (2020). *CABI Invasive Species Compendium*. Wallingford, UK: CAB International. Retrieved from http://www.cabi.org/isc/

Capinha, C., Leung, B., & Anastácio, P. (2011). Predicting worldwide invasiveness for four major problematic decapods: an evaluation of using different calibration sets. *Ecography*, 34(3), 448-459.

Chamberlain, S. A., & Szöcs, E. (2013). taxize: taxonomic search and retrieval in R. *F1000Research*, 2.

Datta, A., Schweiger, O., & Kühn, I. (2020). Origin of climatic data can determine the transferability of species distribution models. *NeoBiota*, 59, 61.

Dobrowski, S. Z., Thorne, J. H., Greenberg, J. A., Safford, H. D., Mynsberge, A. R., Crimmins, S. M., & Swanson, A. K. (2011). Modeling plant ranges over 75 years of climate change in California, USA: temporal transferability and species traits. *Ecological Monographs*, 81(2), 241-257.

Elith, J. (2017). Predicting distributions of invasive species. *Invasive species: Risk assessment and management*, 10(9781139019606.006).

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

Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in ecology and evolution*, *1*(4), 330-342.

Ficetola, G. F., Cagnetta, M., Padoa-Schioppa, E., Quas, A., Razzetti, E., Sindaco, R., & Bonardi, A. (2014). Sampling bias inverts ecogeographical relationships in island reptiles. *Global Ecology and Biogeography*, 23(11), 1303-1313.

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(12), 4302-4315.

Franklin, J. (2009). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press.

GBIF (2020). Global Biodiversity Information Facility Occurrence Download. Retrieved from https://doi.org/10.15468/dl.g85fea

Global Invasive Species Database (GISD) (2015). *The Global Invasive Species Database*. Retrieved from http://www.issg.org/database

Global Administrative Areas (2018). GADM database of Global Administrative Areas, version 3.6. Retrieved from https://gadm.org/data.html

Goncalves, E., Herrera, We., Duarte, M., Bustamante, R. O., Lampo, M., Velasquez, G., ... & García-Rangel, S. (2014). Global invasion of *Lantana camara*: has the climatic niche been conserved across continents?. *PLoS One*, 9(10), e111468.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology letters*, 8(9), 993-1009.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29-36.

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

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(2), 142-152.

Jiménez-Valverde, A. (2012). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, 21(4), 498-507.

Kier, G., Kreft, H., Lee, T. M., Jetz, W., Ibisch, P. L., Nowicki, C., ... & Barthlott, W. (2009). A global assessment of endemism and species richness across island and mainland regions. *Proceedings of the National Academy of Sciences*, 106(23), 9322-9327.

Kummu, M., Taka, M., & Guillaume, J. H. (2018). Gridded global datasets for gross domestic product and Human Development Index over 1990–2015. *Scientific data*, 5, 180004.

Leung, B., Roura-Pascual, N., Bacher, S., Heikkilä, J., Brotons, L., Burgman, M. A., ... & Sol, D. (2012). TEASIng apart alien species risk assessments: a framework for best practices. *Ecology Letters*, 15(12), 1475-1493.

Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2007). AUC: a misleading measure of the performance of predictive distribution models. *Global ecology and Biogeography*, 17(2), 145-151.

Lockwood, J. L., Cassey, P., & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends in ecology & evolution*, 20(5), 223-228.

Lozier, J. D., & Mills, N. J. (2011). Predicting the potential invasive range of light brown apple moth (*Epiphyas postvittana*) using biologically informed and correlative species distribution models. *Biological Invasions*, 13(10), 2409.

Lunetta, R. S., Knight, J. F., Ediriwickrema, J., Lyon, J. G., & Worthy, L. D. (2006). Land-cover change detection using multi-temporal MODIS NDVI data. *Remote sensing of environment*, 105(2), 142-154.

Majekova, M., Paal, T., Plowman, N. S., Bryndova, M., Kasari, L., Norberg, A., ... & Le Bagousse-Pinguet, Y. (2016). Evaluating functional diversity: missing trait data and the importance of species abundance structure and data transformation. *PloS one*, 11(2).

McLoughlin, P. D., Morris, D. W., Fortin, D., Vander Wal, E., & Contasti, A. L. (2010). Considering ecological dynamics in resource selection functions. *Journal of animal ecology*, 79(1), 4-12.

Merow, C., Smith, M. J., Edwards Jr, T. C., Guisan, A., McMahon, S. M., Normand, S., ... & Elith, J. (2014). What do we gain from simplicity versus complexity in species distribution models?. *Ecography*, 37(12), 1267-1281.

Miller, J. A., & Holloway, P. (2015). Incorporating movement in species distribution models. *Progress in Physical Geography*, 39(6), 837-849.

Mooney, H. A., & Cleland, E. E. (2001). The evolutionary impact of invasive species. *Proceedings of the National Academy of Sciences*, 98(10), 5446-5451.

Morán-Ordóñez, A., Lahoz-Monfort, J. J., Elith, J., & Wintle, B. A. (2017). Evaluating 318 continental-scale species distribution models over a 60-year prediction horizon: what factors influence the reliability of predictions?. *Global Ecology and Biogeography*, 26(3), 371-384.

Morlini, I. (2006). On multicollinearity and concurvity in some nonlinear multivariate models. *Statistical Methods and Applications*, 15(1), 3-26.

Nakagawa, S., Johnson, P. C., & Schielzeth, H. (2017). The coefficient of determination R 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, 14(134), 20170213.

Pearce, J. L., & Boyce, M. S. (2006). Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, 43(3), 405-412.

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11), 559-572.

Peterson, A. T., Papes, M., & Kluza, D. A. (2003). Predicting the potential invasive distributions of four alien plant species in North America. *Weed Science*, 51(6), 863-868.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C., & Guisan, A. (2012). Climatic niche shifts are rare among terrestrial plant invaders. *Science*, 335(6074), 1344-1348.

Petitpierre, B., Broennimann, O., Kueffer, C., Daehler, C., & Guisan, A. (2017). Selecting predictors to maximize the transferability of species distribution models: lessons from cross-continental plant invasions. *Global Ecology and Biogeography*, 26(3), 275-287.

Pettorelli, N., Ryan, S., Mueller, T., Bunnefeld, N., Jędrzejewska, B., Lima, M., & Kausrud, K. (2011). The Normalized Difference Vegetation Index (NDVI): unforeseen successes in animal ecology. *Climate research*, 46(1), 15-27.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological modelling*, 190(3-4), 231-259.

Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2), 161-175.

Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological applications*, 19(1), 181-197.

Phillips, S. J. (2017). maxnet: Fitting 'Maxent' Species Distribution Models with 'glmnet'. R package version 0.1.2. Retrieved from https://CRAN.R-project.org/package=maxnet

Pyšek, P., Richardson, D. M., Pergl, J., Jarošík, V., Sixtova, Z., & Weber, E. (2008). Geographical and taxonomic biases in invasion ecology. *Trends in ecology & evolution*, 23(5), 237-244.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Regos, A., Gagne, L., Alcaraz-Segura, D., Honrado, J. P., & Domínguez, J. (2019). Effects of species traits and environmental predictors on performance and transferability of ecological niche models. *Scientific Reports*, 9(1), 1-14.

Runquist, R. D. B., Lake, T., Tiffin, P., & Moeller, D. A. (2019). Species distribution models throughout the invasion history of Palmer amaranth predict regions at risk of future invasion and reveal challenges with modeling rapidly shifting geographic ranges. *Scientific reports*, 9(1), 1-12.

Sardain, A., Sardain, E., & Leung, B. (2019). Global forecasts of shipping traffic and biological invasions to 2050. *Nature Sustainability*, 2(4), 274-282.

Sayre, R., Noble, S., Hamann, S., Smith, R., Wright, D., Breyer, S., ... & Hopkins, D. (2018). A new 30 meter resolution global shoreline vector and associated global islands database for the development of standardized ecological coastal units. *Journal of Operational Oceanography*, 12(sup2), S47-S56.

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., ... & Bacher, S. (2017). No saturation in the accumulation of alien species worldwide. *Nature communications*, 8(1), 1-9.

Sequeira, A. M., Bouchet, P. J., Yates, K. L., Mengersen, K., & Caley, M. J. (2018). Transferring biodiversity models for conservation: opportunities and challenges. *Methods in Ecology and Evolution*, 9(5), 1250-1264.

Shrestha, U. B., & Bawa, K. S. (2014). Impact of climate change on potential distribution of Chinese caterpillar fungus (*Ophiocordyceps sinensis*) in Nepal Himalaya. *PLoS one*, 9(9).

Sofaer, H. R., Jarnevich, C. S., & Flather, C. H. (2018). Misleading prioritizations from modelling range shifts under climate change. *Global ecology and biogeography*, 27(6), 658-666.

Soininen, J., & Luoto, M. (2014). Predictability in species distributions: a global analysis across organisms and ecosystems. *Global ecology and biogeography*, 23(11), 1264-1274.

Spatz, D. R., Zilliacus, K. M., Holmes, N. D., Butchart, S. H., Genovesi, P., Ceballos, G., ... & Croll, D. A. (2017). Globally threatened vertebrates on islands with invasive species. *Science Advances*, 3(10), e1603080.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240(4857), 1285-1293.

Syphard, A. D., & Franklin, J. (2010). Species traits affect the performance of species distribution models for plants in southern California. *Journal of Vegetation Science*, 21(1), 177-189.

Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J. B., Pe'er, G., Singer, A., ... & Gonzalez, A. (2016). Improving the forecast for biodiversity under climate change. *Science*, 353(6304), aad8466.

Václavík, T., & Meentemeyer, R. K. (2011). Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions*, 18(1), 73-83.

Vicente, J. R., Fernandes, R. F., Randin, C. F., Broennimann, O., Gonçalves, J., Marcos, B., ... & Honrado, J. P. (2013). Will climate change drive alien invasive plants into areas of high protection value? An improved model-based regional assessment to prioritise the management of invasions. *Journal of environmental management*, 131, 185-195.

Vitousek, P. M. (1990). Biological invasions and ecosystem processes: towards an integration of population biology and ecosystem studies. In *Ecosystem management* (pp. 183-191). Springer, New York, NY.

Wang, Q., Punchi-Manage, R., Lu, Z., Franklin, S. B., Wang, Z., Li, Y., ... & Xu, Y. (2017). Effects of topography on structuring species assemblages in a subtropical forest. *Journal of Plant Ecology*, 10(3), 440-449.

Wen, L., Saintilan, N., Yang, X., Hunter, S., & Mawer, D. (2015). MODIS NDVI based metrics improve habitat suitability modelling in fragmented patchy floodplains. *Remote Sensing Applications: Society and Environment*, 1, 85-97.

Werkowska, W., Márquez, A. L., Real, R., & Acevedo, P. (2017). A practical overview of transferability in species distribution modeling. *Environmental reviews*, 25(1), 127-133.

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A., & NCEAS Predicting Species Distributions Working Group. (2008). Effects of sample size on the performance of species distribution models. *Diversity and distributions*, 14(5), 763-773.

Wogan, G. O. (2016). Life history traits and niche instability impact accuracy and temporal transferability for historically calibrated distribution models of North American birds. *PLoS One*, 11(3).

Wood, S. N. (2017). *Generalized additive models: an introduction with R*. Chapman and Hall/CRC.

Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., ... & Dormann, C. F. (2018). Outstanding challenges in the transferability of ecological models. Trends in ecology & evolution, 33(10), 790-802.

Yesson, C., Brewer, P. W., Sutton, T., Caithness, N., Pahwa, J. S., Burgess, M., ... & Culham, A. (2007). How global is the global biodiversity information facility?. *PloS one*, 2(11), e1124.

Zhang, L., Liu, S., Sun, P., Wang, T., Wang, G., Zhang, X., & Wang, L. (2015). Consensus forecasting of species distributions: The effects of niche model performance and niche properties. *PLoS One*, 10(3), e0120056.

Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D., ... & Svantesson, S. (2019). CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, 10(5), 744-751.

DATA AVAILABILITY STATEMENT

Species records are available through GBIF (https://www.gbif.org/), with DOIs for occurrences and background sites provided in Supporting Information Text S1 and Table S3, respectively. Distributional data for invasive species are available through GISD (http://www.iucngisd.org/gisd/) and CABI Invasive Species Compendium (https://www.cabi.org/ISC). GADM ESRI shapefiles for global subdivisions are available from http://gadm.org/. Global shoreline vector for islands are available from Sayre et al., 2018 (https://rmgsc.cr.usgs.gov/outgoing/ecosystems/Global/). Bioclimatic and elevation data are available through WorldClim v2 (https://www.worldclim.org/). MODIS NDVI data was accessed through NASA GES DISC (https://giovanni.gsfc.nasa.gov/giovanni/). First records data for invasive species was provided by Seebens et al., 2017's Alien Species First Records data set (https://zenodo.org/record/3690742#.X3uheu17mUk/). Gridded GDP dataset are available through Kummu et al., 2018 (https://datadryad.org/stash/dataset/doi:10.5061/dryad.dk1j0).

TABLES

Table 1. Full list of variables and acronyms used to predict the transferability of SDMs, with a description of how the factors were generated. Rationale for why each factor may influence SDM transferability is provided for each factor, with sources. Variables marked with an asterisk (*) denote factors generated for both the native and exotic ranges.

Variable (acronym)	Description	Rationale
Native AUC (NAUC)	Performance of the SDM on its own range measured as AUC, using 10-fold cross-validation. Logit-transformed.	Models that fail to characterize its own range are likely to fail when extrapolating to new environments (Morán-Ordóñez et al., 2017).
Number of occurrences (NOCC; EOCC)*	Number of occurrences used to fit or validate the model. Multiple occurrences found in a single cell were counted as a single occurrence. Log-transformed.	Ranges with small sample sizes may fail to capture complex ecological relationships and be prone to stochasticity, leading to increased parameter uncertainty and unstable results (Wisz et al., 2008; Yates et al., 2018).
Continent (CONT)	Nominal variable of the invaded continent. Multi-continental invasions were treated as their own level.	Generalities in environmental conditions and biotic composition within large regional areas, as well as differences in quality of invasive species records and biodiversity information (Pyšek et al., 2008) may affect SDM transferability.
Taxonomic class (CLASS)	Nominal variable of the species' taxonomic class.	SDM transferability may differ between groups of species, or their habitat, that may share similar characteristics (Regos et al., 2019).
Sessile (SESS)	Binary variable on whether the species is sessile (1) or mobile (0)	Species traits such as the behaviour and movement of species may affect our ability to accurately records species occurrences, and influence the dispersal ability of species to reach all available sites (Syphard & Franklin, 2010; Miller & Holloway, 2015).
Environmental breadth (NEB; EEB)*	Convex hull area of occurrence points projected into environmental PCA space. Log-transformed.	Generalist species may not be restricted by the environmental factors considered in the SDM, and result in poor transferability (Brotons et al., 2004; Zhang et al., 2015).
Environmental coverage (NEC; EEC)*	Proportion of total environmental space of the background sites and occurrences occupied by only the occurrences.	A large environmental coverage may lead to SDMs being unable to distinguish between presences and their background environment (Brotons et al., 2004; Zhang et al., 2015).
Gross domestic product (NGDP; EGDP)*	Sum of the GDP of all defined regions in the range. Log-transformed.	Increased propagule pressure (e.g. through trade) may lead to increased dispersal to all suitable locations for a given range ('exposure'), resulting in species closer to equilibrium with their environment (Lockwood et al., 2005; Leung et al., 2012).
Year of first record (YEAR)	Earliest recorded occurrence of the species within the defined invaded range.	Newer invaders may still be spreading, and not yet reached all suitable habitats, violating assumptions of equilibrium (Runquist et al., 2019).
Geographic distance between ranges (DIST)	Haversine distance (in km) between the centroids of the native and exotic ranges, including background environment.	Species may experience new environmental conditions or community compositions outside of the range in the fitting data leading to unpredictable responses (Werkowska et al., 2017; Yates et al., 2018).
Environmental similarity between ranges (ESIM)	Proportion of negative MES values in the exotic range with the native range as reference sites.	
Island (NISL; EISL)*	Binary predictor on whether all occurrences were found on an island (1) or not (0). Native and exotic.	Ecological characteristics of islands differ significantly from the mainland - higher endemism, lower species richness and restricted distributions - and may greatly affect species distribution (Kier et al., 2009).

Table 2. Estimated model parameters, with standard error, *t*-value and *P*-values of the fixed effects for the linear mixed model fitted using the full dataset for GAMs and MaxEnt with taxonomic class as a categorical predictor.

	GAM				**MaxEnt**			
Variable	**Estimate**	**Std. error**	***t*-statistic**	***P*-value**	**Estimate**	**Std. error**	***t*-statistic**	***P*-value**
Native AUC (NAUC)*	0.089	0.019	4.592	< 0.001	0.058	0.018	3.237	0.001
Native number of occurrences (NOCC)	-0.028	0.026	-1.051	0.295	-0.001	0.024	-0.055	0.956
Exotic number of occurrences (EOCC)^G	0.047	0.022	2.097	0.036	0.036	0.022	1.627	0.105
Continent (CONT)*	-	-	-	< 0.001	-	-	-	< 0.001
Taxonomic class (CLASS)*	-	-	-	0.017	-	-	-	0.010
Native environmental breadth (NEB)*	0.153	0.028	5.492	< 0.001	0.182	0.027	6.632	< 0.001
Exotic environmental breadth (EEB)*	-0.108	0.032	-3.414	< 0.001	-0.121	0.032	-3.835	< 0.001
Native environmental coverage (NEC)	0.011	0.027	0.419	0.677	-0.012	0.025	-0.484	0.629
Exotic environmental coverage (EEC)*	-0.153	0.027	-5.670	< 0.001	-0.142	0.027	-5.284	< 0.001
Native GDP (NGDP)*	-0.074	0.025	-2.975	0.003	-0.056	0.023	-2.435	0.015
Exotic GDP (EGDP)	0.059	0.034	1.741	0.082	0.058	0.034	1.726	0.084
Year of first records (YEAR)^G	0.035	0.016	2.146	0.032	0.020	0.016	1.258	0.209
Geographic distance (DIST)*	-0.055	0.018	-3.006	0.003	-0.057	0.018	-3.123	0.002
Environmental similarity (ESIM)	0.023	0.018	1.257	0.211	0.013	0.019	0.656	0.513
Native island (NISL)	-0.035	0.137	-0.258	0.797	0.032	0.127	0.252	0.801
Exotic island (EISL)*	-0.270	0.056	-4.862	< 0.001	-0.302	0.056	-5.441	< 0.001

* Denotes factors significant for both GAM and MaxEnt

^{G} Denotes factors significant for only GAM

Table 3. Mean performance of the species distribution models divided between taxonomic groups, measured as area under the receiver operating characteristic curve (AUC), in both the native (fitting) range and the extrapolated (validation) range.

	Native					**Exotic**				
		GAM		**MaxEnt**			**GAM**		**MaxEnt**	
Taxa	**N**	**AUC**	**Std. dev.**	**AUC**	**Std. dev.**	**N**	**AUC**	**Std. dev.**	**AUC**	**Std. dev.**
Aves	21	0.751	0.080	0.772	0.080	58	0.638	0.117	0.661	0.114
Mammalia	13	0.761	0.110	0.774	0.107	30	0.639	0.112	0.641	0.095
Insecta	19	0.777	0.072	0.799	0.048	34	0.673	0.117	0.676	0.117
Plants	591	0.814	0.074	0.834	0.068	1790	0.689	0.124	0.690	0.122
Pinopsida	5	0.907	0.075	0.921	0.059	9	0.674	0.180	0.696	0.144
Liliopsida	161	0.819	0.070	0.836	0.063	465	0.685	0.123	0.686	0.120
Magnoliopsida	416	0.810	0.075	0.832	0.069	1301	0.689	0.123	0.691	0.123
Polypodiopsida	9	0.825	0.038	0.858	0.035	15	0.760	0.162	0.794	0.119
Reptilia	4	0.792	0.087	0.804	0.095	6	0.816	0.155	0.738	0.187
TOTAL	**648**	**0.809**	**0.076**	**0.830**	**0.070**	**1918**	**0.686**	**0.124**	**0.689**	**0.122**

FIGURES

Figure 1. Flow chart outlining the analyses performed in the study.

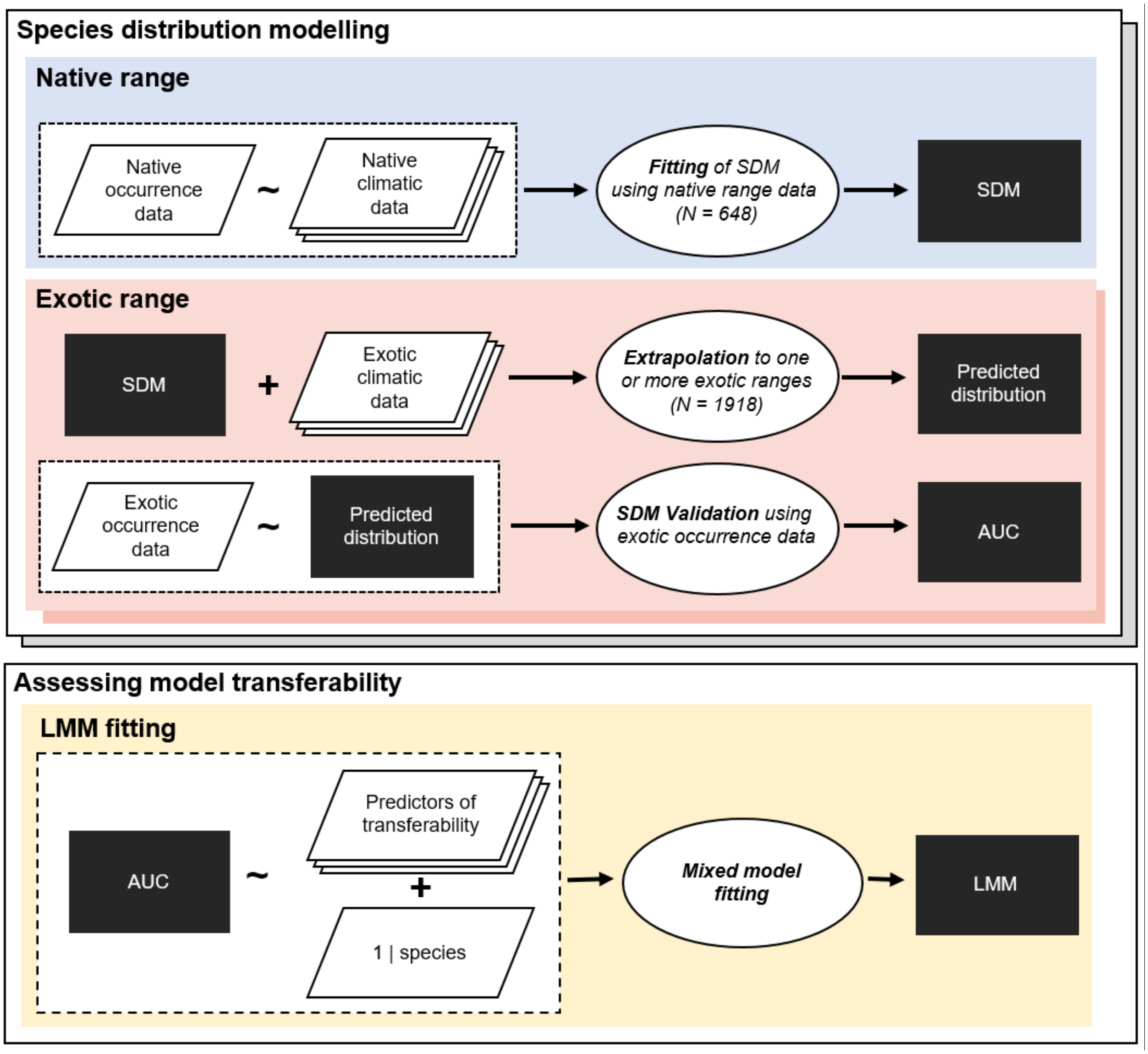


Figure 2. Density histogram of the performance of GAMs (a) and MaxEnt (b) measured as area under the receiver operating characteristic curve (AUC) when predicting the native range, or fitting dataset (blue), using 10-fold cross-validation and when extrapolating to the exotic range, or validation dataset (red).

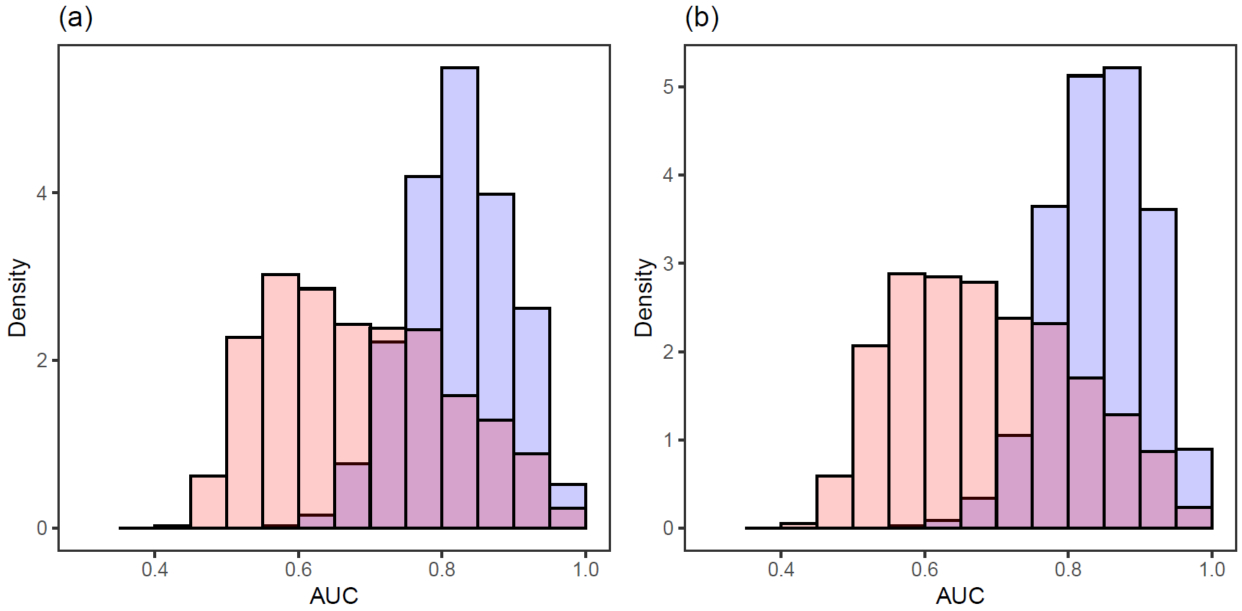


Figure 3. Dot-whisker plot of the parameter estimates for the binary and continuous predictors included in the linear mixed model for GAM (black) and MaxEnt (orange), with taxonomic class and continent as categorical predictors. Whiskers represent the 95 percent confidence intervals of the parameter estimates.

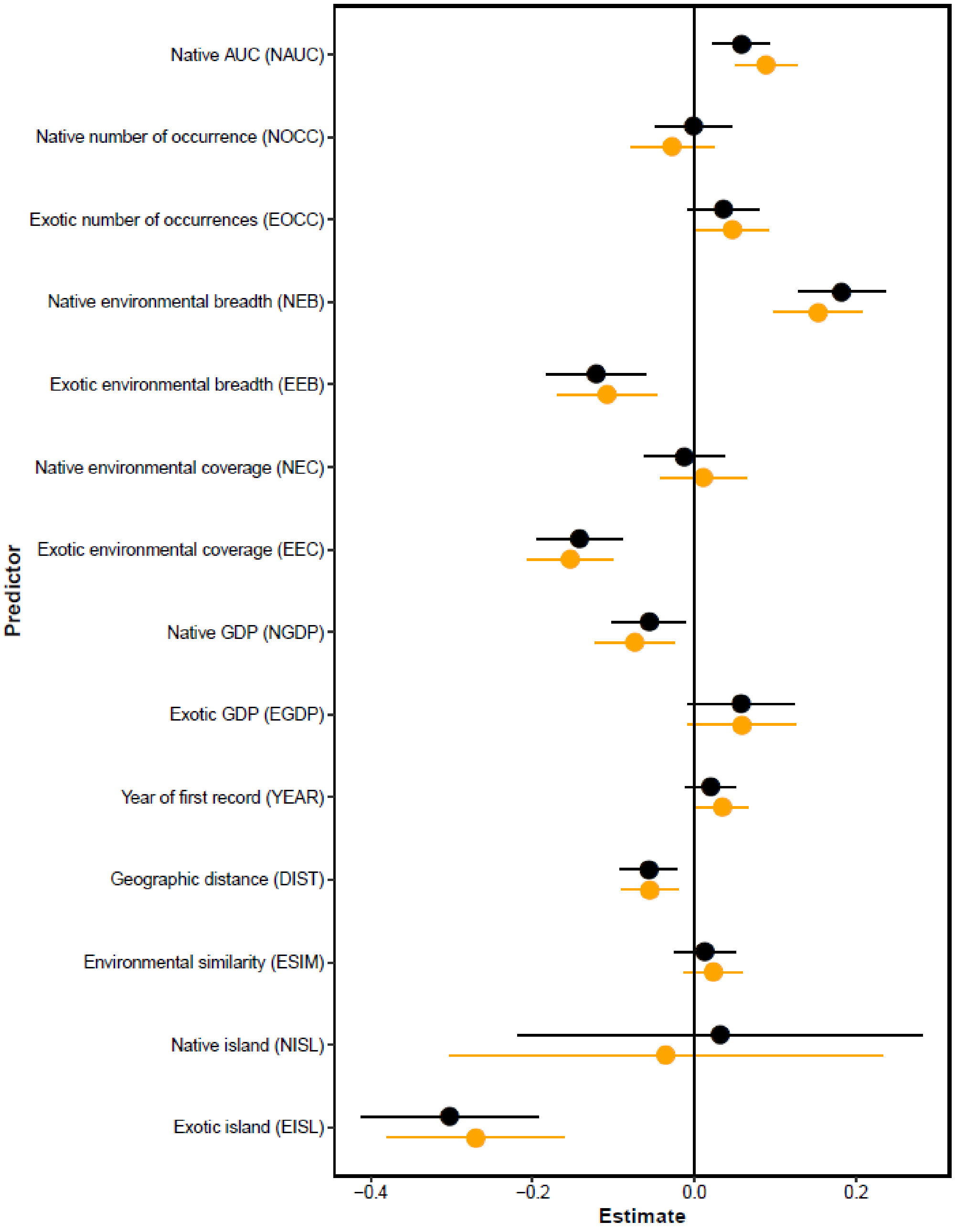
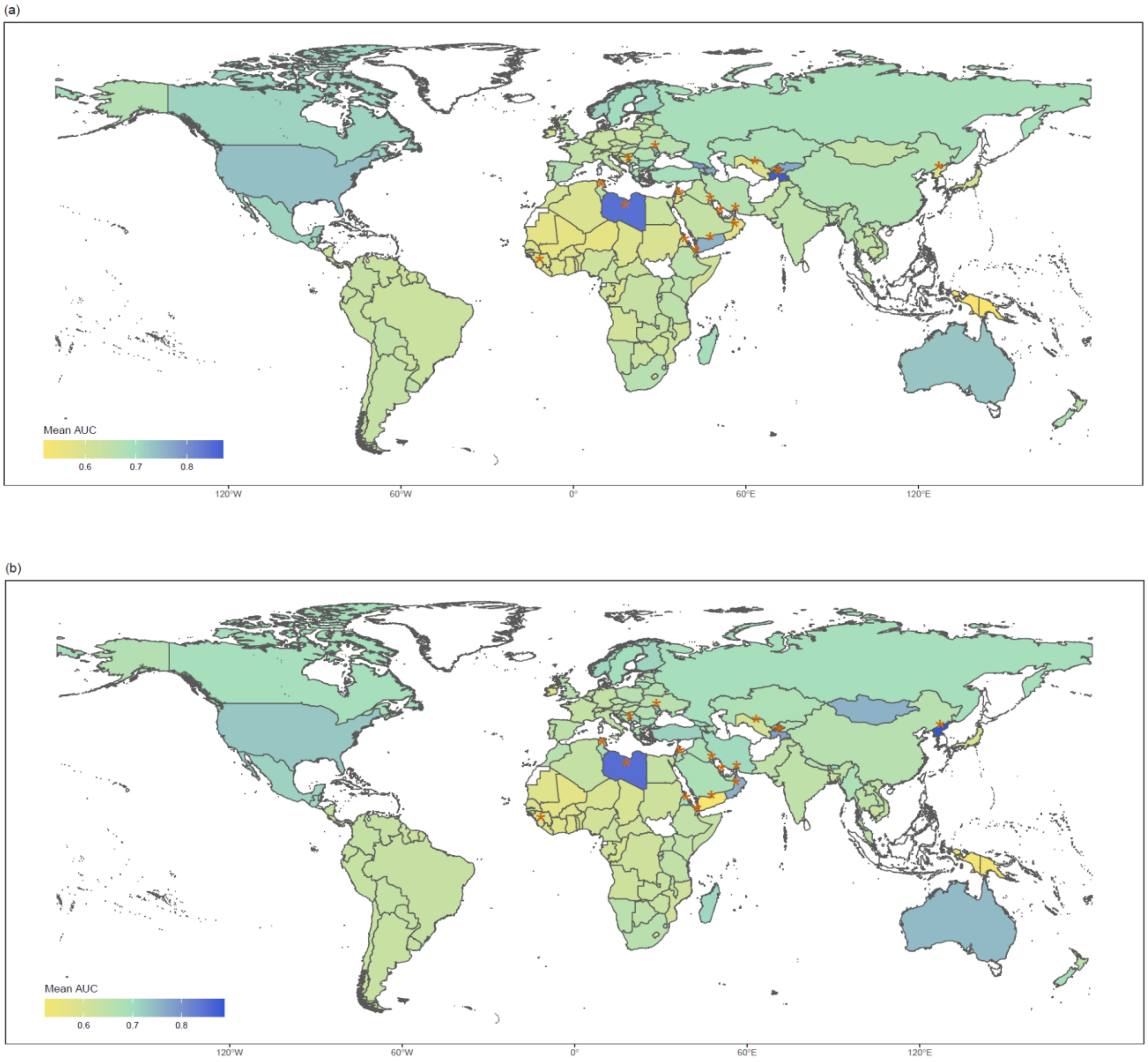


Figure 4. Map of the mean discriminatory performance (area under the receiver operating characteristic curve; AUC) of the extrapolated models using (a) GAM and (b) MaxEnt. Regions marked with an asterisk (*) represent areas that were invaded by fewer than three species.



GENERAL CONCLUSION

SDMs remain one of the most valuable mathematical tools in ecology for addressing large-scale questions in global change and biodiversity. While the potential uses of SDMs continue to be explored, further caution is warranted given the various potential causes for concern as models are used to predict novel environments (Sequeira et al., 2018; Werkowska et al., 2017; Yates et al., 2018). While several studies have assessed the spatial transferability of SDMs (e.g. Goncalves et al., 2014; Petitpierre et al., 2012), many of these studies have been restricted to a limited number of species, taxonomic groups or regions and have resulted in a lack of consensus. A strong focus has been placed on improving model predictions across new contexts, and identifying which models may be transferable prior to their extrapolation (Yates et al., 2018). Therefore, a large-scale systematic analysis of SDM transferability is needed to understand to what degree such models are currently transferable, and under what conditions they may succeed.

In this thesis, I demonstrated that despite a strong model performance in the native (fitting) range, more than half of the models showed poor spatial transferability in the exotic range. The SDMs were implemented using presence-only data and climate-based abiotic variables, which has become the most common approach given their wide availability and potential for global application (Elith and Leathwick, 2009). Response variables with greater biological information, such as abundance or presence-absence data, as well as more mechanistic or ecologically relevant predictors, may confer better transferability (Yates et al., 2018). However, the collection of such data is often challenging, requiring rigorous sampling design (Elith and Leathwick, 2009), and are therefore constrained by their limited availability relative to the easily accessible presence-only databases (e.g. GBIF, 2020) and high resolution global climate information (e.g. Fick & Hijmans, 2017). Thus, caution is warranted when using SDMs to make predictions across spatial landscapes, particularly through these popular approaches.

While model performances were overall poor, I demonstrated that transferability was significantly influenced by several factors, including the number of occurrences in the exotic range, the taxonomic class and the invaded continent. Several studies have cautioned against extrapolating beyond the conditions used to fit the models, as SDMs may fail to capture these unobserved species-environment relationships (Werkowska et al., 2017; Yates et al., 2018). Environmental conditions did not differ considerably between the native and exotic ranges, with respect to the climatic variables used to fit the SDMs. Despite these differences, I observed poorer model performance across greater geographic distances and when extrapolating to islands, suggesting that other differences may exist between ranges beyond the climatic predictors, or that the species-environment relationship had changed between ranges ("nice shift"; Tingley et al., 2014). The impacts of invasive species on biodiversity loss are particularly more severe on islands globally (Spatz et al., 2017), highlighting concerns for the prediction of invaders using ecological models as a management tool.

Several predictors were contrary to their hypothesized direction, which requires further exploration. Contrary to other studies (Brotons et al., 2004; Brotons et al., 2007; Zhang et al., 2015), I found that SDMs tended to perform better with larger environmental breadths. Additionally, I found that higher exposure (the opportunity to reach suitable habitats) led to poorer model performance, where models performed better for more recent invasions and high GDP in the native range (relating to greater trade and thus greater propagule pressure). This is contrary to the assumption that species closer to equilibrium with their environment are more predictable (Václavík & Meentemeyer, 2011), however this result provides a silver lining as more recent invasions that are often of more concern to managers may be predictable.

Further examination of additional predictors should be considered, as this thesis considered only a subset of potential factors. The mobility of species did not significantly affect transferability, however given that taxonomic class was significant, other taxa-specific biological predictors may influence transferability, such as biotic interaction and dispersal ability (Urban et al., 2016), biogeography and habitat preference (Regos et al., 2019), as well as body size (Soininen & Luoto, 2014). Other functional trait-based approaches may be predictive, however such data are limited in their availability (Majekova et al., 2016). Changes in environmental correlation structures between ranges may also provide further insights on why the SDMs failed (Werkowska et al., 2017), particularly given that environmental values did not generally differ.

REFERENCES

Brotons, L., Thuiller, W., Araújo, M. B., & Hirzel, A. H. (2004). Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, 27(4), 437-448.

Brotons, L., Herrando, S., & Pla, M. (2007). Updating bird species distribution at large spatial scales: applications of habitat modelling to data from long-term monitoring programs. *Diversity and distributions*, 13(3), 276-288.

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

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(12), 4302-4315.

GBIF (2020). Global Biodiversity Information Facility. Available from https://www.gbif.org/

Goncalves, E., Herrera, We., Duarte, M., Bustamante, R. O., Lampo, M., Velasquez, G., ... & García-Rangel, S. (2014). Global invasion of *Lantana camara*: has the climatic niche been conserved across continents?. *PLoS One*, 9(10), e111468.

Majekova, M., Paal, T., Plowman, N. S., Bryndova, M., Kasari, L., Norberg, A., ... & Le Bagousse-Pinguet, Y. (2016). Evaluating functional diversity: missing trait data and the importance of species abundance structure and data transformation. *PloS one*, 11(2).

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C., & Guisan, A. (2012). Climatic niche shifts are rare among terrestrial plant invaders. *Science*, 335(6074), 1344-1348.

Regos, A., Gagne, L., Alcaraz-Segura, D., Honrado, J. P., & Domínguez, J. (2019). Effects of species traits and environmental predictors on performance and transferability of ecological niche models. *Scientific Reports*, 9(1), 1-14.

Sequeira, A. M., Bouchet, P. J., Yates, K. L., Mengersen, K., & Caley, M. J. (2018). Transferring biodiversity models for conservation: opportunities and challenges. *Methods in Ecology and Evolution*, 9(5), 1250-1264.

Soininen, J., & Luoto, M. (2014). Predictability in species distributions: a global analysis across organisms and ecosystems. *Global ecology and biogeography*, 23(11), 1264-1274.

Spatz, D. R., Zilliacus, K. M., Holmes, N. D., Butchart, S. H., Genovesi, P., Ceballos, G., ... & Croll, D. A. (2017). Globally threatened vertebrates on islands with invasive species. *Science Advances*, 3(10), e1603080.

Tingley, R., Vallinoto, M., Sequeira, F., & Kearney, M. R. (2014). Realized niche shift during a global biological invasion. *Proceedings of the National Academy of Sciences*, 111(28), 10233-10238.

Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J. B., Pe'er, G., Singer, A., ... & Gonzalez, A. (2016). Improving the forecast for biodiversity under climate change. *Science*, 353(6304), aad8466.

Václavík, T., & Meentemeyer, R. K. (2011). Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions*, 18(1), 73-83.

Werkowska, W., Márquez, A. L., Real, R., & Acevedo, P. (2017). A practical overview of transferability in species distribution modeling. *Environmental reviews*, 25(1), 127-133.

Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., ... & Dormann, C. F. (2018). Outstanding challenges in the transferability of ecological models. Trends in ecology & evolution, 33(10), 790-802.

Zhang, L., Liu, S., Sun, P., Wang, T., Wang, G., Zhang, X., & Wang, L. (2015). Consensus forecasting of species distributions: The effects of niche model performance and niche properties. *PLoS One*, 10(3), e0120056.

SUPPORTING INFORMATION

Text S1. Full list of the GBIF DOIs for the occurrence records used in the analyses, prior to filtering the data.

https://doi.org/10.15468/dl.g85fea

https://doi.org/10.15468/dl.m5ym58

https://doi.org/10.15468/dl.ujb33g

https://doi.org/10.15468/dl.k57nmd

https://doi.org/10.15468/dl.fmmw5u

https://doi.org/10.15468/dl.nrskxc

https://doi.org/10.15468/dl.s4atgj

https://doi.org/10.15468/dl.rs4mz5

https://doi.org/10.15468/dl.xvgpsn

https://doi.org/10.15468/dl.qvnt2u

https://doi.org/10.15468/dl.fjfeb2

https://doi.org/10.15468/dl.7f22vx

https://doi.org/10.15468/dl.vcdyuy

https://doi.org/10.15468/dl.bby79p

https://doi.org/10.15468/dl.n6vt57

https://doi.org/10.15468/dl.qkpvf4

REFERENCES

GBIF (2020). Global Biodiversity Information Facility Occurrence Download. Retrieved from https://www.gbif.org/

Text S2. Examination of the Continuous Boyce index (CBI) as an alternative measure of model transferability.

Several criticisms have been raised against the use of AUC as a measure of model performance (Jiménez-Valverde, 2012; Lobo et al., 2007), and that a weakness of AUC is the equal weighting of omission and commission errors. Another limitation is that AUC only provides a measure of relative probability - that is, the focus of AUC is on discrimination rather than calibration. However, we argue that this is the appropriate limit to inferences for presence-only data, because we do not know how many presences are unobserved; we only know the relative probability of occurrence in comparison with other locations.

The Continuous Boyce Index (CBI) has been commonly used as an alternative measure of SDM performance in place of AUC for presence-only modelling, as it relies solely on species presences (Hirzel et al., 2006). The Boyce index compares the model's predicted suitability of presence locations to the overall predictions of suitable areas (Boyce et al., 2002). CBI values range between -1 and 1, with positive values indicating predicted suitability consistent with species occurrences, negative values indicating low predicted suitability at sites with presences. Values that are near 0 indicate predicted suitability consistent with a random model.

One of the central issues of presence-only data is imperfect detection, where a potentially large fraction of occurrences may be unobserved (Fithian et al., 2014). Further, this is particularly important in extrapolating to new locations, where the fraction of observed presences may differ between ranges. Thus, a metric that is robust to this confounding factor would be preferable in the context of presence-only data.

We conducted theoretical analyses to test the sensitivity of the metric to the fraction of unobserved occurrences using simulations. We modeled an underlying functional relationship to generate occurrence probabilities using a simple logistic equation, serially reducing the fraction of occurrences observed at five percent intervals (i.e., simulating presence-only data), and randomly choosing 10,000 background points to fit the SDM. We generated GAMs at each fraction using the sampled species occurrences and background sites, with AUC and CBI values generated from the fitted values. We repeated the process 50 times, generating a new environment and species occurrence each time, and took the mean and standard deviation of AUC and CBI values. We calculated the CBI value using the 'ecospat.boyce' function from the 'ecospat' R package (Di Cola et al., 2017).

We found that AUC was robust to the fraction of unobserved occurrences, while CBI was sensitive to this proportion (Figure S1). The mean CBI values at 95 percent sampled had a mean value of 0.976 (standard deviation s = 0.025), and decreased as the proportion of sampled presences decreased to a mean of 0.726 (s = 0.119). Conversely, AUC remained consistent as the proportion decreased, with variance increasing marginally at very low proportions, with a mean AUC of 0.870 (s = 0.004) and 0.873 (s = 0.016) for 95 percent and 5 percent sampled presences, respectively. Therefore, we believe that AUC was preferable for our study, and applied it as the main measure of model transferability.

Nonetheless, we include the CBI analysis in the Supplementary Material, examining model transferability. Using 10-fold cross-validation, we obtained a mean CBI of 0.826 (standard deviation s = 0.122; N = 648) for GAMs, and mean CBI of 0.853 (s = 0.145) for MaxEnt models in the native range. When extrapolating to the exotic ranges, we observed a decline in performance, with a mean CBI of 0.356 (s = 0.561; N = 1918) for GAMs and mean CBI of 0.360 (s = 0.543) for MaxEnt models.

REFERENCES

Boyce, M. S., Vernier, P. R., Nielsen, S. E., & Schmiegelow, F. K. (2002). Evaluating resource selection functions. *Ecological modelling*, 157(2-3), 281-300.

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(6), 774-787.

Fithian, W., Elith, J., Hastie, T., & Keith, D. A. (2015). Bias correction in species distribution models: pooling survey and collection data for multiple species. *Methods in ecology and evolution*, 6(4), 424-438.

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(2), 142-152.

Jiménez-Valverde, A. (2012). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, 21(4), 498-507.

Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2007). AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17(2), 145-151.

Text S3. Comparison of random cross-validation and spatial-block cross-validation within the native fitting range

We examined the discriminatory performance within the native range calculated using the mean AUC from 5-fold spatial-block cross-validation (Fithian et al., 2015) in comparison with random partitioning. Spatial-block cross-validation is the technique of splitting the study side into a checkerboard, whose blocks (which include presences and background sites) that are then randomly assigned to folds. Therefore, spatial-block cross-validation has been suggested as a better reflection of potential SDM transferability than random partitioning, which may lack spatial independence due to spatial autocorrelation (Bahn & McGill, 2013). We generated the folds using the 'blockCV' package (Valavi et al., 2019), with a block size of 200×200 km^2 or 100×100 km^2 in cases of more spatially aggregated occurrences (that is, cases where 200-km blocks resulted in folds containing 0 occurrences, or more than half of the occurrences).

Mean native range AUC values generated using spatial-block cross-validation were highly correlated with random paritioning ($r = 0.958$ for GAM; $r = 0.946$ for MaxEnt), with little difference in terms of overall model discriminatory performance (Figure S2). SDMs generated using GAMs had a mean AUC of 0.786 ($s = 0.081$; $N = 648$) using spatial-block cross-validation, compared to a mean AUC of 0.809 ($s = 0.076$) using random partitioning, while SDMs fitted using MaxEnt had a mean AUC of 0.810 ($s = 0.078$) compared to 0.830 ($s = 0.070$). The transferability models with native range AUC calculated using spatial-block instead of random partitioning as a predictor did not significantly differ in terms of parameter estimate direction and magnitude, as well as overall variance explained (marginal R^2 of 0.226 and 0.237 for GAM and MaxEnt respectively, and a conditional R^2 of 0.323 and 0.292; Nakagawa et al., 2017).

REFERENCES

Bahn, V., & McGill, B. J. (2013). Testing the predictive performance of distribution models. *Oikos*, 122(3), 321-331.

Fithian, W., Elith, J., Hastie, T., & Keith, D. A. (2015). Bias correction in species distribution models: pooling survey and collection data for multiple species. *Methods in ecology and evolution*, 6(4), 424-438.

Nakagawa, S., Johnson, P. C., & Schielzeth, H. (2017). The coefficient of determination R 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, 14(134), 20170213.

Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2019). block CV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods in Ecology and Evolution*, 10(2), 225-232.

Text S4. Description on the methodology used to generate environmental breadth and environmental coverage of a species.

To generate the measure of environmental breadth and environmental coverage, we first generated a 2-dimensional environment space using the first two axes of a principal component analysis (PCA; Pearson, 1901) with the 9 environmental predictors included in the SDM fitting using 10 000 global reference points (PC1 = 37.969 % deviance explained; PC2 = 23.753 % deviance explained). We then projected the occurrence and background environment cells into this defined two-dimensional environmental space and drew convex hulls around the points (Supporting Information Figure S3). Projecting the species occurrences into a previously defined environment space allowed comparison between species when considering their environmental breadth and coverage. We measured the environmental breadth of a species as the area of the polygon generated by the convex hull around the occurrences in environment space. Environmental coverage was generated as the environmental breadth of a species divided by the environmental breadth of all cells (i.e. the background environmental cells and the occurrence cells).

REFERENCES

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11), 559-572.

SUPPLEMENTARY TABLES

Table S1. Description of the data sources used in the analyses, for both the species distribution modelling and the predictors of transferability.

Datum	Description	Source
Occurrence and distribution data		
Species sightings	Records of species presences (1) or as background environment using target-group background sites (0)	Global Biodiversity Information Facility (GBIF, 2020)
Range information	Distributional information of species at the country-level, classifying species within a region as native or exotic.	CABI Invasive Species Compendium (CABI, 2020); IUCN Global Invasive Species Database (GISD, 2015)
Country subdivisions	Administrative areas of all countries at level 1 (province/state/equivalent), with contiguous subdivisions within countries sharing a land border combined to form a region.	Global Administrative Areas, version 3.6 (GADM, 2018)
Environmental variables		
Bioclimatic variables	19 bioclimatic variables reduced to 7 uncorrelated variables at 2.5-arcmin resolution (bio1 = annual mean temperature, bio2 = mean of monthly maximum and minimum temperatures, bio7 = temperature annual range, bio12 = annual precipitation, bio14 = precipitation of the driest month, bio15 = precipitation seasonality, bio19 = precipitation of the coldest quarter)	WorldClim version 2 (Fick & Hijmans, 2017)
Elevation	Derived from STRM elevation data; 2.5-arcmin resolution	
Vegetation	Normalized Difference Vegetation Index, NDVI, at 0.05-degree, or 3-arcmin resolution	MODIS NDVI through the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC; Acker and Leptoukh, 2007)
Predictors of transferability		
Class	Taxonomic class of the species	'taxize' package in R (Chamberlain & Szöcs, 2013)
Islands	Binary value indicating whether administrative region is an island (1) or not (0)	Global shoreline vector and islands dataset (Sayre et al., 2018)
Year of first record	Earliest record of non-indigenous species within the invaded range	Alien Species First Records version 1.2 (Seebens et al., 2017); Global Biodiversity Information Facility (GBIF, 2020)
Gross domestic product	Sum of gridded cells of gross domestic product, GDP within regions at 5-arcmin resolution for 2015	Gridded global datasets for Gross Domestic Product and Human Development Index over 1990–2015 (Kummu et al., 2018)

Table S2. List of the 648 species included in the analysis along with their taxonomic class and the total number of occurrences classified in their native and exotic ranges.

Species	Class	Taxa	Occurrence points	
			Native	**Exotic**
Abelmoschus moschatus	Magnoliopsida	Plant	256	133
Abutilon grandifolium	Magnoliopsida	Plant	60	27
Abutilon theophrasti	Magnoliopsida	Plant	21	2895
Acacia auriculiformis	Magnoliopsida	Plant	655	140
Acacia cyclops	Magnoliopsida	Plant	379	187
Acacia dealbata	Magnoliopsida	Plant	5824	1739
Acacia decurrens	Magnoliopsida	Plant	678	42
Acacia longifolia	Magnoliopsida	Plant	3179	347
Acacia mangium	Magnoliopsida	Plant	27	127
Acacia mearnsii	Magnoliopsida	Plant	3302	706
Acacia melanoxylon	Magnoliopsida	Plant	8442	652
Acacia retinodes	Magnoliopsida	Plant	400	106
Acanthospermum australe	Magnoliopsida	Plant	330	99
Acanthospermum hispidum	Magnoliopsida	Plant	234	524
Acer negundo	Magnoliopsida	Plant	96	6156
Acer platanoides	Magnoliopsida	Plant	23682	2920
Acer pseudoplatanus	Magnoliopsida	Plant	19109	10428
Achillea millefolium	Magnoliopsida	Plant	81423	441
Acridotheres tristis	Aves	Vertebrate	22993	8499
Adenanthera pavonina	Magnoliopsida	Plant	84	26
Aegilops cylindrica	Liliopsida	Plant	46	295
Aegilops triuncialis	Liliopsida	Plant	812	111
Aeschynomene americana	Magnoliopsida	Plant	708	109
Agapanthus praecox	Liliopsida	Plant	55	635
Agave americana	Liliopsida	Plant	385	3892
Ageratina adenophora	Magnoliopsida	Plant	91	1082
Ageratum conyzoides	Magnoliopsida	Plant	808	1212
Ageratum houstonianum	Magnoliopsida	Plant	215	1059
Agropyron cristatum	Liliopsida	Plant	526	972
Agrostis capillaris	Liliopsida	Plant	47938	2356
Agrostis gigantea	Liliopsida	Plant	11436	848
Ailanthus altissima	Magnoliopsida	Plant	32	4885
Aira caryophyllea	Liliopsida	Plant	3132	2565
Akebia quinata	Magnoliopsida	Plant	374	89
Albizia julibrissin	Magnoliopsida	Plant	366	2363

Albizia lebbeck	Magnoliopsida	Plant	188	667
Alectoris chukar	Aves	Vertebrate	2566	4232
Alhagi maurorum	Magnoliopsida	Plant	105	77
Allamanda cathartica	Magnoliopsida	Plant	536	477
Alliaria petiolata	Magnoliopsida	Plant	21547	4503
Alnus glutinosa	Magnoliopsida	Plant	39562	244
Alocasia macrorrhizos	Liliopsida	Plant	43	26
Alopecurus pratensis	Liliopsida	Plant	12112	12497
Alpinia purpurata	Liliopsida	Plant	45	273
Alpinia zerumbet	Liliopsida	Plant	175	100
Alternanthera philoxeroides	Magnoliopsida	Plant	130	795
Alternanthera sessilis	Magnoliopsida	Plant	29	4307
Alysicarpus vaginalis	Magnoliopsida	Plant	492	55
Alyssum desertorum	Magnoliopsida	Plant	173	315
Amaranthus spinosus	Magnoliopsida	Plant	106	223
Ambrosia artemisiifolia	Magnoliopsida	Plant	2690	2884
Ambrosia tenuifolia	Magnoliopsida	Plant	32	48
Ambrosia trifida	Magnoliopsida	Plant	2653	200
Ammophila arenaria	Liliopsida	Plant	1661	458
Amorpha fruticosa	Magnoliopsida	Plant	27	238
Andropogon virginicus	Liliopsida	Plant	81	918
Annona cherimola	Magnoliopsida	Plant	20	24
Annona glabra	Magnoliopsida	Plant	385	63
Annona squamosa	Magnoliopsida	Plant	26	336
Anoplolepis gracilipes	Insecta	Invertebrate	110	21
Anser indicus	Aves	Vertebrate	1958	1792
Anthemis cotula	Magnoliopsida	Plant	308	1279
Anthoxanthum odoratum	Liliopsida	Plant	40614	4033
Antigonon leptopus	Magnoliopsida	Plant	757	377
Arctotheca calendula	Magnoliopsida	Plant	146	5751
Ardisia crenata	Magnoliopsida	Plant	348	266
Ardisia elliptica	Magnoliopsida	Plant	75	72
Argemone mexicana	Magnoliopsida	Plant	516	533
Aristolochia elegans	Magnoliopsida	Plant	27	155
Arivela viscosa	Magnoliopsida	Plant	59	369
Arrhenatherum elatius	Liliopsida	Plant	25048	320
Arthraxon hispidus	Liliopsida	Plant	452	206
Artocarpus altilis	Magnoliopsida	Plant	43	118
Arundo donax	Liliopsida	Plant	91	3200
Asclepias curassavica	Magnoliopsida	Plant	302	1930
Asparagus asparagoides	Liliopsida	Plant	115	2963

Asparagus officinalis	Liliopsida	Plant	3805	1288
Asparagus setaceus	Liliopsida	Plant	175	384
Asystasia gangetica	Magnoliopsida	Plant	485	132
Atriplex semibaccata	Magnoliopsida	Plant	2743	414
Austrocylindropuntia subulata	Magnoliopsida	Plant	44	418
Avena barbata	Liliopsida	Plant	1576	5145
Axonopus fissifolius	Liliopsida	Plant	259	1052
Azadirachta indica	Magnoliopsida	Plant	88	1632
Azolla filiculoides	Polypodiopsida	Plant	338	1198
Baccharis halimifolia	Magnoliopsida	Plant	709	443
Bassia hyssopifolia	Magnoliopsida	Plant	32	188
Bauhinia purpurea	Magnoliopsida	Plant	106	26
Bauhinia variegata	Magnoliopsida	Plant	144	243
Bellis perennis	Magnoliopsida	Plant	29002	1029
Berberis aquifolium	Magnoliopsida	Plant	1173	7444
Berberis darwinii	Magnoliopsida	Plant	83	106
Berberis thunbergii	Magnoliopsida	Plant	226	3703
Bidens frondosa	Magnoliopsida	Plant	1141	3554
Bidens pilosa	Magnoliopsida	Plant	1480	4234
Bixa orellana	Magnoliopsida	Plant	1005	32
Boerhavia coccinea	Magnoliopsida	Plant	648	884
Boerhavia diffusa	Magnoliopsida	Plant	180	133
Bombus terrestris	Insecta	Invertebrate	10200	123
Bothriochloa ischaemum	Liliopsida	Plant	661	416
Bothriochloa pertusa	Liliopsida	Plant	25	493
Bougainvillea spectabilis	Magnoliopsida	Plant	620	265
Brachypodium distachyon	Liliopsida	Plant	1399	805
Brachypodium sylvaticum	Liliopsida	Plant	21485	46
Branta canadensis	Aves	Vertebrate	149911	25519
Brassica nigra	Magnoliopsida	Plant	1487	541
Brassica rapa	Magnoliopsida	Plant	24	3057
Brassica tournefortii	Magnoliopsida	Plant	118	3671
Briza maxima	Liliopsida	Plant	1814	4644
Bromus diandrus var. rigidus	Liliopsida	Plant	348	53
Bromus hordeaceus	Liliopsida	Plant	19161	4599
Bromus inermis	Liliopsida	Plant	3676	1968
Bromus japonicus	Liliopsida	Plant	542	828
Bromus madritensis	Liliopsida	Plant	1543	896
Bromus rubens	Liliopsida	Plant	997	4456
Bromus secalinus	Liliopsida	Plant	365	1098
Bromus sterilis	Liliopsida	Plant	11021	297

Bromus tectorum	Liliopsida	Plant	693	4464
Broussonetia papyrifera	Magnoliopsida	Plant	396	551
Brugmansia suaveolens	Magnoliopsida	Plant	200	295
Bubulcus ibis	Aves	Vertebrate	11493	42214
Buddleja davidii	Magnoliopsida	Plant	59	1731
Buddleja madagascariensis	Magnoliopsida	Plant	96	33
Butomus umbellatus	Liliopsida	Plant	4859	312
Caesalpinia decapetala	Magnoliopsida	Plant	110	640
Caesalpinia pulcherrima	Magnoliopsida	Plant	501	863
Caladium bicolor	Liliopsida	Plant	1581	165
Calliandra houstoniana	Magnoliopsida	Plant	1238	66
Callisia fragrans	Liliopsida	Plant	70	22
Callitropsis arizonica	Pinopsida	Plant	294	110
Calluna vulgaris	Magnoliopsida	Plant	39007	111
Calopogonium mucunoides	Magnoliopsida	Plant	472	227
Calotropis procera	Magnoliopsida	Plant	1536	1334
Cananga odorata	Magnoliopsida	Plant	133	93
Canna indica	Liliopsida	Plant	1832	818
Capsella bursa-pastoris	Magnoliopsida	Plant	28137	3432
Cardamine flexuosa	Magnoliopsida	Plant	11747	609
Cardiospermum grandiflorum	Magnoliopsida	Plant	183	199
Cardiospermum halicacabum	Magnoliopsida	Plant	887	612
Carduus nutans	Magnoliopsida	Plant	4212	1778
Carduus pycnocephalus	Magnoliopsida	Plant	1294	1335
Carpobrotus edulis	Magnoliopsida	Plant	194	1416
Cassia fistula	Magnoliopsida	Plant	92	272
Cassytha filiformis	Magnoliopsida	Plant	2271	52
Casuarina cunninghamiana	Magnoliopsida	Plant	1543	73
Casuarina equisetifolia	Magnoliopsida	Plant	419	218
Catharanthus roseus	Magnoliopsida	Plant	49	2891
Cecropia schreberiana ssp. antillarum	Magnoliopsida	Plant	473	21
Cedrela odorata	Magnoliopsida	Plant	993	68
Celastrus orbiculatus	Magnoliopsida	Plant	459	1344
Cenchrus biflorus	Liliopsida	Plant	1552	51
Cenchrus macrourus	Liliopsida	Plant	123	140
Cenchrus pedicellatus	Liliopsida	Plant	1329	814
Cenchrus polystachios	Liliopsida	Plant	973	640
Centaurea debeauxii	Magnoliopsida	Plant	1737	23
Centaurea diffusa	Magnoliopsida	Plant	189	254
Centaurea melitensis	Magnoliopsida	Plant	934	2919
Centaurea solstitialis	Magnoliopsida	Plant	443	1129

Centella asiatica	Magnoliopsida	Plant	3219	42
Centratherum punctatum	Magnoliopsida	Plant	642	51
Ceratitis capitata	Insecta	Invertebrate	157	23
Cervus elaphus	Mammalia	Vertebrate	9933	190
Cervus nippon	Mammalia	Vertebrate	102	530
Chloris gayana	Liliopsida	Plant	178	1230
Chloris virgata	Liliopsida	Plant	622	936
Chondrilla juncea	Magnoliopsida	Plant	2096	1820
Christella dentata	Polypodiopsida	Plant	1116	453
Chromolaena odorata	Magnoliopsida	Plant	1467	820
Cinnamomum camphora	Magnoliopsida	Plant	191	1095
Cinnamomum verum	Magnoliopsida	Plant	29	88
Cirsium arvense	Magnoliopsida	Plant	39900	3341
Cirsium vulgare	Magnoliopsida	Plant	36995	17538
Clematis terniflora	Magnoliopsida	Plant	333	525
Clematis vitalba	Magnoliopsida	Plant	16398	558
Clerodendrum bungei	Magnoliopsida	Plant	43	66
Clerodendrum chinense	Magnoliopsida	Plant	21	93
Clidemia hirta	Magnoliopsida	Plant	1211	81
Clitoria ternatea	Magnoliopsida	Plant	85	714
Coccinella septempunctata	Insecta	Invertebrate	14363	4050
Coccinia grandis	Magnoliopsida	Plant	208	31
Cocos nucifera	Liliopsida	Plant	183	671
Coix lacryma-jobi	Liliopsida	Plant	85	375
Colocasia esculenta	Liliopsida	Plant	362	893
Columba livia	Aves	Vertebrate	35571	116223
Conium maculatum	Magnoliopsida	Plant	4452	1971
Cornus sericea	Magnoliopsida	Plant	3054	1665
Cortaderia jubata	Liliopsida	Plant	61	181
Cortaderia selloana	Liliopsida	Plant	192	2784
Corvus splendens	Aves	Vertebrate	16428	438
Corythucha ciliata	Insecta	Invertebrate	20	131
Cosmos caudatus	Magnoliopsida	Plant	135	81
Cosmos sulphureus	Magnoliopsida	Plant	449	629
Cotoneaster horizontalis	Magnoliopsida	Plant	36	2745
Crassocephalum crepidioides	Magnoliopsida	Plant	104	613
Crataegus monogyna	Magnoliopsida	Plant	45152	1919
Crotalaria retusa	Magnoliopsida	Plant	253	532
Cryptostegia grandiflora	Magnoliopsida	Plant	39	472
Cryptostegia madagascariensis	Magnoliopsida	Plant	43	94
Ctenosaura similis	Reptilia	Vertebrate	594	42

Cucumis anguria	Magnoliopsida	Plant	36	272
Cupaniopsis anacardioides	Magnoliopsida	Plant	763	38
Cuscuta epithymum	Magnoliopsida	Plant	2081	192
Cyanthillium cinereum	Magnoliopsida	Plant	2749	291
Cyathula prostrata	Magnoliopsida	Plant	210	246
Cyclachaena xanthiifolia	Magnoliopsida	Plant	175	298
Cygnus olor	Aves	Vertebrate	34689	10556
Cylindropuntia fulgida	Magnoliopsida	Plant	366	51
Cynara cardunculus	Magnoliopsida	Plant	464	820
Cynodon dactylon	Liliopsida	Plant	616	10239
Cynodon nlemfuensis	Liliopsida	Plant	47	119
Cynoglossum officinale	Magnoliopsida	Plant	2934	599
Cyperus difformis	Liliopsida	Plant	1516	151
Cyperus imbricatus	Liliopsida	Plant	138	42
Cyperus papyrus	Liliopsida	Plant	151	128
Cyperus rotundus	Liliopsida	Plant	351	354
Cyrtomium falcatum	Polypodiopsida	Plant	709	116
Cytisus scoparius	Magnoliopsida	Plant	20792	1910
Dactylis glomerata	Liliopsida	Plant	50529	4424
Dactyloctenium aegyptium	Liliopsida	Plant	1981	1498
Dalbergia sissoo	Magnoliopsida	Plant	32	50
Datura quercifolia	Magnoliopsida	Plant	88	270
Delairea odorata	Magnoliopsida	Plant	35	799
Delonix regia	Magnoliopsida	Plant	25	1252
Descurainia sophia	Magnoliopsida	Plant	3975	748
Desmodium incanum	Magnoliopsida	Plant	738	138
Dichanthium annulatum	Liliopsida	Plant	357	295
Dichrostachys cinerea	Magnoliopsida	Plant	425	929
Digitaria bicornis	Liliopsida	Plant	347	385
Digitaria ciliaris	Liliopsida	Plant	272	1649
Digitaria eriantha	Liliopsida	Plant	445	61
Digitaria sanguinalis	Liliopsida	Plant	5507	917
Dimorphotheca sinuata	Magnoliopsida	Plant	80	107
Dioscorea alata	Liliopsida	Plant	41	77
Dioscorea bulbifera	Liliopsida	Plant	744	186
Diprion similis	Insecta	Invertebrate	25	427
Dovyalis caffra	Magnoliopsida	Plant	40	378
Drymaria cordata	Magnoliopsida	Plant	237	335
Duranta erecta	Magnoliopsida	Plant	910	906
Dysphania ambrosioides	Magnoliopsida	Plant	599	1261
Echinochloa pyramidalis	Liliopsida	Plant	331	75

Echinocystis lobata	Magnoliopsida	Plant	1323	660
Echium plantagineum	Magnoliopsida	Plant	1542	7297
Ehrharta calycina	Liliopsida	Plant	199	1082
Ehrharta erecta	Liliopsida	Plant	143	1507
Elaeagnus angustifolia	Magnoliopsida	Plant	284	898
Elaeagnus umbellata	Magnoliopsida	Plant	438	2545
Elephantopus mollis	Magnoliopsida	Plant	1066	83
Eleusine indica	Liliopsida	Plant	420	2115
Elodea canadensis	Liliopsida	Plant	164	7602
Elodea nuttallii	Liliopsida	Plant	183	3961
Emilia sonchifolia	Magnoliopsida	Plant	140	797
Epilobium ciliatum	Magnoliopsida	Plant	7059	8804
Eragrostis atrovirens	Liliopsida	Plant	305	45
Eragrostis ciliaris	Liliopsida	Plant	354	514
Eragrostis japonica	Liliopsida	Plant	309	88
Eragrostis lehmanniana	Liliopsida	Plant	267	283
Eragrostis pilosa	Liliopsida	Plant	802	484
Eragrostis plana	Liliopsida	Plant	109	24
Eragrostis tenella	Liliopsida	Plant	55	273
Eragrostis unioloides	Liliopsida	Plant	130	28
Eremochloa ophiuroides	Liliopsida	Plant	23	56
Erigeron karvinskianus	Magnoliopsida	Plant	401	1304
Erinaceus europaeus	Mammalia	Vertebrate	1104	9717
Eriobotrya japonica	Magnoliopsida	Plant	50	4462
Erodium botrys	Magnoliopsida	Plant	293	1534
Erodium cicutarium	Magnoliopsida	Plant	17185	6834
Etlingera elatior	Liliopsida	Plant	63	96
Eucalyptus camaldulensis	Magnoliopsida	Plant	6279	428
Euonymus alatus	Magnoliopsida	Plant	542	1124
Euonymus fortunei	Magnoliopsida	Plant	230	552
Euonymus japonicus	Magnoliopsida	Plant	68	189
Eupatorium cannabinum	Magnoliopsida	Plant	2159	17714
Euphorbia esula	Magnoliopsida	Plant	1339	219
Euphorbia terracina	Magnoliopsida	Plant	548	551
Euphorbia tirucalli	Magnoliopsida	Plant	968	47
Fallopia convolvulus	Magnoliopsida	Plant	23	2099
Ficus microcarpa	Magnoliopsida	Plant	318	63
Ficus pumila	Magnoliopsida	Plant	128	24
Fimbristylis dichotoma	Liliopsida	Plant	146	261
Fimbristylis littoralis	Liliopsida	Plant	718	183
Flemingia strobilifera	Magnoliopsida	Plant	156	28

Foeniculum vulgare	Magnoliopsida	Plant	5334	3449
Forficula auricularia	Insecta	Invertebrate	6010	864
Gaillardia pulchella	Magnoliopsida	Plant	75	252
Galega officinalis	Magnoliopsida	Plant	857	24
Galinsoga parviflora	Magnoliopsida	Plant	411	658
Galinsoga quadriradiata	Magnoliopsida	Plant	256	6336
Gallus gallus	Aves	Vertebrate	3312	259
Gaultheria shallon	Magnoliopsida	Plant	536	130
Genetta genetta	Mammalia	Vertebrate	112	2814
Genista monspessulana	Magnoliopsida	Plant	167	1453
Gleditsia triacanthos	Magnoliopsida	Plant	21	1708
Gliricidia sepium	Magnoliopsida	Plant	607	186
Glyceria declinata	Liliopsida	Plant	3989	236
Glyceria maxima	Liliopsida	Plant	7639	147
Gomphocarpus physocarpus	Magnoliopsida	Plant	168	560
Grevillea robusta	Magnoliopsida	Plant	594	1163
Gunnera tinctoria	Magnoliopsida	Plant	96	532
Gypsophila paniculata	Magnoliopsida	Plant	262	141
Haemorhous mexicanus	Aves	Vertebrate	3891	93798
Halyomorpha halys	Insecta	Invertebrate	107	2092
Harmonia axyridis	Insecta	Invertebrate	312	16769
Hedera helix	Magnoliopsida	Plant	45707	1887
Hedychium coronarium	Liliopsida	Plant	113	462
Heliconia bihai	Liliopsida	Plant	148	40
Heliotropium curassavicum	Magnoliopsida	Plant	1148	565
Helminthotheca echioides	Magnoliopsida	Plant	3787	3404
Hemidactylus frenatus	Reptilia	Vertebrate	483	1138
Hemidactylus mabouia	Reptilia	Vertebrate	411	489
Heteropogon contortus	Liliopsida	Plant	2681	608
Holcus lanatus	Liliopsida	Plant	25497	9465
Houttuynia cordata	Magnoliopsida	Plant	398	31
Hydrilla verticillata	Liliopsida	Plant	81	229
Hydrocharis morsus-ranae	Liliopsida	Plant	3985	275
Hyoscyamus niger	Magnoliopsida	Plant	1920	223
Hyparrhenia rufa	Liliopsida	Plant	746	976
Hypericum perforatum	Magnoliopsida	Plant	34372	5912
Hypochaeris radicata	Magnoliopsida	Plant	14344	14304
Icerya purchasi	Insecta	Invertebrate	34	28
Impatiens walleriana	Magnoliopsida	Plant	25	1087
Imperata cylindrica	Liliopsida	Plant	5596	398
Indigofera hirsuta	Magnoliopsida	Plant	1010	293

Indigofera spicata	Magnoliopsida	Plant	191	106
Indigofera suffruticosa	Magnoliopsida	Plant	1040	40
Ipomoea hederifolia	Magnoliopsida	Plant	811	163
Ipomoea nil	Magnoliopsida	Plant	435	370
Ipomoea purpurea	Magnoliopsida	Plant	1247	1602
Ipomoea quamoclit	Magnoliopsida	Plant	317	580
Iris pseudacorus	Liliopsida	Plant	26144	1145
Ischaemum rugosum	Liliopsida	Plant	141	77
Jacaranda mimosifolia	Magnoliopsida	Plant	93	1396
Jacobaea vulgaris	Magnoliopsida	Plant	25043	1880
Jatropha gossypiifolia	Magnoliopsida	Plant	463	609
Juncus effusus	Liliopsida	Plant	30798	1335
Juncus ensifolius	Liliopsida	Plant	202	79
Juncus tenuis	Liliopsida	Plant	419	9369
Lachnagrostis filiformis	Liliopsida	Plant	5189	37
Lagarosiphon major	Liliopsida	Plant	29	358
Lagerstroemia indica	Magnoliopsida	Plant	175	367
Lagerstroemia speciosa	Magnoliopsida	Plant	91	32
Lamium amplexicaule	Magnoliopsida	Plant	8801	2970
Lantana camara	Magnoliopsida	Plant	3821	6552
Lawsonia inermis	Magnoliopsida	Plant	70	57
Leiothrix lutea	Aves	Vertebrate	976	103
Lemna minuta	Liliopsida	Plant	228	1950
Leonotis nepetifolia	Magnoliopsida	Plant	25	803
Leonurus japonicus	Magnoliopsida	Plant	182	94
Lepidium latifolium	Magnoliopsida	Plant	747	438
Lepidium perfoliatum	Magnoliopsida	Plant	95	414
Lepidium virginicum	Magnoliopsida	Plant	1789	1129
Leptochloa panicea ssp. mucronata	Liliopsida	Plant	365	28
Lepus europaeus	Mammalia	Vertebrate	9143	7977
Lespedeza cuneata	Magnoliopsida	Plant	726	763
Leucaena leucocephala	Magnoliopsida	Plant	747	2155
Leucanthemum vulgare	Magnoliopsida	Plant	15478	13145
Ligustrum lucidum	Magnoliopsida	Plant	63	1686
Ligustrum obtusifolium	Magnoliopsida	Plant	345	134
Ligustrum sinense	Magnoliopsida	Plant	137	2041
Lilioceris lilii	Insecta	Invertebrate	576	4500
Linaria dalmatica	Magnoliopsida	Plant	34	565
Linaria vulgaris	Magnoliopsida	Plant	29201	2030
Linepithema humile	Insecta	Invertebrate	54	428
Lolium multiflorum	Liliopsida	Plant	1093	4749

Lolium perenne	Liliopsida	Plant	25956	4881
Lolium rigidum	Liliopsida	Plant	1095	3732
Lolium temulentum	Liliopsida	Plant	87	70
Lonicera japonica	Magnoliopsida	Plant	311	4227
Lonicera morrowii	Magnoliopsida	Plant	44	671
Ludwigia peruviana	Magnoliopsida	Plant	555	274
Lupinus angustifolius	Magnoliopsida	Plant	933	216
Lupinus polyphyllus	Magnoliopsida	Plant	1435	24780
Lycium ferocissimum	Magnoliopsida	Plant	80	4502
Lygodium japonicum	Polypodiopsida	Plant	1303	737
Lygodium microphyllum	Polypodiopsida	Plant	628	84
Lythrum salicaria	Magnoliopsida	Plant	30569	2790
Macroptilium atropurpureum	Magnoliopsida	Plant	657	471
Macrothelypteris torresiana	Polypodiopsida	Plant	252	834
Malva pusilla	Magnoliopsida	Plant	229	525
Mangifera indica	Magnoliopsida	Plant	304	1275
Manilkara zapota	Magnoliopsida	Plant	588	74
Marrubium vulgare	Magnoliopsida	Plant	1971	6589
Martynia annua	Magnoliopsida	Plant	391	149
Medicago lupulina	Magnoliopsida	Plant	17089	2859
Medicago polymorpha	Magnoliopsida	Plant	1566	4700
Melaleuca quinquenervia	Magnoliopsida	Plant	910	49
Melia azedarach	Magnoliopsida	Plant	154	4907
Melicoccus bijugatus	Magnoliopsida	Plant	71	49
Melilotus officinalis	Magnoliopsida	Plant	2720	4242
Melinis minutiflora	Liliopsida	Plant	93	648
Melinis repens	Liliopsida	Plant	666	2995
Mentha pulegium	Magnoliopsida	Plant	622	713
Mesembryanthemum crystallinum	Magnoliopsida	Plant	447	1122
Microstegium vimineum	Liliopsida	Plant	338	1306
Mimosa pigra	Magnoliopsida	Plant	994	861
Mimosa pudica	Magnoliopsida	Plant	1152	709
Mirabilis jalapa	Magnoliopsida	Plant	695	4332
Miscanthus sinensis	Liliopsida	Plant	552	135
Molothrus bonariensis	Aves	Vertebrate	15885	1486
Momordica charantia	Magnoliopsida	Plant	282	1827
Moringa oleifera	Magnoliopsida	Plant	129	374
Mucuna pruriens	Magnoliopsida	Plant	174	20
Muntingia calabura	Magnoliopsida	Plant	1063	106
Murdannia nudiflora	Liliopsida	Plant	43	303
Mustela nivalis	Mammalia	Vertebrate	227	2468

Mustela vison	Mammalia	Vertebrate	1770	8108
Myiopsitta monachus	Aves	Vertebrate	7684	1529
Myocastor coypus	Mammalia	Vertebrate	193	1907
Myriophyllum spicatum	Magnoliopsida	Plant	5675	545
Mythimna unipuncta	Insecta	Invertebrate	1359	286
Nassella neesiana	Liliopsida	Plant	60	518
Nassella tenuissima	Liliopsida	Plant	28	31
Nasturtium microphyllum	Magnoliopsida	Plant	1600	139
Neonotonia wightii	Magnoliopsida	Plant	178	117
Nephrolepis cordifolia	Polypodiopsida	Plant	484	138
Nicotiana glauca	Magnoliopsida	Plant	223	3566
Nicotiana tabacum	Magnoliopsida	Plant	35	179
Nymphoides peltata	Magnoliopsida	Plant	1085	133
Ocimum gratissimum	Magnoliopsida	Plant	259	170
Oldenlandia corymbosa	Magnoliopsida	Plant	223	345
Oldenlandia lancifolia	Magnoliopsida	Plant	177	128
Ondatra zibethicus	Mammalia	Vertebrate	2119	2475
Onopordum acanthium	Magnoliopsida	Plant	3681	1152
Onopordum illyricum	Magnoliopsida	Plant	139	61
Opuntia ficus-indica	Magnoliopsida	Plant	343	1606
Orobanche minor	Magnoliopsida	Plant	1329	677
Oxyura jamaicensis	Aves	Vertebrate	38372	862
Paederia foetida	Magnoliopsida	Plant	574	32
Panicum repens	Liliopsida	Plant	155	468
Parkinsonia aculeata	Magnoliopsida	Plant	1048	2707
Parthenium hysterophorus	Magnoliopsida	Plant	1377	1068
Parthenocissus quinquefolia	Magnoliopsida	Plant	9207	2484
Paspalum conjugatum	Liliopsida	Plant	764	308
Paspalum dilatatum	Liliopsida	Plant	91	5537
Paspalum distichum	Liliopsida	Plant	498	1972
Paspalum notatum	Liliopsida	Plant	173	397
Paspalum paniculatum	Liliopsida	Plant	386	111
Paspalum scrobiculatum	Liliopsida	Plant	1826	30
Paspalum urvillei	Liliopsida	Plant	153	1215
Paspalum vaginatum	Liliopsida	Plant	507	98
Passer domesticus	Aves	Vertebrate	23909	141202
Passiflora caerulea	Magnoliopsida	Plant	370	1773
Passiflora edulis	Magnoliopsida	Plant	384	1126
Passiflora foetida	Magnoliopsida	Plant	2327	1243
Passiflora suberosa	Magnoliopsida	Plant	583	346
Passiflora tarminiana	Magnoliopsida	Plant	57	96

Pastinaca sativa	Magnoliopsida	Plant	10487	1223
Paulownia tomentosa	Magnoliopsida	Plant	40	1589
Pentas lanceolata	Magnoliopsida	Plant	93	32
Persicaria maculosa	Magnoliopsida	Plant	18051	1608
Persicaria perfoliata	Magnoliopsida	Plant	295	438
Persicaria punctata	Magnoliopsida	Plant	2167	29
Phalaris aquatica	Liliopsida	Plant	108	4361
Phalaris paradoxa	Liliopsida	Plant	186	784
Phasianus colchicus	Aves	Vertebrate	2473	76556
Pheidole megacephala	Insecta	Invertebrate	108	94
Phleum pratense	Liliopsida	Plant	29249	2177
Phormium tenax	Liliopsida	Plant	817	82
Phyllanthus urinaria	Magnoliopsida	Plant	209	408
Physalis angulata	Magnoliopsida	Plant	496	662
Physalis peruviana	Magnoliopsida	Plant	273	1459
Pinus elliottii	Pinopsida	Plant	224	177
Pinus halepensis	Pinopsida	Plant	2870	579
Pinus pinaster	Pinopsida	Plant	4827	157
Pinus radiata	Pinopsida	Plant	118	1985
Piper aduncum	Magnoliopsida	Plant	1656	34
Pithecellobium dulce	Magnoliopsida	Plant	780	277
Pittosporum tenuifolium	Magnoliopsida	Plant	770	102
Pittosporum undulatum	Magnoliopsida	Plant	2757	65
Plantago coronopus	Magnoliopsida	Plant	6082	1861
Plantago lanceolata	Magnoliopsida	Plant	42408	9822
Pluchea carolinensis	Magnoliopsida	Plant	371	39
Poa annua	Liliopsida	Plant	6272	10310
Poa compressa	Liliopsida	Plant	10132	1121
Poa nemoralis	Liliopsida	Plant	28847	245
Poa pratensis	Liliopsida	Plant	37565	3518
Polycarpon tetraphyllum	Magnoliopsida	Plant	835	1347
Polygonum aviculare ssp. depressum	Magnoliopsida	Plant	6086	714
Polypogon monspeliensis	Liliopsida	Plant	881	3259
Populus alba	Magnoliopsida	Plant	4267	660
Porphyrio porphyrio	Aves	Vertebrate	13157	209
Portulaca pilosa	Magnoliopsida	Plant	956	1266
Potamogeton crispus	Liliopsida	Plant	5751	841
Potamogeton perfoliatus	Liliopsida	Plant	6111	55
Procyon lotor	Mammalia	Vertebrate	6480	1415
Prosopis glandulosa	Magnoliopsida	Plant	2039	85
Prosopis juliflora	Magnoliopsida	Plant	385	488

Prosopis velutina	Magnoliopsida	Plant	73	94
Prunella vulgaris	Magnoliopsida	Plant	46590	2263
Prunus serotina	Magnoliopsida	Plant	6723	4071
Pseudelephantopus spicatus	Magnoliopsida	Plant	608	46
Psittacula krameri	Aves	Vertebrate	14215	1904
Pteris tripartita	Polypodiopsida	Plant	130	37
Pteris vittata	Polypodiopsida	Plant	489	212
Pueraria phaseoloides	Magnoliopsida	Plant	81	120
Pycnonotus jocosus	Aves	Vertebrate	8533	456
Pyracantha coccinea	Magnoliopsida	Plant	1782	862
Pyrus calleryana	Magnoliopsida	Plant	53	1179
Python bivittatus	Reptilia	Vertebrate	27	162
Rhamnus alaternus	Magnoliopsida	Plant	3647	210
Rhamnus cathartica	Magnoliopsida	Plant	9557	2062
Rhamnus frangula	Magnoliopsida	Plant	28367	1178
Rhaponticum repens	Magnoliopsida	Plant	42	447
Rhinocyllus conicus	Insecta	Invertebrate	267	113
Rhodomyrtus tomentosa	Magnoliopsida	Plant	143	21
Rhus typhina	Magnoliopsida	Plant	2278	1616
Rivina humilis	Magnoliopsida	Plant	1450	248
Rosa bracteata	Magnoliopsida	Plant	22	187
Rosa multiflora	Magnoliopsida	Plant	308	2476
Rosa rugosa	Magnoliopsida	Plant	51	10632
Rosmarinus officinalis	Magnoliopsida	Plant	47	1256
Rottboellia cochinchinensis	Liliopsida	Plant	563	233
Rubus ellipticus	Magnoliopsida	Plant	41	34
Rubus niveus	Magnoliopsida	Plant	59	37
Rubus parviflorus	Magnoliopsida	Plant	2366	27
Rubus plicatus	Magnoliopsida	Plant	25351	186
Rubus rosifolius	Magnoliopsida	Plant	1410	191
Rubus ulmifolius	Magnoliopsida	Plant	8642	928
Rudbeckia laciniata	Magnoliopsida	Plant	115	894
Ruellia simplex	Magnoliopsida	Plant	417	350
Rumex acetosella	Magnoliopsida	Plant	37327	6857
Rumex obtusifolius	Magnoliopsida	Plant	26583	2326
Rusa unicolor	Mammalia	Vertebrate	77	458
Russelia equisetiformis	Magnoliopsida	Plant	60	22
Saccharum ravennae	Liliopsida	Plant	363	52
Sacciolepis indica	Liliopsida	Plant	879	80
Sagina procumbens	Magnoliopsida	Plant	26939	380
Sagittaria latifolia	Liliopsida	Plant	791	123

Sagittaria platyphylla	Liliopsida	Plant	158	137
Salix babylonica	Magnoliopsida	Plant	20	1361
Salix cinerea	Magnoliopsida	Plant	26545	260
Salix fragilis	Magnoliopsida	Plant	7110	1086
Samanea saman	Magnoliopsida	Plant	220	130
Sambucus nigra	Magnoliopsida	Plant	35779	498
Sansevieria hyacinthoides	Liliopsida	Plant	111	62
Schinus terebinthifolius	Magnoliopsida	Plant	585	559
Schismus arabicus	Liliopsida	Plant	59	527
Sciurus carolinensis	Mammalia	Vertebrate	4469	7777
Sechium edule	Magnoliopsida	Plant	206	86
Senecio glomeratus	Magnoliopsida	Plant	360	30
Senecio madagascariensis	Magnoliopsida	Plant	67	1819
Senecio squalidus	Magnoliopsida	Plant	190	837
Senecio viscosus	Magnoliopsida	Plant	2860	161
Senecio vulgaris	Magnoliopsida	Plant	28568	4113
Senna alata	Magnoliopsida	Plant	312	662
Senna hirsuta	Magnoliopsida	Plant	576	193
Senna obtusifolia	Magnoliopsida	Plant	1125	1163
Senna occidentalis	Magnoliopsida	Plant	1236	2024
Senna septemtrionalis	Magnoliopsida	Plant	195	733
Senna siamea	Magnoliopsida	Plant	26	1773
Senna spectabilis	Magnoliopsida	Plant	580	794
Sesbania punicea	Magnoliopsida	Plant	78	285
Setaria palmifolia	Liliopsida	Plant	493	242
Setaria parviflora	Liliopsida	Plant	1298	1362
Setaria verticillata	Liliopsida	Plant	2830	532
Sicyos angulatus	Magnoliopsida	Plant	410	82
Sida acuta	Magnoliopsida	Plant	1357	1181
Sida linifolia	Magnoliopsida	Plant	607	186
Silene gallica	Magnoliopsida	Plant	2014	2327
Silybum marianum	Magnoliopsida	Plant	2137	4417
Sisymbrium irio	Magnoliopsida	Plant	980	2768
Solanum elaeagnifolium	Magnoliopsida	Plant	2732	1055
Solanum erianthum	Magnoliopsida	Plant	102	87
Solanum mauritianum	Magnoliopsida	Plant	336	1780
Solanum quitoense	Magnoliopsida	Plant	106	60
Solanum seaforthianum	Magnoliopsida	Plant	41	680
Solanum sisymbriifolium	Magnoliopsida	Plant	503	247
Solanum viarum	Magnoliopsida	Plant	162	147
Solenopsis invicta	Insecta	Invertebrate	54	1287

Solidago canadensis	Magnoliopsida	Plant	974	11162
Solidago gigantea	Magnoliopsida	Plant	1124	5309
Soliva sessilis	Magnoliopsida	Plant	30	800
Sonchus asper	Magnoliopsida	Plant	18336	7336
Sonchus oleraceus	Magnoliopsida	Plant	14914	18764
Sorghum halepense	Liliopsida	Plant	47	3185
Spartium junceum	Magnoliopsida	Plant	2060	629
Spathodea campanulata	Magnoliopsida	Plant	170	411
Spermacoce verticillata	Magnoliopsida	Plant	1584	355
Sphagneticola trilobata	Magnoliopsida	Plant	1018	690
Spiraea chamaedryfolia	Magnoliopsida	Plant	170	967
Spiraea japonica	Magnoliopsida	Plant	147	554
Sporobolus africanus	Liliopsida	Plant	171	1620
Stachys arvensis	Magnoliopsida	Plant	1672	1055
Stachytarpheta cayennensis	Magnoliopsida	Plant	1065	308
Stenotaphrum secundatum	Liliopsida	Plant	144	834
Streptopelia decaocto	Aves	Vertebrate	31941	103993
Sus scrofa	Mammalia	Vertebrate	4722	5521
Symphyotrichum novi-belgii	Magnoliopsida	Plant	467	2098
Syngonium podophyllum	Liliopsida	Plant	1528	115
Syzygium cumini	Magnoliopsida	Plant	46	215
Taeniatherum caput-medusae	Liliopsida	Plant	675	193
Tagetes erecta	Magnoliopsida	Plant	779	2753
Tagetes minuta	Magnoliopsida	Plant	209	245
Tamarix aphylla	Magnoliopsida	Plant	48	612
Tamarix gallica	Magnoliopsida	Plant	800	32
Tamarix parviflora	Magnoliopsida	Plant	101	81
Tamarix ramosissima	Magnoliopsida	Plant	102	1301
Tanacetum vulgare	Magnoliopsida	Plant	37184	2911
Tecoma capensis	Magnoliopsida	Plant	244	178
Tecoma stans	Magnoliopsida	Plant	1997	1033
Terminalia catappa	Magnoliopsida	Plant	231	745
Thespesia populnea	Magnoliopsida	Plant	126	45
Thevetia peruviana	Magnoliopsida	Plant	473	1440
Threskiornis aethiopicus	Aves	Vertebrate	5579	470
Thunbergia alata	Magnoliopsida	Plant	241	1350
Thunbergia fragrans	Magnoliopsida	Plant	63	223
Thunbergia grandiflora	Magnoliopsida	Plant	27	252
Tibouchina urvilleana	Magnoliopsida	Plant	248	65
Tithonia diversifolia	Magnoliopsida	Plant	360	1234
Tithonia rotundifolia	Magnoliopsida	Plant	120	232

Toona ciliata	Magnoliopsida	Plant	669	143
Toxicodendron succedaneum	Magnoliopsida	Plant	149	31
Tradescantia fluminensis	Liliopsida	Plant	204	1762
Tradescantia pallida	Liliopsida	Plant	240	620
Tradescantia spathacea	Liliopsida	Plant	299	171
Tradescantia zebrina	Liliopsida	Plant	379	879
Trapa natans	Magnoliopsida	Plant	552	123
Triadica sebifera	Magnoliopsida	Plant	166	1636
Trifolium dubium	Magnoliopsida	Plant	17131	3433
Trifolium repens	Magnoliopsida	Plant	60678	11810
Triumfetta rhomboidea	Magnoliopsida	Plant	742	325
Turbina corymbosa	Magnoliopsida	Plant	518	22
Tussilago farfara	Magnoliopsida	Plant	41605	3598
Typha latifolia	Liliopsida	Plant	26506	86
Ulex europaeus	Magnoliopsida	Plant	13078	2858
Urena lobata	Magnoliopsida	Plant	1244	427
Urochloa panicoides	Liliopsida	Plant	136	401
Urtica dioica	Magnoliopsida	Plant	57496	120
Vachellia nilotica	Magnoliopsida	Plant	947	363
Verbascum thapsus	Magnoliopsida	Plant	21821	9707
Verbena litoralis	Magnoliopsida	Plant	825	291
Verbena rigida	Magnoliopsida	Plant	42	1391
Verbesina encelioides	Magnoliopsida	Plant	1254	374
Vernicia fordii	Magnoliopsida	Plant	48	122
Vespula germanica	Insecta	Invertebrate	2024	550
Vicia villosa	Magnoliopsida	Plant	4634	2883
Vincetoxicum nigrum	Magnoliopsida	Plant	335	199
Vulpia bromoides	Liliopsida	Plant	2905	4412
Vulpia myuros	Liliopsida	Plant	4791	5982
Wasmannia auropunctata	Insecta	Invertebrate	696	72
Wisteria floribunda	Magnoliopsida	Plant	466	42
Xanthium spinosum	Magnoliopsida	Plant	89	2971
Xanthogaleruca luteola	Insecta	Invertebrate	108	127
Xanthosoma sagittifolium	Liliopsida	Plant	122	66
Youngia japonica	Magnoliopsida	Plant	719	688
Zingiber zerumbet	Liliopsida	Plant	43	23
Zinnia peruviana	Magnoliopsida	Plant	1170	295
Zizania latifolia	Liliopsida	Plant	69	26
Ziziphus mauritiana	Magnoliopsida	Plant	34	1043

Table S3. Full list of GBIF DOIs for downloading occurrence records used to sample target-group background sites, with classes and countries as ISO3 corresponding to each DOI link, semi-colon separated.

DOI	Class	ISO3
https://doi.org/10.15468/dl.scehaa	Insecta	SWE
https://doi.org/10.15468/dl.3yjkmb	Insecta	FRA
https://doi.org/10.15468/dl.bsm8hv	Insecta	GBR
https://doi.org/10.15468/dl.24awy7	Insecta	USA
https://doi.org/10.15468/dl.5bks73	Insecta	CAN
https://doi.org/10.15468/dl.q57a8f	Insecta	DEU
https://doi.org/10.15468/dl.93aa9t	Insecta	ITA
https://doi.org/10.15468/dl.2w5ybx	Insecta	CZE;RUS;BGR;CHL;GRC;HUN;KOR;CHN;UKR;TUR;SRB;GEO;SVK;ROU;SVN;HRV
https://doi.org/10.15468/dl.9evn73	Insecta	NLD
https://doi.org/10.15468/dl.mgppp4	Insecta	DNK
https://doi.org/10.15468/dl.9kerem	Insecta	PRT
https://doi.org/10.15468/dl.a2epjr	Insecta	ZAF;ISR;MEX;ARG;BLR;IDN;BRA;DZA;LKA;IND;MTQ;PHL;THA;TWN;GUF;ISL;PER;PRI;KAZ;IRN;CRI;DOM;HTI;COL;MYS;BRN;LTU;CUB;TTO;MKD;MOZ;AFG;BOL;MNE;KHM;MAR;PRY;ARM;MDA;GTM;PAN;JOR;LVA;CYP;SYR;VNM
https://doi.org/10.15468/dl.zqv7zc	Insecta	CHE
https://doi.org/10.15468/dl.t3retf	Insecta	BEL
https://doi.org/10.15468/dl.hg3x44	Insecta	ESP
https://doi.org/10.15468/dl.rgnegs	Insecta	AUS
https://doi.org/10.15468/dl.v22rjk	Insecta	AUT
https://doi.org/10.15468/dl.9a8f5k	Insecta	POL
https://doi.org/10.15468/dl.eeb65r	Insecta	NZL
https://doi.org/10.15468/dl.s2z5d9	Insecta	NOR
https://doi.org/10.15468/dl.ptj88b	Insecta	FIN
https://doi.org/10.15468/dl.cz5dxs	Insecta	UZB;MNG;ALB
https://doi.org/10.15468/dl.q498ed	Insecta	IRL
https://doi.org/10.15468/dl.kt7vb3	Insecta	LUX
https://doi.org/10.15468/dl.jvtb89	Insecta	JPN
https://doi.org/10.15468/dl.8hd77s	Insecta	EST
https://doi.org/10.15468/dl.smwn42	Mammalia	SWE
https://doi.org/10.15468/dl.8w892v	Mammalia	FRA
https://doi.org/10.15468/dl.t72fra	Mammalia	IMN;NZL;MEX;RUS;SRB;BGR;NLD;ZAF;CRI;TCD;AUS;LKA;TWN;PRT;LUX;GEO;BGD;SVK;DZA;BWA;SDN;BLR;HUN;NAM;SVN;HND;MNG;ROU;BTN;LTU;SLV;UKR;PAN;EST;ALB;AZE;EGY;KAZ;LVA;GTM;ETH;JEY;OMN;AND;MKD;PER;MAR;UZB;KHM;BLZ;KGZ;NIC;TUN;MNE;COL;ECU;BIH;ESH;NER
https://doi.org/10.15468/dl.gyk7hf	Mammalia	GBR
https://doi.org/10.15468/dl.ujydgk	Mammalia	USA
https://doi.org/10.15468/dl.ef9445	Mammalia	CAN
https://doi.org/10.15468/dl.akg67z	Mammalia	DEU

https://doi.org/10.15468/dl.n7yhx4	Mammalia	ARG;AUT;ITA;IND;MYS;CHL;POL;JPN;FIN;BRA;IRL;KEN;ISR;CHN;CHE;IRN;PRY;GHA;CIV;IDN;UGA
https://doi.org/10.15468/dl.qdu7dp	Mammalia	URY;THA;KOR;NPL;CZE;HRV;VNM;GRC;TUR;PAK;MMR;IRQ;ZMB;SEN;GMB
https://doi.org/10.15468/dl.hu69jv	Mammalia	DNK
https://doi.org/10.15468/dl.ym3k5p	Mammalia	ESP
https://doi.org/10.15468/dl.d4fdrv	Mammalia	BEL
https://doi.org/10.15468/dl.md73q3	Mammalia	NOR
https://doi.org/10.15468/dl.54jtyn	Pinopsida	FRA;ZAF;PRT;USA;NZL;ITA;ISR;AUS;ESP;IRN;GRC;BEL;MEX;PSE;GBR;TUR;MLT;ALB;SYR;CYP;LBN;CHE;MAR;DZA;JOR
https://doi.org/10.15468/dl.bnhhfx	Polypodiopsida	ZAF;TWN;CHN;PNG;ESP;PRT;ITA;MEX;IDN;ECU;PER;COL;MYS;PHL;MDG;BRA
https://doi.org/10.15468/dl.pwhn5q	Polypodiopsida	AUS
https://doi.org/10.15468/dl.ypgzhd	Polypodiopsida	GBR
https://doi.org/10.15468/dl.gs3vd7	Polypodiopsida	DEU;HKG;BEL;CHL;REU;ASM;URY;MYT;MAR;CYP;CAN;GRC;IRL;ISR
https://doi.org/10.15468/dl.qd957y	Polypodiopsida	USA
https://doi.org/10.15468/dl.2xb4p6	Polypodiopsida	NZL
https://doi.org/10.15468/dl.e78gvx	Polypodiopsida	FRA
https://doi.org/10.15468/dl.dw8b63	Polypodiopsida	JPN
https://doi.org/10.15468/dl.pu2hhr	Polypodiopsida	NLD
https://doi.org/10.15468/dl.rprua9	Polypodiopsida	VNM;ARG;KHM;SWZ;KOR;MOZ;CPV;GTM;NPL;TGO;ZWE;PRY;GNQ
https://doi.org/10.15468/dl.yu9s5k	Reptilia	USA
https://doi.org/10.15468/dl.tceyhv	Reptilia	PAN;ZAF;HND;CRI;TZA;CUB;PRI;TWN;COL;BRA;GTM;CHN;ECU;PER;ARG
https://doi.org/10.15468/dl.dnxzqe	Reptilia	MEX
https://doi.org/10.15468/dl.kxtujh	Reptilia	IND;AUS;BHS;KHM;NCL;PNG;JAM;REU;MMR;PHL;NPL;LAO;BGD;BTN
https://doi.org/10.15468/dl.nzjm8h	Reptilia	BLZ;THA;NIC;VNM;GUY;MOZ;SLV;PRY;STP;GHA;ZWE;SUR;NGA;BOL;VEN;ZMB;GUF;SWZ;CMR;SEN;GAB;COD
https://doi.org/10.15468/dl.mgrq9f	Aves	SWE
https://doi.org/10.15468/dl.qwadev	Aves	SWE
https://doi.org/10.15468/dl.qvhyh6	Aves	SWE
https://doi.org/10.15468/dl.789zrr	Aves	SWE
https://doi.org/10.15468/dl.tpfxv4	Aves	SWE
https://doi.org/10.15468/dl.h6ys6p	Aves	SWE
https://doi.org/10.15468/dl.eurguf	Aves	SWE
https://doi.org/10.15468/dl.8kguu5	Aves	SWE
https://doi.org/10.15468/dl.2mrcvp	Aves	SWE
https://doi.org/10.15468/dl.pgspdw	Aves	SWE
https://doi.org/10.15468/dl.rd8ns4	Aves	ZAF
https://doi.org/10.15468/dl.93tn2f	Aves	LKA;THA;MYT;BGD;MYS;REU;ISR;KAZ;MNG;ARE;SAU;NPL;MDG;LAO;KHM;NCL;OMN;VNM;MUS;MMR;IRN;UZB;AFG;KGZ;KWT;JOR;BTN;PAK;IRQ;LBN;TJK;SYR;TKM;LUX;BHS;EST
https://doi.org/10.15468/dl.3pyptv	Aves	IND
https://doi.org/10.15468/dl.yg5y25	Aves	USA
https://doi.org/10.15468/dl.r8yrh3	Aves	USA
https://doi.org/10.15468/dl.btmhg7	Aves	USA
https://doi.org/10.15468/dl.728cmq	Aves	USA

https://doi.org/10.15468/dl.58bpns	Aves	USA
https://doi.org/10.15468/dl.jy4epc	Aves	USA
https://doi.org/10.15468/dl.cquss5	Aves	USA
https://doi.org/10.15468/dl.vjwhjj	Aves	USA
https://doi.org/10.15468/dl.2setyf	Aves	USA
https://doi.org/10.15468/dl.e8bu9y	Aves	USA
https://doi.org/10.15468/dl.tk4ank	Aves	USA
https://doi.org/10.15468/dl.xmbj69	Aves	USA
https://doi.org/10.15468/dl.f3pbgh	Aves	USA
https://doi.org/10.15468/dl.93xkf3	Aves	USA
https://doi.org/10.15468/dl.e6ycq7	Aves	USA
https://doi.org/10.15468/dl.7xbgu3	Aves	USA
https://doi.org/10.15468/dl.em8usn	Aves	USA
https://doi.org/10.15468/dl.6rh8nc	Aves	USA
https://doi.org/10.15468/dl.kywuy2	Aves	AUS
https://doi.org/10.15468/dl.a7vac3	Aves	AUS
https://doi.org/10.15468/dl.cmkdgz	Aves	AUS
https://doi.org/10.15468/dl.ezyvay	Aves	AUS
https://doi.org/10.15468/dl.gqs288	Aves	AUS
https://doi.org/10.15468/dl.ucqq36	Aves	DEU
https://doi.org/10.15468/dl.zche68	Aves	NZL
https://doi.org/10.15468/dl.svnva4	Aves	RUS
https://doi.org/10.15468/dl.n4a3uc	Aves	FRA
https://doi.org/10.15468/dl.z3jdew	Aves	FRA
https://doi.org/10.15468/dl.atebj6	Aves	FRA
https://doi.org/10.15468/dl.xfzd6e	Aves	CHN
https://doi.org/10.15468/dl.tjyhnd	Aves	NLD
https://doi.org/10.15468/dl.kzfrf9	Aves	NLD
https://doi.org/10.15468/dl.t62pjb	Aves	NLD
https://doi.org/10.15468/dl.42cukm	Aves	BEL
https://doi.org/10.15468/dl.v52zzm	Aves	BEL
https://doi.org/10.15468/dl.45suvj	Aves	BEL
https://doi.org/10.15468/dl.rxyvq6	Aves	TUR
https://doi.org/10.15468/dl.vtgb78	Aves	AUT
https://doi.org/10.15468/dl.p339bx	Aves	CZE
https://doi.org/10.15468/dl.guhetd	Aves	CHE
https://doi.org/10.15468/dl.rdrnaa	Aves	GBR
https://doi.org/10.15468/dl.atawz7	Aves	GBR
https://doi.org/10.15468/dl.fn7uhe	Aves	GBR
https://doi.org/10.15468/dl.7fajg9	Aves	FIN
https://doi.org/10.15468/dl.x2m8v4	Aves	FIN
https://doi.org/10.15468/dl.mqaw7r	Aves	FIN
https://doi.org/10.15468/dl.bekcgz	Aves	UKR
https://doi.org/10.15468/dl.9q36mp	Aves	ITA
https://doi.org/10.15468/dl.bwfs8w	Aves	DNK
https://doi.org/10.15468/dl.8ppeqw	Aves	DNK
https://doi.org/10.15468/dl.ktx97h	Aves	DNK
https://doi.org/10.15468/dl.e5wcx6	Aves	DNK
https://doi.org/10.15468/dl.2nz6vd	Aves	DNK
https://doi.org/10.15468/dl.jkevau	Aves	CAN
https://doi.org/10.15468/dl.wnn2ww	Aves	CAN
https://doi.org/10.15468/dl.tuxms8	Aves	CAN
https://doi.org/10.15468/dl.u3rmh2	Aves	CAN
https://doi.org/10.15468/dl.ykxwsr	Aves	CAN

https://doi.org/10.15468/dl.p9ebfw	Aves	CAN
https://doi.org/10.15468/dl.yqnhhq	Aves	CAN
https://doi.org/10.15468/dl.npwbbz	Aves	CAN
https://doi.org/10.15468/dl.6wqz3v	Aves	CAN
https://doi.org/10.15468/dl.kr3f9c	Aves	NOR
https://doi.org/10.15468/dl.uu4c2g	Aves	NOR
https://doi.org/10.15468/dl.7795dh	Aves	NOR
https://doi.org/10.15468/dl.758btm	Aves	NOR
https://doi.org/10.15468/dl.j6ss9w	Aves	NOR
https://doi.org/10.15468/dl.m6mprq	Aves	MEX
https://doi.org/10.15468/dl.gz63j8	Aves	LVA;LTU;HRV;BLR;SVK;SVN;MNE;MKD;NGA;MDA;CMR;COD;RWA;AGO
https://doi.org/10.15468/dl.45eugh	Aves	SJM;COL;ARG
https://doi.org/10.15468/dl.txqkq7	Aves	IRL
https://doi.org/10.15468/dl.hktxwm	Aves	POL
https://doi.org/10.15468/dl.rxfhxm	Aves	PRT
https://doi.org/10.15468/dl.z8rg39	Aves	ESP
https://doi.org/10.15468/dl.zjtfjb	Aves	ESP
https://doi.org/10.15468/dl.q6fwhv	Aves	ESP
https://doi.org/10.15468/dl.7m8a7z	Aves	DZA;BRA;GUY;TUN;KOR;CUB;HND;ECU
https://doi.org/10.15468/dl.nw77ye	Aves	BGR
https://doi.org/10.15468/dl.ksfnr2	Aves	UGA
https://doi.org/10.15468/dl.9u3j6k	Aves	PAN;TWN
https://doi.org/10.15468/dl.cb8e3a	Aves	GRC
https://doi.org/10.15468/dl.zqrjn6	Aves	EGY;BLZ;PRK;NIC;MAR;BWA;CYP;ALB;PER;SEN;ZWE;BIH;SYC;NAM;CPV;BOL
https://doi.org/10.15468/dl.nd8jre	Aves	ROU
https://doi.org/10.15468/dl.dfeqa8	Aves	SRB
https://doi.org/10.15468/dl.raxcec	Aves	GMB;VEN;FRO;BEN;ISL;GTM;CHL;IMN;MLT;SLV;PRY;IDN;HKG;PRI;SGP
https://doi.org/10.15468/dl.abfpek	Aves	HUN
https://doi.org/10.15468/dl.pbj3u3	Aves	GHA
https://doi.org/10.15468/dl.f394r9	Aves	KEN
https://doi.org/10.15468/dl.mvadan	Aves	TTO;TZA;URY;DOM;GUF;SUR;PHL;PNG;SOM;VIR;BES;BFA;QAT;TCD;GNB;GEO;AZE;MRT;BHR;MLI;ETH;MWI;ZMB;MOZ;SWZ;SDN
https://doi.org/10.15468/dl.22acs9	Aves	JPN
https://doi.org/10.15468/dl.dzxx2y	Liliopsida	SWE
https://doi.org/10.15468/dl.w7wuv3	Liliopsida	FRA
https://doi.org/10.15468/dl.cxgvbt	Liliopsida	GBR
https://doi.org/10.15468/dl.edbg2x	Liliopsida	COL
https://doi.org/10.15468/dl.dbeycv	Liliopsida	USA
https://doi.org/10.15468/dl.yy698u	Liliopsida	TWN;UKR;POL;ROU;IDN;PHL;PAN;ZWE;PRI;BOL;GAB;PNG;BEN;KEN;CZE;IND;BGD;BWA;MOZ;MAR;GTM;NPL;GRC;TUR;NIC;IRL;VEN;THA;MLI;CIV;KOR;MDG;TZA;HND;CHL;ETH;VNM;BLZ;MYS;MWI;PAK;BFA;PRY;SLV;NGA;UGA;GUF;TGO;IRN;COD;AFG;ZMB;CMR;GUY;LAO;NCL;PSE;SEN;GHA
https://doi.org/10.15468/dl.zz3f6a	Liliopsida	NZL
https://doi.org/10.15468/dl.g2e5h4	Liliopsida	ZAF
https://doi.org/10.15468/dl.vdteqy	Liliopsida	MEX
https://doi.org/10.15468/dl.6ufw3e	Liliopsida	RUS
https://doi.org/10.15468/dl.5z3us7	Liliopsida	ITA;AUT;CRI;ECU;ARG;ISR

https://doi.org/10.15468/dl.9wxqbb	Liliopsida	CAN
https://doi.org/10.15468/dl.w34t73	Liliopsida	DEU
https://doi.org/10.15468/dl.994rd4	Liliopsida	LTU;BLR;CHE;LVA;REU;SVN;EST;NAM;GLP;SRB;MNE;ARE;BIH;LUX;ISL;SJM
https://doi.org/10.15468/dl.3femm5	Liliopsida	AUS
https://doi.org/10.15468/dl.s7rq2v	Liliopsida	BEL
https://doi.org/10.15468/dl.kk5ndq	Liliopsida	KAZ;KGZ;URY;DOM;IRQ;SOM;HRV;SVK;SUR;LKA;VUT;SAU;HUN;DZA;CYP;BGR;SLB;EGY;GEO;FSM;SWZ;CUB;ARM;MNG;BHS;GNQ;ALB;MMR;KHM;AZE;LBN;HTI;TUN;CPV;TJK;JOR;RWA;ATG;TTO;LSO;WSM;COG;OMN;HKG;MKD;SYR;AGO;TKM;ASM;BTN;TON;SGP;SYC
https://doi.org/10.15468/dl.zv9sdh	Liliopsida	FIN
https://doi.org/10.15468/dl.2v846r	Liliopsida	PRT
https://doi.org/10.15468/dl.kkktrv	Liliopsida	ESP
https://doi.org/10.15468/dl.j2yt6d	Liliopsida	NOR
https://doi.org/10.15468/dl.3srh2c	Liliopsida	NLD
https://doi.org/10.15468/dl.pvrern	Liliopsida	DNK
https://doi.org/10.15468/dl.8qnwzj	Liliopsida	BRA
https://doi.org/10.15468/dl.mack8m	Liliopsida	PER
https://doi.org/10.15468/dl.6kum3z	Liliopsida	CHN
https://doi.org/10.15468/dl.gn433r	Liliopsida	JPN
https://doi.org/10.15468/dl.xvwmpt	Magnoliopsida	SWE
https://doi.org/10.15468/dl.r4tqxx	Magnoliopsida	DNK
https://doi.org/10.15468/dl.sqfuqj	Magnoliopsida	FRA
https://doi.org/10.15468/dl.4kjd9n	Magnoliopsida	GBR
https://doi.org/10.15468/dl.nbx4ch	Magnoliopsida	COL
https://doi.org/10.15468/dl.tzbaxd	Magnoliopsida	NZL
https://doi.org/10.15468/dl.garrrd	Magnoliopsida	CAN
https://doi.org/10.15468/dl.u6x9j3	Magnoliopsida	USA
https://doi.org/10.15468/dl.3n5txa	Magnoliopsida	RUS
https://doi.org/10.15468/dl.2vzsbh	Magnoliopsida	CHE
https://doi.org/10.15468/dl.wrhw2h	Magnoliopsida	ITA
https://doi.org/10.15468/dl.rqmavn	Magnoliopsida	POL
https://doi.org/10.15468/dl.hdezsu	Magnoliopsida	LTU;DZA;BLR;NGA;MNG;PRI;SVK;GTM;MWI;KAZ;HND;ROU;GEO;HRV;BGR;HUN;LVA;PHL;DOM;VNM;LBR;URY;SWZ;JAM;TUN;KEN;AZE;CUB;SYR;SUR;BLZ;BWA;JOR;NPL;IRN;CYP;MLI;UGA;GNQ;ZWE;AFG;MOZ;ZMB;ARM;LBN;GUF;SLV;SEN;TGO;RWA;PYF;AGO;CAF;COG;PRK;BFA;HKG;NER;VUT;TCD;GIN;SDN;YEM
https://doi.org/10.15468/dl.dztge3	Magnoliopsida	MEX
https://doi.org/10.15468/dl.b752bn	Magnoliopsida	TWN
https://doi.org/10.15468/dl.tpcg3z	Magnoliopsida	ZAF
https://doi.org/10.15468/dl.wsqz73	Magnoliopsida	ECU
https://doi.org/10.15468/dl.kwcugz	Magnoliopsida	PRT
https://doi.org/10.15468/dl.hyb9wc	Magnoliopsida	FIN
https://doi.org/10.15468/dl.suek67	Magnoliopsida	IRL;CHL;NIC;UKR;KOR;PAN;EST;COD;TUR;SVN;MAR;PAK;PRY;NCL;THA;GHA;GUY;ETH;CIV
https://doi.org/10.15468/dl.hem47d	Magnoliopsida	ESP
https://doi.org/10.15468/dl.2922bk	Magnoliopsida	DEU
https://doi.org/10.15468/dl.em7p57	Magnoliopsida	AUS
https://doi.org/10.15468/dl.87bpsk	Magnoliopsida	MYS;IND;TZA;VEN;GAB
https://doi.org/10.15468/dl.9vh3tz	Magnoliopsida	BEL

https://doi.org/10.15468/dl.gu7kw7	Magnoliopsida	AUT
https://doi.org/10.15468/dl.2v67xv	Magnoliopsida	CZE
https://doi.org/10.15468/dl.k9e6hm	Magnoliopsida	DMA;KNA;CYM;SOM;SRB;MNE;ALB;LAO;MKD;SLB;TLS;VIR;OMN;TTO;LKA;GMB;IRQ;HTI;GRD;BHS;BTN;KHM;FJI;GUM;BGD;MUS;MMR;LCA;VGB;EGY;BRB;SAU;LSO;TJK;SGP;UZB;SYC;MLT
https://doi.org/10.15468/dl.6h9t4a	Magnoliopsida	BEN
https://doi.org/10.15468/dl.83ys52	Magnoliopsida	IDN
https://doi.org/10.15468/dl.nqw2sc	Magnoliopsida	ISR
https://doi.org/10.15468/dl.qtkqdq	Magnoliopsida	PER
https://doi.org/10.15468/dl.ja5kgm	Magnoliopsida	NOR
https://doi.org/10.15468/dl.xt392e	Magnoliopsida	ARG
https://doi.org/10.15468/dl.v7u6jz	Magnoliopsida	BRA
https://doi.org/10.15468/dl.29q7pe	Magnoliopsida	NLD
https://doi.org/10.15468/dl.e6ctmd	Magnoliopsida	CHN
https://doi.org/10.15468/dl.8dtqzv	Magnoliopsida	GRC
https://doi.org/10.15468/dl.3z9dyn	Magnoliopsida	BIH;REU;GLP;KGZ;ATG;NAM;MTQ;MDA;BES;AND;ARE;LIE;FRO;PSE;SJM;KWT;QAT;SMR;MSR;LBY
https://doi.org/10.15468/dl.ew6ckc	Magnoliopsida	BOL
https://doi.org/10.15468/dl.f8a3qr	Magnoliopsida	LUX
https://doi.org/10.15468/dl.f7x8q4	Magnoliopsida	CRI
https://doi.org/10.15468/dl.z2nkxw	Magnoliopsida	PNG
https://doi.org/10.15468/dl.44x4z2	Magnoliopsida	JPN
https://doi.org/10.15468/dl.wx6zv9	Magnoliopsida	CMR
https://doi.org/10.15468/dl.fud689	Magnoliopsida	ISL
https://doi.org/10.15468/dl.x7h8sq	Magnoliopsida	MDG
https://doi.org/10.15468/dl.s6g489	Insecta	KEN;NGA;BEN;PRK;BLZ;LCA;VCT;VIR;GRD;DMA;BRB;BES;JAM;CUW;GLP;KNA;LIE;HND;TZA;BIH;MMR;SEN;NIC;GAB;VEN;GUY;ZWE;NAM;CMR;COD;TGO;CIV;ZMB;GIN;BWA;SDN;ETH;MWI;BDI;UGA;URY;SUR
https://doi.org/10.15468/dl.pxefdy	Mammalia	BEN;CYP;PHL;SGP;ARM;GIN;VEN
https://doi.org/10.15468/dl.z6aq22	Pinopsida	BHS;BRA;BLZ;COL;NIC;HND;PAN;TUN;BOL;CUB;SVN;GUY;MNE;NLD;CRI;LBY;TCA
https://doi.org/10.15468/dl.5m5trk	Polypodiopsida	TZA;GAB;CXR;CUB;SLB;BOL;CRI;ETH;CIV;VEN;PAN;NIC;COD;NCL;SGP;NIU;GUM;UGA;KEN;BEN;CMR;ZMB;GHA;COG;LBR;AGO;GIN;DOM;THA;GUF;BRN;NGA;COM;BDI;LAO;BLZ;YEM;MMR;BFA;PLW;MUS;MWI;TON;LKA;FJI;PYF;WSM;RWA;VUT;MNP;SLV;IND;CAF;HND;LSO;NAM;SEN
https://doi.org/10.15468/dl.b5s6vb	Reptilia	GIN;NAM;MWI;BEN;RWA
https://doi.org/10.15468/dl.47gmcu	Aves	FJI;PYF;VUT;PSE;WSM;ARM;GRL;SPM;TCA;GLP;HTI;CRI;BRB;JAM;LCA;MTQ;CUW;STP;LIE;LSO;GUM;LBY;CIV;GAB;GIN;ESH;LBR;ERI;GGY;TGO;COK;BRN;AND;TLS;YEM;VCT;MDV;KNA;VGB;GRD;MAC;AIA;ATG;DJI;FLK;TON;SLB;NER;BDI;SLE;CAF
https://doi.org/10.15468/dl.mx6wgh	Liliopsida	MTQ;CXR;GIN;LBR;NER;TCD;GMB;BDI;CAF;DJI;MRT;SDN;LBY;YEM;KWT;DMA;UZB;PYF;SLE;COM;MLT;MDA;MUS;PRK;GNB;LIE;TLS;JAM;NIU;WLF;AND

https://doi.org/10.15468/dl.pqhuk3	Magnoliopsida	FSM;BRN;CUW;BDI;COK;CPV;MRT;MNP;ABW;STP;SPM;TON;TCA;GNB;CXR;COM;DJI;NRU;PLW;SLE;ERI;MAC;TKM

Table S4. Pairwise correlation table between all continuous and binary predictors included in the transferability model for GAMs.

	NAUC	**NOCC**	**EOCC**	**NEB**	**EEB**	**NEC**	**EEC**	**NGDP**	**EGDP**	**YEAR**	**DIST**	**ESIM**	**NISL**	**EISL**	**SESS**
NAUC	NA	-0.013	-0.069	-0.287	-0.188	-0.360	-0.170	0.052	-0.022	0.021	0.033	-0.047	-0.090	0.064	0.141
NOCC	-0.013	NA	0.185	0.511	0.010	0.574	0.166	0.486	-0.143	-0.123	0.075	-0.292	-0.144	0.194	-0.148
EOCC	-0.069	0.185	NA	0.102	0.436	0.155	0.575	0.148	0.295	-0.301	-0.131	-0.068	-0.007	-0.198	-0.232
NEB	-0.287	0.511	0.102	NA	0.167	0.556	0.221	0.528	-0.009	-0.047	0.100	-0.402	-0.194	-0.007	-0.165
EEB	-0.188	0.010	0.436	0.167	NA	0.087	0.501	0.009	0.456	-0.120	0.077	-0.005	0.037	-0.463	-0.012
NEC	-0.360	0.574	0.155	0.556	0.087	NA	0.221	0.111	-0.059	-0.082	-0.020	-0.051	0.024	0.091	-0.215
EEC	-0.170	0.166	0.575	0.221	0.501	0.221	NA	0.090	0.032	-0.234	-0.134	-0.150	0.010	0.026	-0.205
NGDP	0.052	0.486	0.148	0.528	0.009	0.111	0.090	NA	-0.025	-0.071	0.104	-0.301	-0.387	0.097	-0.102
EGDP	-0.022	-0.143	0.295	-0.009	0.456	-0.059	0.032	-0.025	NA	-0.173	-0.276	0.132	0.036	-0.566	-0.078
YEAR	0.021	-0.123	-0.301	-0.047	-0.120	-0.082	-0.234	-0.071	-0.173	NA	0.192	0.041	-0.006	0.042	0.083
DIST	0.033	0.075	-0.131	0.100	0.077	-0.020	-0.134	0.104	-0.276	0.192	NA	-0.090	-0.054	0.159	0.158
ESIM	-0.047	-0.292	-0.068	-0.402	-0.005	-0.051	-0.150	-0.301	0.132	0.041	-0.090	NA	0.199	-0.047	0.036
NISL	-0.090	-0.144	-0.007	-0.194	0.037	0.024	0.010	-0.387	0.036	-0.006	-0.054	0.199	NA	-0.028	0.036
EISL	0.064	0.194	-0.198	-0.007	-0.463	0.091	0.026	0.097	-0.566	0.042	0.159	-0.047	-0.028	NA	0.025
SESS	0.141	-0.148	-0.232	-0.165	-0.012	-0.215	-0.205	-0.102	-0.078	0.083	0.158	0.036	0.036	0.025	NA

Table S5. Summary table for models calculated using species mobility (SESS) as a binary predictor of SDM performance, in place of taxonomic class (CLASS). Parameter estimates, standard error, t-statistic and *P*-values are provided.

	GAM				**MaxEnt**			
Variable	**Estimate**	**Std. error**	***t*-statistic**	***P*-value**	**Estimate**	**Std. error**	***t*-statistic**	***P*-value**
Native AUC (NAUC)	0.086	0.019	4.489	< 0.001	0.058	0.018	3.240	0.001
Native number of occurrences (NOCC)	-0.031	0.026	-1.185	0.238	-0.005	0.024	-0.226	0.822
Exotic number of occurrences (EOCC)	0.048	0.022	2.149	0.032	0.037	0.022	1.680	0.094
Continent (CONT)	-	-	-	< 0.001	-	-	-	< 0.001
Sessile (SESS)	0.121	0.070	1.720	0.086	0.115	0.066	1.736	0.083
Native environmental breadth (NEB)	0.155	0.028	5.641	< 0.001	0.184	0.027	6.800	< 0.001
Exotic environmental breadth (EEB)	-0.114	0.032	-3.592	< 0.001	-0.128	0.032	-4.035	< 0.001
Native environmental coverage (NEC)	0.011	0.027	0.411	0.683	-0.010	0.026	-0.381	0.704
Exotic environmental coverage (EEC)	-0.154	0.027	-5.685	< 0.001	-0.142	0.027	-5.290	< 0.001
Native GDP (NGDP)	-0.074	0.025	-3.000	0.003	-0.055	0.023	-2.391	0.017
Exotic GDP (EGDP)	0.061	0.034	1.809	0.071	0.061	0.034	1.804	0.071
Year of first records (YEAR)	0.034	0.016	2.140	0.033	0.020	0.016	1.280	0.201
Geographic distance (DIST)	-0.054	0.018	-2.930	0.003	-0.055	0.018	-3.011	0.003
Environmental similarity (ESIM)	0.025	0.018	1.366	0.174	0.015	0.019	0.763	0.447
Native island (NISL)	-0.035	0.138	-0.252	0.801	0.036	0.129	0.282	0.778
Exotic island (EISL)	-0.271	0.056	-4.885	< 0.001	-0.306	0.056	-5.520	< 0.001

Table S6. Predictor estimates for the transferability model with GAMs and MaxEnt, including taxonomic class and continent as categorical predictors. Mean and standard deviation are provided for the scaled factors. Predictors marked with an asterisk (*) were log-transformed and predictors marked with two asterisks (**) were logit-transformed before scaling.

			GAM			**MaxEnt**		
Variable	**Mean**	**Std. dev**	**Estimate**	**Std. error**	**t-statistic**	**Estimate**	**Std. error**	**t-statistic**
(Intercept)			1.0256	0.1414	7.2530	0.9802	0.1330	7.3719
Native AUC (NAUC)**	1.6286	0.5512	0.0886	0.0193	4.5919	0.0583	0.0180	3.2373
Native number of occurrences (NOCC)*	6.6843	2.0848	-0.0277	0.0264	-1.0511	-0.0013	0.0242	-0.0551
Exotic number of occurrences (EOCC)*	5.0330	1.5050	0.0470	0.0224	2.0972	0.0359	0.0221	1.6270
Continent (CONT)								
Africa			-0.4141	0.0751	-5.5146	-0.4403	0.0745	-5.9072
Asia			-0.4997	0.0675	-7.4008	-0.5014	0.0673	-7.4537
Asia;Europe			-0.2986	0.0969	-3.0813	-0.1893	0.0958	-1.9754
Asia;Oceania			-0.7395	0.2900	-2.5498	-0.6362	0.2873	-2.2147
Europe			-0.3358	0.0675	-4.9747	-0.3543	0.0666	-5.3167
North America;South America			-0.4299	0.0837	-5.1383	-0.4430	0.0832	-5.3225
Oceania			-0.0133	0.0689	-0.1937	0.0818	0.0685	1.1940
South America			-0.4831	0.0684	-7.0602	-0.4483	0.0680	-6.5886
Taxonomic class (CLASS)								
Aves			-0.0665	0.1653	-0.4024	0.0012	0.1540	0.0078
Insecta			-0.0535	0.1789	-0.2992	0.0431	0.1684	0.2559
Liliopsida			0.0746	0.1387	0.5380	0.0863	0.1296	0.6662
Magnoliopsida			0.1317	0.1355	0.9718	0.1650	0.1266	1.3028
Pinopsida			-0.0727	0.2636	-0.2758	0.0506	0.2506	0.2018
Polypodiopsida			0.4386	0.2204	1.9903	0.6333	0.2093	3.0259
Reptilia			0.7356	0.3019	2.4364	0.4163	0.2885	1.4429
Native environmental breadth (NEB)*	2.3361	0.7345	0.1529	0.0278	5.4916	0.1819	0.0274	6.6316
Exotic environmental breadth (EEB)*	1.3130	1.0438	-0.1081	0.0317	-3.4136	-0.1215	0.0317	-3.8352
Native environmental coverage (NEC)	0.4412	0.1754	0.0114	0.0271	0.4189	-0.0123	0.0254	-0.4841
Exotic environmental coverage (EEC)	0.3377	0.1806	-0.1532	0.0270	-5.6702	-0.1421	0.0269	-5.2843
Native GDP (NGDP)*	29.8166	1.3399	-0.0735	0.0247	-2.9749	-0.0559	0.0229	-2.4349
Exotic GDP (EGDP)*	28.3193	1.8381	0.0591	0.0339	1.7412	0.0580	0.0336	1.7264
Year of first records (YEAR)	1897.1674	78.6150	0.0347	0.0162	2.1464	0.0199	0.0158	1.2579
Geographic distance (DIST)	10898.6147	4590.9176	-0.0549	0.0183	-3.0062	-0.0565	0.0181	-3.1226
Environmental similarity (ESIM)	0.1563	0.2080	0.0233	0.0185	1.2571	0.0126	0.0192	0.6558
Native island (NISL)			-0.0352	0.1366	-0.2579	0.0321	0.1275	0.2521
Exotic island (EISL)			-0.2703	0.0556	-4.8617	-0.3024	0.0556	-5.4412

Table S7. Mean predictive performance of GAM and MaxEnt species distribution models for each country in the study, with standard deviation and number of invading species.

Country	**ISO3**	**Mean AUC (GAM)**	**Standard deviation (GAM)**	**Mean AUC (MaxEnt)**	**Standard deviation (MaxEnt)**	**Number of species**
Albania	ALB	0.699	0.162	0.746	0.130	5
Algeria	DZA	0.596	0.084	0.638	0.114	4
Angola	AGO	0.612	0.081	0.618	0.097	12
Argentina	ARG	0.640	0.099	0.633	0.098	120
Armenia	ARM	0.722	0.168	0.671	0.134	3
Australia	AUS	0.731	0.132	0.747	0.125	348
Austria	AUT	0.630	0.084	0.646	0.077	56
Azerbaijan	AZE	0.765	0.139	0.679	0.098	6
Bangladesh	BGD	0.649	0.079	0.623	0.065	13
Belarus	BLR	0.656	0.112	0.667	0.097	21
Belgium	BEL	0.635	0.089	0.641	0.083	62
Belize	BLZ	0.713	0.121	0.727	0.109	49
Benin	BEN	0.591	0.064	0.584	0.077	28
Bhutan	BTN	0.692	0.105	0.721	0.115	12
Bolivia	BOL	0.647	0.086	0.637	0.087	88
Bosnia and Herzegovina	BIH	0.619	0.105	0.675	0.093	6
Botswana	BWA	0.627	0.079	0.647	0.088	36
Brazil	BRA	0.628	0.092	0.634	0.094	163
Bulgaria	BGR	0.701	0.152	0.692	0.091	10
Burkina Faso	BFA	0.584	0.051	0.581	0.077	19
Burundi	BDI	0.661	0.099	0.670	0.105	9
Cambodia	KHM	0.631	0.075	0.629	0.088	23
Cameroon	CMR	0.631	0.078	0.625	0.089	28
Canada	CAN	0.715	0.116	0.694	0.117	166
Central African Republic	CAF	0.622	0.111	0.617	0.118	10
Chad	TCD	0.588	0.118	0.596	0.111	9
Chile	CHL	0.634	0.098	0.632	0.097	77
China	CHN	0.676	0.100	0.666	0.107	77
Colombia	COL	0.638	0.095	0.640	0.093	126
Costa Rica	CRI	0.632	0.087	0.633	0.091	84
Côte d'Ivoire	CIV	0.585	0.071	0.581	0.071	25
Croatia	HRV	0.630	0.079	0.676	0.080	13
Czech Republic	CZE	0.639	0.090	0.651	0.085	52

Democratic Republic of the Congo	COD	0.638	0.083	0.618	0.089	34
Djibouti	DJI	0.612	0.041	0.586	0.073	2
Ecuador	ECU	0.636	0.088	0.635	0.090	125
Egypt	EGY	0.628	0.041	0.638	0.050	3
Eritrea	ERI	0.614	NA	0.700	NA	1
Estonia	EST	0.635	0.093	0.655	0.067	19
Ethiopia	ETH	0.653	0.086	0.644	0.087	57
Finland	FIN	0.714	0.129	0.711	0.119	46
France	FRA	0.647	0.108	0.638	0.096	105
French Guiana	GUF	0.616	0.080	0.622	0.079	45
Gabon	GAB	0.616	0.069	0.613	0.088	20
Gambia	GMB	0.573	0.067	0.572	0.072	6
Georgia	GEO	0.784	0.139	0.708	0.088	8
Germany	DEU	0.631	0.086	0.641	0.086	75
Ghana	GHA	0.587	0.072	0.584	0.067	23
Greece	GRC	0.687	0.145	0.674	0.120	20
Guatemala	GTM	0.699	0.117	0.707	0.110	87
Guinea	GIN	0.588	0.066	0.582	0.075	20
Guinea-Bissau	GNB	0.550	0.054	0.554	0.062	5
Guyana	GUY	0.625	0.076	0.623	0.081	46
Hungary	HUN	0.634	0.088	0.642	0.075	29
India	IND	0.653	0.101	0.649	0.105	85
Indonesia	IDN	0.519	0.034	0.530	0.039	5
Iran	IRN	0.669	0.158	0.706	0.110	6
Iraq	IRQ	0.639	0.100	0.683	0.161	3
Ireland	IRL	0.595	0.081	0.582	0.092	15
Italy	ITA	0.645	0.100	0.639	0.091	92
Japan	JPN	0.608	0.088	0.593	0.093	43
Jordan	JOR	0.577	NA	0.686	NA	1
Kazakhstan	KAZ	0.685	0.128	0.679	0.120	7
Kenya	KEN	0.655	0.089	0.645	0.094	73
Kuwait	KWT	0.581	0.006	0.764	0.110	2
Kyrgyzstan	KGZ	0.768	0.097	0.670	0.134	3
Laos	LAO	0.644	0.077	0.652	0.082	21
Latvia	LVA	0.652	0.098	0.678	0.086	21
Lesotho	LSO	0.607	0.070	0.615	0.084	16
Liberia	LBR	0.586	0.055	0.584	0.058	14
Libya	LBY	0.845	NA	0.852	NA	1
Lithuania	LTU	0.654	0.099	0.687	0.079	20
Luxembourg	LUX	0.651	0.069	0.668	0.068	13

Madagascar	MDG	0.695	0.127	0.710	0.098	11
Malawi	MWI	0.649	0.090	0.645	0.091	59
Malaysia	MYS	0.649	0.085	0.653	0.091	28
Mali	MLI	0.566	0.067	0.567	0.065	14
Mauritania	MRT	0.571	0.058	0.579	0.068	7
Mexico	MEX	0.714	0.122	0.717	0.114	202
Moldova	MDA	0.744	0.050	0.734	0.054	2
Mongolia	MNG	0.647	0.102	0.759	0.142	3
Montenegro	MNE	0.627	0.115	0.621	0.105	2
Morocco	MAR	0.603	0.084	0.666	0.103	4
Mozambique	MOZ	0.621	0.081	0.617	0.077	33
Myanmar	MMR	0.670	0.077	0.681	0.090	21
Namibia	NAM	0.647	0.104	0.658	0.121	33
Nepal	NPL	0.658	0.079	0.669	0.096	35
Netherlands	NLD	0.625	0.078	0.642	0.082	47
New Zealand	NZL	0.661	0.113	0.670	0.109	234
Nicaragua	NIC	0.624	0.088	0.634	0.094	72
Niger	NER	0.567	0.066	0.594	0.067	12
Nigeria	NGA	0.620	0.079	0.612	0.091	34
North Korea	PRK	0.586	NA	0.887	NA	1
Norway	NOR	0.715	0.128	0.709	0.121	52
Oman	OMN	0.581	0.006	0.764	0.110	2
Pakistan	PAK	0.645	0.108	0.649	0.112	30
Panama	PAN	0.631	0.088	0.633	0.090	95
Papua New Guinea	PNG	0.519	0.034	0.530	0.039	5
Paraguay	PRY	0.629	0.084	0.637	0.088	49
Peru	PER	0.636	0.089	0.634	0.092	106
Poland	POL	0.632	0.086	0.664	0.083	39
Portugal	PRT	0.655	0.114	0.646	0.102	69
Qatar	QAT	0.581	0.006	0.764	0.110	2
Republic of Congo	COG	0.594	0.053	0.597	0.071	10
Romania	ROU	0.672	0.119	0.671	0.090	20
Russia	RUS	0.690	0.125	0.693	0.119	38
Rwanda	RWA	0.669	0.095	0.650	0.093	33
Saudi Arabia	SAU	0.639	0.100	0.683	0.161	3
Senegal	SEN	0.583	0.065	0.593	0.071	18
Serbia	SRB	0.683	0.094	0.704	0.050	5
Sierra Leone	SLE	0.575	NA	0.594	NA	1
Slovakia	SVK	0.654	0.098	0.676	0.071	14
Slovenia	SVN	0.627	0.070	0.649	0.084	19

Somalia	SOM	0.614	0.053	0.639	0.072	12
South Africa	ZAF	0.667	0.102	0.662	0.107	121
Spain	ESP	0.669	0.117	0.654	0.100	103
Sudan	SDN	0.596	0.052	0.621	0.064	6
Suriname	SUR	0.615	0.094	0.622	0.087	28
Swaziland	SWZ	0.670	0.104	0.683	0.103	33
Sweden	SWE	0.702	0.124	0.698	0.120	56
Switzerland	CHE	0.631	0.084	0.643	0.075	45
Tajikistan	TJK	0.871	NA	0.780	NA	1
Tanzania	TZA	0.655	0.092	0.647	0.095	73
Thailand	THA	0.650	0.084	0.653	0.091	34
Togo	TGO	0.598	0.057	0.585	0.068	21
Tunisia	TUN	0.653	0.086	0.685	0.171	2
Turkey	TUR	0.691	0.157	0.719	0.104	17
Uganda	UGA	0.646	0.090	0.641	0.093	53
Ukraine	UKR	0.663	0.093	0.661	0.094	32
United Arab Emirates	ARE	0.639	0.100	0.683	0.161	3
United Kingdom	GBR	0.644	0.099	0.649	0.105	90
United States	USA	0.739	0.128	0.732	0.129	402
Uruguay	URY	0.651	0.098	0.636	0.112	39
Uzbekistan	UZB	0.590	NA	0.606	NA	1
Venezuela	VEN	0.627	0.090	0.635	0.083	89
Vietnam	VNM	0.657	0.086	0.648	0.091	37
Yemen	YEM	0.755	NA	0.521	NA	1
Zambia	ZMB	0.638	0.075	0.629	0.085	47
Zimbabwe	ZWE	0.622	0.077	0.634	0.078	38

SUPPLEMENTARY FIGURES

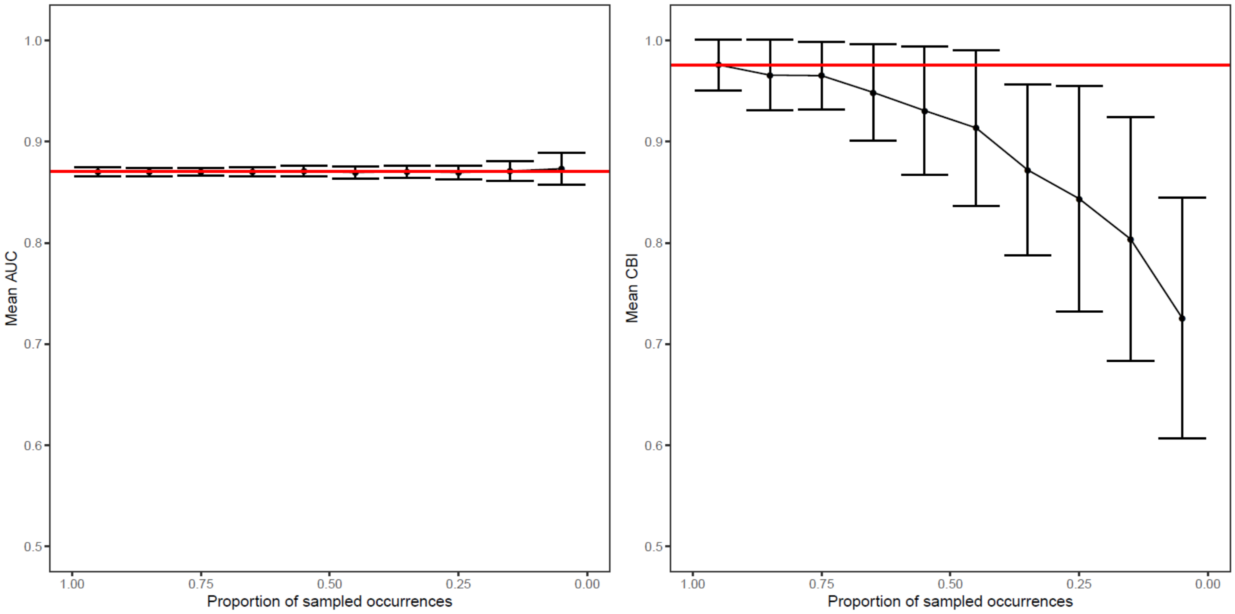


Figure S1. Change in area under the receiver operating characteristic curve (AUC) and Continuous Boyce Index (CBI) values at decreasing proportion of sampled occurrences for the simulated species-environment. Points represent the mean values at the given proportion of sampled presences, with the red line representing the mean value at 95 percent sampled presences and error bars representing the standard deviation at each fraction.

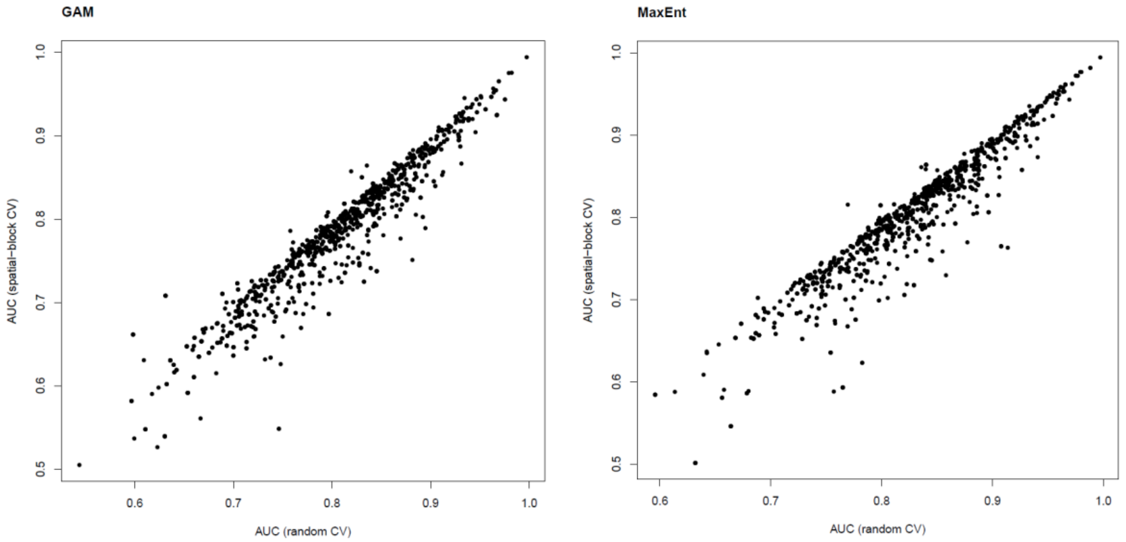


Figure S2. Comparison of native range model performance using 10-fold random cross-validation and 5-fold spatial-block cross-validation, measured as the mean area under the receiver operating characteristic curve (AUC).

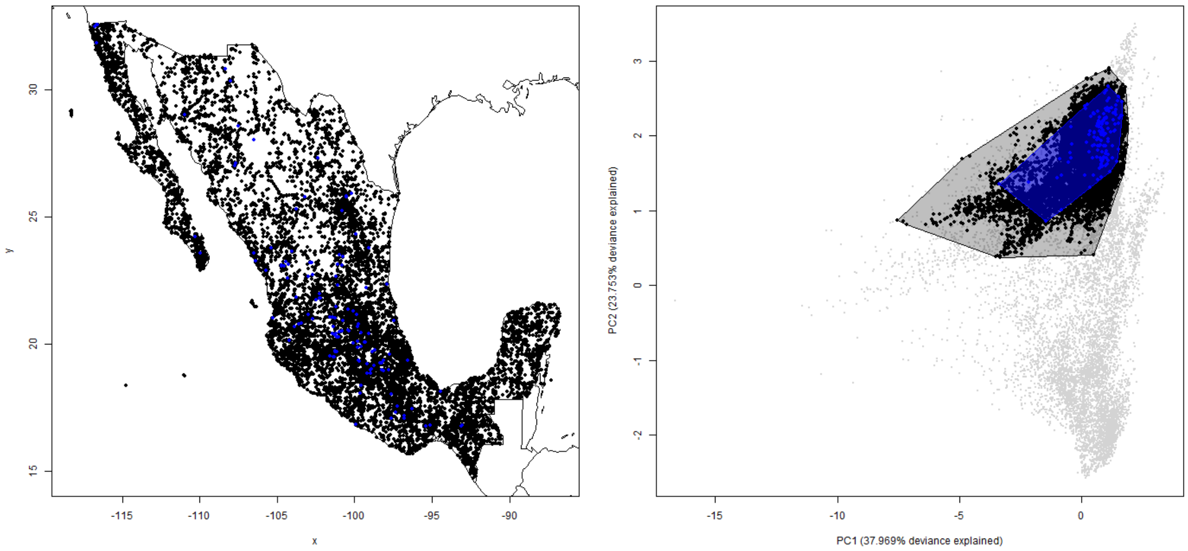


Figure S3. Example of PCA-based environmental breadth and coverage as a predictor for the prickly pear (*Opuntia ficus-indica*). The grey points represent the 10,000 global reference site points used to generate the PCA environment space for all species. The blue and black points represent the occurrences and sampled background sites in geographic and environmental space, respectively.