

ECOGRAPHY

Review

Species distribution models rarely predict the biology of real populations

Julie A. Lee-Yaw, Jenny L. McCune, Samuel Pironon and Seema N. Sheth

RUNNER-UP

Ecography E4 award

J. A. Lee-Yaw (<https://orcid.org/0000-0001-9843-8438>) ✉ (julie.lee yaw@uleth.ca) and J. L. McCune, Dept of Biological Sciences, Univ. of Lethbridge, Lethbridge, Alberta, Canada. – S. Pironon, Royal Botanic Gardens, Kew, Richmond, UK. – S. N. Sheth (<https://orcid.org/0000-0001-8284-7608>), Dept of Plant and Microbial Biology, North Carolina State Univ., Raleigh, NC, USA.

Ecography

2022: e05877

doi: 10.1111/ecog.05877

Subject Editor: Damien Fordham

Editor-in-Chief:

David Nogués-Bravo

Accepted 25 November 2021



Species distribution models (SDMs) are widely used in ecology. In theory, SDMs capture (at least part of) species' ecological niches and can be used to make inferences about the distribution of suitable habitat for species of interest. Because habitat suitability is expected to influence population demography, SDMs have been used to estimate a variety of population parameters, from occurrence to genetic diversity. However, a critical look at the ability of SDMs to predict independent data across different aspects of population biology is lacking. Here, we systematically reviewed the literature, retrieving 201 studies that tested predictions from SDMs against independent assessments of occurrence, abundance, population performance, and genetic diversity. Although there is some support for the ability of SDMs to predict occurrence (~53% of studies depending on how support was assessed), the predictive performance of these models declines progressively from occurrence to abundance, to population mean fitness, to genetic diversity. At the same time, we observed higher success among studies that evaluated performance for single versus multiple species, pointing to a possible publication bias. Thus, the limited accuracy of SDMs reported here may reflect the best-case scenario. We discuss the limitations of these models and provide specific recommendations for their use for different applications going forward. However, we emphasize that predictions from SDMs, especially when used to inform conservation decisions, should be treated as hypotheses to be tested with independent data rather than as stand-ins for the population parameters we seek to know.

Keywords: abundance, ecological niche, genetic diversity, habitat suitability, independent data, occurrence, performance, population growth

Introduction

Modelling species' distributions and the demographics of populations is a mainstay of ecology. In theory, species are expected to occur in places where conditions permit population growth rates to be stable or positive (Hutchinson 1957, Maguire 1973, Pulliam 2000, Chase and Leibold 2003). Within this set of conditions, *more*



www.ecography.org

© 2022 The Author. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

favourable environments are expected to result in a greater excess of births (e.g. greater survival and reproductive success of individuals) and thus abundance and population mean fitness are expected to scale positively with habitat suitability (Brown 1984). Large populations are less prone to the loss of alleles due to genetic drift (Frankham 1996, Leimu et al. 2006, Mccusker and Bentzen 2010) and thus habitat suitability – through effects on population size – may also influence the amount of genetic diversity within populations (Eckert et al. 2008), which may in turn feedback to influence population persistence (Agashe 2009).

One way to make predictions about population demographics is to use information about the environmental conditions that underlie species occurrence (i.e. presence or presence–absence) to model the ecological niche (or part thereof) of a species of interest (Soberón 2007, Peterson 2011, Sillero 2011, Peterson and Soberón 2012). We hereafter use the term species distribution model (SDM) to refer to this correlative approach but note that other names exist, in particular, ecological niche model (Box 1; see Elith and Leathwick 2009, Peterson 2011, Peterson and Soberón 2012 for full discussion on terminology). In some cases, the output of an SDM is treated as probability of occurrence and the spatial projection of the model is used to make predictions about the distribution of populations (Fig. 1, Box 2). In other cases, model output is treated as relative habitat suitability either because of decisions made during model construction (Peterson and Soberón 2012) or by convention (see Elith et al. 2011 for treatment of maxent output), with predictions then used to make inferences about potential differences in the demographic performance of populations across space (Fig. 1). A major appeal of these models is that they require only a set of locality records and corresponding environmental data to generate, both of which are often freely available. Thus, there has been increasing interest in using these methods to infer population parameters that are difficult to measure directly (Vaz et al. 2015).

However, there are several reasons why SDMs may fail to predict the location or demography of real populations. SDMs are often generated with coarse-scale environmental datasets (e.g. GIS layers at > 1 km resolution). On the one hand, environmental conditions at broad spatial scales may define the overall spatial limits of species' niches (Soberón 2007) and thus SDMs may broadly predict where species are or could be (Box 2). On the other hand, at a given point in space, accessibility of the site, local conditions and species interactions may also factor into whether a species is present and how well populations perform (Pearson and Dawson 2003, Peterson 2011). Thus, the scale at which many SDMs are calibrated may make them better at predicting the overall distributions of species than population parameters at a given site.

Furthermore, SDMs rely on occurrence data as input and this has two implications for their ability to predict other parameters. First, occurrence datasets may fail to adequately capture the niche of species (Loiselle et al. 2007). Specifically, dispersal can lead to populations in sink habitats (Pulliam

2000) and conversely, dispersal limitation may exclude species from large areas of suitable habitat (Svenning et al. 2008, Pagel et al. 2020). These issues limit our ability to fully estimate niches from observed occurrences. Second, the processes governing occurrence may differ from those governing other population parameters (Boulangéat et al. 2012). Thus, even when the conditions that determine presence are adequately represented in an occurrence dataset, SDMs may better predict local occurrence than other aspects of population biology (i.e. the full demographic niche of species). This issue may be especially pronounced when the parameter in question depends on other parameters that are themselves only loosely linked to occurrence (e.g. genetic diversity is theoretically linked to abundance, which at some scales, is the sum of occurrence; Fig. 1).

Regardless of their intended application, model evaluation is a critical step when developing an SDM. Many studies employ some form of cross-validation or internal model evaluation, whereby one or more subsets of the locality data used to build the model are withheld and used to assess model performance using a variety of metrics (Guisan and Zimmermann 2000). Apart from well-known issues with such internal validation (Chatfield 1995, Elith and Burgman 2002, Vaughan and Ormerod 2005, Elith et al. 2006, Lobo et al. 2008, Roberts et al. 2017), the need to evaluate SDMs with independent data is often overlooked (Elith and Burgman 2002). In the absence of independent validation, it is difficult to fully assess the utility of SDMs for informing relative differences in population demography across space.

Towards improving our understanding of the accuracy of SDMs, we compile and summarize results from studies that have compared predictions from SDMs to independent data on occurrence, population abundance, population mean fitness and performance, and genetic diversity. We assess how common it is for studies to find the expected relationship between SDM predictions and independent data for each parameter in turn (Fig. 1) and discuss specific limitations and recommendations for the use of these models for each purpose. By consolidating information from across the different applications of these models, we gain greater insight into the overall utility and limitations of SDMs in ecology and conservation biology.

Literature survey

To assess support for the use of SDMs to predict different aspects of population biology, we searched Web of Science for studies testing the relationship between SDM predictions and independent data on occurrence, abundance, population mean fitness and performance, and genetic diversity. We restricted our search to studies published after 1980 (i.e. after the advent of modern, statistical SDM methods). To reduce the effects of model extrapolation and transfer, we focused on species in their native range and removed studies where predictions were made for very different regions or time periods. We removed studies examining recent range dynamics in

Box 1. Glossary of terms

Abundance: The number of individuals of a species present at a site.

AUC: Area under the receiver operating characteristic (ROC) curve. A threshold-independent measure of model discrimination ability. Ranges from 0 to 1, where values of 0.5 indicate a model no better than random at predicting presences versus absences.

Calibration: ‘Model calibration’ often refers to internal validation of SDMs during model generation (see cross-validation). With independent data, ‘calibration performance’ refers to the degree of correlation between SDM predictions and probability of occurrence quantified during independent surveys, e.g. whether sites with predicted probabilities of 0.60 have a 60% chance of being occupied and whether this is twice as likely as for sites with predicted probabilities of 0.30 (Vaughan and Ormerod 2005).

SDM: Used here to refer to both ENMs and SDMs, both of which are correlative in nature.

SDM prediction: The output of a SDM for each cell in the study region. Depending on model algorithm, SDM output may represent probability of occurrence or habitat suitability but in many cases is either transformed or interpreted as (relative) habitat suitability.

Commission error: Predicting a species to be present at location where it is in fact absent (false positive).

Cross-validation: Testing the ability of a SDM to predict presences and absences of a species using a withheld portion of the locality data used to build the model.

Discrimination ability: The ability of a SDM to correctly distinguish observed presences and absences of a species across the study region.

Dispersal limitation: The absence of individuals from locations due to physiological or physical barriers to movement or due to insufficient time for colonization.

Distance to niche centroid (as per Lira-Noriega and Manthey 2014): a measure calculated by extracting environmental conditions underlying all sites from within the predicted geographic range of species (i.e. based on SDM predictions); estimating the multivariate centroid of this environmental space; and calculating the geometric distance between a location of interest and this centroid in environmental space. Note: in other studies, environmental data is initially extracted from species observations or other estimates of the range (i.e. not based on SDM predictions).

Ecological niche: Many definitions; see Chase and Leopold (2003) for review, and Soberón (2007) for discussion in relation to SDMs specifically. Many parameters in this review are concerned with the demographic niche – environmental conditions that support non-negative intrinsic population growth.

Finite rate of increase (λ): Per capita growth rate for a population, measured as the ratio of population size from one time interval to that of a preceding time interval.

Genetic diversity: The diversity of alleles and genotypes within a population as represented by allelic richness, nucleotide diversity, heterozygosity, inbreeding coefficients, and/or relatedness.

Habitat suitability: The ability and degree to which environmental conditions can support viable populations.

Independent data: Measurements of geographic range, occurrence, abundance, population mean fitness or genetic diversity collected independently of occurrence records used to build SDMs. Ideally does not include spatial or other biases, and are collected from sites spanning the full range of SDM predictions.

Individual fitness: Multiple definitions exist, unified by the goal of assessing the relative ability of an individual to pass its genes on to the next generation.

Intrinsic rate of increase (r): The maximum per capita growth rate for a population.

Occurrence: Presence of a given species at a given site. Occurrence datasets can include both presences and absences.

Omission errors: Predicting a species to be absent from a location where it is in fact present (false negative).

Population mean fitness (λ): Fitness averaged across all individuals in a population.

Range position: The position of a site relative to the centre or the edges of a species’ geographic distribution.

ROC: Receiver operating characteristic curve. A curve drawn by graphing sensitivity versus 1-specificity at each possible threshold of occurrence.

SDM and ENM (species distribution model and ecological niche model): Terms that have been used to describe correlative models that associate environmental variables with known occurrences of species (reviewed in Peterson 2011, Peterson and Soberón 2012). Although in theory and ideally in practice, modeling the distribution is not the same as modeling the niche (Pulliam 2000, Peterson 2011, Peterson and Soberón 2012), authors use both terms when generating models to predict the parameters covered in this review.

Sensitivity: The proportion of observed presences that were correctly predicted to be present by the SDM.

Specificity: The proportion of observed absences that were correctly predicted to be absent by the SDM.

True positive rate: The proportion of observed presences that were predicted present by the SDM.

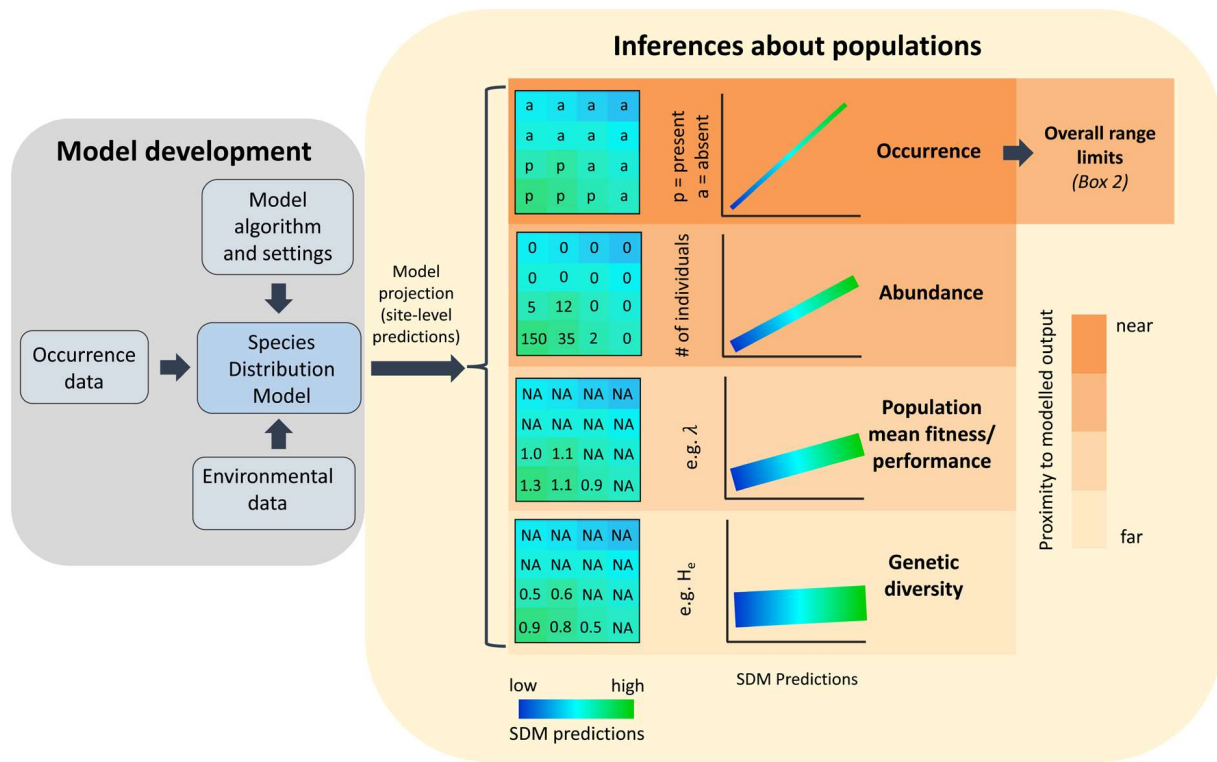


Figure 1. The use of species distribution models (SDMs) to predict different parameters in population biology. The left-hand panel shows considerations that go into SDM development (these have been reviewed previously; see main text for references). These models are used to generate site-level predictions of probability of occurrence or habitat suitability ('SDM predictions'). In theory, more suitable sites are expected to be associated with greater occurrence, abundance, population mean fitness or performance, and genetic diversity, leading to the expectation of positive relationships between SDM predictions and each of these parameters as shown in the right-hand panel. However, for reasons outlined in the main text, we expect the ability of SDMs to predict independent data to decrease as we move away from predicting occurrence (changes in strength of relationship in plots in the right-hand panel).

response to climate change, and studies that did not include independent estimates of the parameter of interest or that did not specifically examine the relationship between SDM predictions and the independent data. We supplemented our search with articles citing or cited by those retrieved, as well as with relevant studies from our own libraries. Full details of the literature search, including additional parameter-specific considerations, are provided in the Supporting information. Overall, we retrieved 1827 articles of potential relevance for our review (Supporting information). Of these, 201 met our criteria for inclusion upon examination. Although we are unlikely to have captured all relevant studies in our search, the papers reviewed here span multiple journals, disciplines, biomes, countries and taxa, and are thus expected to be generally representative of studies testing SDM predictions with independent data.

For each study, we used the results presented in figures or tables and author conclusions to determine whether SDMs predicted independent data on the population parameter of interest ('yes' or 'no', or percentage of 'yes' if more than one model and/or species was considered). We took the models and results from each study at face-value (i.e. we assumed that authors had done due diligence in parameterizing their models and making decisions appropriate for the taxa at

hand). Small sample sizes, as well as differences in the methods employed, evaluation metrics used and results presented by the different studies precluded us from obtaining comparable effect sizes across papers, even among those assessing the same type of parameter. However, our goal was not to provide a single, overall estimate of the fit of SDM predictions to data nor to explore the effects of different modeling decisions on model fit (for reviews on the latter see Hallgren et al. 2019, Elith et al. 2020). Rather, we simply ask: how often do SDMs, as presented in the literature, predict independent data on occurrence, abundance, population mean fitness and performance, and genetic diversity?

Results

Occurrence within the range

We retrieved 101 papers that evaluated SDM predictions with independent data on occurrence (Supporting information). The majority of these studies included some measure of the discriminatory ability of models – that is their ability to discriminate between occupied and unoccupied sites (Vaughan and Ormerod 2005). For instance, the area under

Box 2. Prediction and explanation of species' distributional limits

Although our review is focused on site-level predictions, we note that SDMs have been used extensively to estimate species geographic ranges and to understand the drivers of species distributions (referred to as 'prediction' and 'explanation' respectively by Elith and Leathwick 2009). Formal tests of the extent to which SDM predictions match independent estimates of the range are rare. Willems and Hill (2009) found that SDMs correctly predicted 81.8% of pixels in a reference distribution map for Vervet monkeys across Africa. Mainali et al. (2020) likewise reported close agreement between SDM predictions and expert range maps of 330 butterfly species. However, SDMs often over- or under-predict observed range limits. For example, Marcer et al. (2013) found that SDMs tended to overpredict the range polygons of seven Mediterranean plants; whereas SDMs tended to underpredict the altitudinal limits of 50% of the European tree species considered by Mellert et al. (2011). Sarquis et al. (2018) found that, despite moderate matching scores between SDMs and an expert map, SDMs consistently underpredicted the range of a well-studied viper in a particular region of Argentina (Sarquis et al. 2018). Qualitative comparisons between the distribution of occurrences and SDM prediction maps are more common and often reveal regions where SDMs over- or under-predict range limits (i.e. global commission and omission errors respectively; Duan et al. 2014, Ortega-Huerta and Vega-Rivera 2017, Peterson et al. 2018).

Several factors may influence agreement between SDM predictions and alternative estimates of species' distributions. Besides methodological considerations around the SDMs themselves (Loiselle et al. 2007, Elith and Leathwick 2009, Phillips et al. 2009, Peterson et al. 2011), erroneous and/or overly-simplistic independent estimates of observed ranges may limit our ability to validate SDM predictions (see Burgman and Fox 2003, O'Hara et al. 2017, Peterson et al. 2018, Rotenberry and Balasubramaniam 2020 for criticisms of range polygons in particular). At the same time, species' ranges reflect the culmination of a number of different processes (Gaston 2003) and many of the factors that shape species' overall distributions may not be well-captured by most SDMs. Combining SDMs with information about dispersal constraints and species interactions that may exclude target species from entire geographic areas is likely necessary to effectively estimate species' broad-scale distributions (Guisan and Thuiller 2005, Peterson 2006, Heikkinen et al. 2007, Raedig and Kreft 2011, Jetz et al. 2012, Sangermano and Eastman 2012, Domisch et al. 2016, Fourcade 2016, Merow et al. 2017, Ortega-Huerta and Vega-Rivera 2017, Calixto-Pérez 2018).

The same limitations of SDMs for estimating species' distributions may make them useful for investigating specific explanations for range limits. In theory, SDMs calibrated with climatic variables should inform the extent to which species' ranges are in equilibrium with climate. SDMs that predict suitable habitat beyond observed range limits suggest that dispersal limitation or biotic factors may be limiting the range. Several studies have used SDMs in this way (Kozak and Wiens 2006, Svenning et al. 2008, Dullinger et al. 2012, Cunningham et al. 2016, Lee-Yaw et al. 2018) but independent assessments of these inferences are rare. Lee-Yaw and colleagues (2016) compared SDM explanations of range limits to those inferred from over-the-edge transplant experiments for 40 species. For the majority of species (77.5%), SDM predictions aligned with transplant results, with both the suitability of sites and performance of transplanted individuals declining across range limits, suggesting range limits are niche limits (Lee-Yaw et al. 2016). However, the SDMs used in that study rarely predicted suitable habitat beyond range limits (i.e. rarely predicted dispersal limitation) and although the number of transplant experiments reporting dispersal limitation was also small, the SDMs failed to predict these cases. In contrast, (Bayly and Angert 2019) found that SDM predictions of dispersal limitation aligned with inferences from over-the-edge transplant experiments in monkeyflowers. Specifically, both suitability and population growth rates were high beyond observed range limits, but only when models were calibrated with microhabitat variables. These studies suggest that climate-based SDMs may not always be able to detect dispersal limitation; However, more studies comparing inferences about the drivers of range limits from SDMs with independent data are needed to better evaluate the explanatory power of these models.

the receiver operating characteristic (ROC) curve (AUC: Hosmer and Lemeshow 2000) was calculated with independent data for 53 (52%) studies. SDMs often distinguished independent presences from absences reasonably well based on this threshold-independent metric, with 34 of these studies (64%) reporting AUC > 0.70 (Swets 1988) for at least 75% of species considered. Threshold-dependent measures of discrimination ability (e.g. sensitivity, specificity, true positive rate, etc.) were also commonly reported (used in 46 studies). Here, support for the predictive performance of SDMs was more variable, with only 50% (23 out of 46) of studies

finding clear support for model predictions. Notably, support for the discriminatory ability of SDMs was lower in studies examining multiple versus single species (e.g. AUC values lower than 0.70 for a substantial proportion of species in multi-species studies; Fig. 2; Supporting information).

Although the discrimination ability of SDMs is important, testing whether there is a correlation between continuous model predictions and independent estimates of occurrence across sites is equally important for many downstream applications (Vaughan and Ormerod 2005, Gogol-Prokurat 2011, Halvorsen 2012, Lawson et al. 2014). Only 19 of the studies

we reviewed tested this ‘calibration’ (sensu Vaughan and Ormerod 2005) component of model accuracy with independent data – with most simply testing the significance of logistic regressions of SDM predictions against independent presence–absence data. Although some studies have reported significant relationships in this regard (van Manen et al. 2002, Boetsch et al. 2003, Edvardsen et al. 2011), others have found more mixed support (van Manen et al. 2005, McCune 2016); Supporting information). Gogol-Prokurat (2011) used a stricter test in rare plants that assessed not only whether there was a significant positive relationship between presence–absence and SDM predictions, but whether the relationship was linear. Only 28% of the 36 SDMs tested in that study showed probability of occurrence to be directly proportional to SDM predictions, as indicated by a significant linear fit (Gogol-Prokurat 2011). Thus, SDMs sometimes (but not always) predict the relative probability of occurrence at different sites.

Finally, a few studies have tested the ability of SDMs to predict occurrence using other methods. For example, Wauchope-Drumm et al. (2020) compared the average predicted suitability scores for independent presences versus absences for a rare marsupial, finding that mean suitability was significantly higher for independent presences than absences. Likewise, Pinto et al. (2016) found that predicted suitability was higher for observed flapper skate tracks (based on geolocator data) than for random tracks. Finally, some studies have compared the probability of finding a target species in field surveys guided by SDM predictions versus random or expert-based designs and reported greater success when relying on SDM predictions (Loyn et al. 2001, van Manen et al. 2005, Aitken et al. 2007, Aizpurua et al. 2017). Collectively then, on the basis of discriminatory ability, calibration accuracy, and other measures of performance, SDMs often predict occurrence data quite well (Fig. 2).

Limitations

Although SDMs can provide accurate predictions of species occurrences, a substantial number of studies (~38%) included at least one SDM that did not accurately predict independent data on occurrence, with ~9% of studies finding no support for SDM predictions (Supporting information; Fig. 2). A number of factors may influence the ability of SDMs to accurately predict occurrence. With respect to model generation, input occurrence datasets often represent haphazard or opportunistic records, not data from surveys carefully designed to model the full response of species to the environmental conditions that govern occurrence (Phillips et al. 2009). Furthermore, models based on data from one region may not be appropriate for making predictions in other areas (model transferability is often low; Randin et al. 2006, Boiffin et al. 2017) and thus, if occurrence data are not available for a region of interest, the applicability of these models may be limited. In addition to these limitations, models may be misled by presences in locations that are not suitable (i.e. arising from the observation of vagrant individuals; relict populations in long-lived species; or sink populations

maintained by dispersal; Van Horne 1983, Pulliam 1988, Jenkins et al. 2003) or by absences in locations that are suitable (i.e. arising from dispersal limitation, biotic interactions, environmental stochasticity; Boetsch et al. 2003, Franklin 2010, Gogol-Prokurat 2011). These same issues may plague independent evaluation datasets, especially when data from test locations are based on a single survey (i.e. presence or absence not verified over multiple surveys/years). Thus, both SDM predictions and our ability to evaluate these predictions may be influenced by the extent to which occurrence records accurately reflect species’ responses to the environmental predictors in the model.

Conclusions and recommendations

SDMs do well when it comes to predicting occurrence for many species. However, we emphasize that it is necessary to validate models with independent data before they are used for decision-making and other applications. In particular, we note that cross-validated AUC – whereby different subsets of available locality data are used for training and evaluating models – has repeatedly been found to be an unreliable indicator of model performance when compared to evaluation based on independent data (Fois et al. 2018, Gregr et al. 2019, McCune et al. 2020). Careful selection of training and testing data to minimize autocorrelation between datasets is expected to reduce this effect (Roberts et al. 2017). However, independent data collected using formal survey methods (i.e. stratified random sampling) and with the specific purpose of validating model predictions remains the most appropriate test of model performance ahead of the downstream application of such models (Elith and Burgman 2002). Furthermore, a model with reasonable discrimination ability may not accurately predict independent estimates of (continuous) probability of presence (Gogol-Prokurat 2011). Thus, SDM predictions should not be used to rank different sites (i.e. for conservation priority) without first assessing their calibration performance (Vaughan and Ormerod 2005). Finally, model parameterization (choice of predictors and modeling algorithm) can have a dramatic effect on model performance (Fois et al. 2018, McCune et al. 2020), and thus we recommend studies explore these effects with independent data before selecting a final model (or set of models) for use.

Abundance

In a previous meta-analysis based on 30 studies Weber et al. (2017), concluded that SDM predictions may serve as a reasonable proxy for abundance. Here, we provide an updated summary of results from studies examining the relationship between SDM predictions and abundance based on 20 recent studies, along with the 20 studies from Weber et al. (2017) that met our criteria and relied on occurrence-based SDMs (Supporting information). Across all studies, only 50% found support for a positive relationship between SDM predictions and independent estimates of abundance in 75% or more of species examined (Supporting information); with only 42%

finding unambiguous support for this relationship across all species examined (Fig. 2; Supporting information).

Notably, whereas most single-species studies reported a positive relationship between abundance and predicted habitat suitability, studies examining very large numbers of species (> 100) often reported little to no support for this pattern (Fig. 2; Supporting information). For example, Dallas and Hastings (2018) found that SDM predictions were often unrelated to abundance in North American mammals and tree species (positive correlations were found for only 13 out of 404 species). Likewise, Santini et al. (2019) found mostly non-significant or negative correlations between SDM predictions and range-wide patterns of abundance in a global analysis of 108 mammals and birds. Finally, Sporbert et al. (2020) found weak and varying relationships between abundance and climatic suitability from SDMs for 517 European plants. Thus, SDMs do poorly when it comes to predicting abundance for many species.

Limitations

Several factors other than habitat suitability can influence abundance at a given site including species' interactions, transient dynamics and stochasticity, metapopulation dynamics, Allee effects, and disturbance (see discussions by Stevens and Amarilla-Stevens 2012, Weber et al. 2017, Holt 2020, Osorio-Olvera et al. 2020). For this reason, VanDerWal et al. (2009) proposed that the relationship between habitat suitability and abundance is likely to be 'wedge-shaped' and that SDMs may thus be better at predicting the upper limits of abundance rather than actual abundance at a given site. They found this was true for the 69 vertebrate species they examined (VanDerWal et al. 2009) and other studies have since reported similar wedge-shaped relationships across a variety of taxa (Muñoz et al. 2015, Acevedo et al. 2017, Braz et al. 2020). Such wedge-shaped patterns suggest that the processes driving abundance at a given site are often not well-captured by SDMs. Furthermore, the lack of linear or wedge-shaped relationships between SDM predictions and abundance in the 517 plants examined by Sporbert et al. (2020) – the largest study of this kind to date – is a notable example of how SDMs may fail to predict even the upper limits of abundance for a number of species.

Conclusions and recommendations

Although SDMs can sometimes predict local abundance, predictions from these models are often a poor proxy for abundance. At the same time, the wedge-shaped relationship observed between abundance and SDM predictions for some species suggests that SDMs may be useful for narrowing in on places where there is the potential for abundance to be high (VanDerWal et al. 2009). However, even here, care is needed when using SDMs to make predictions. A few strategies may improve the predictive performance of SDMs when it comes to abundance. For instance, some studies have had success modeling abundance when including estimates of abundance (i.e. counts of occurrences in an area) as a predictor or using

a two-step approach whereby the probability of presence is modeled first (assuming a binomial distribution) and then abundance is modeled for sites deemed to be suitable (e.g. two-step modeling of zero-inflated data or hurdle model: Potts and Elith 2006, Boulangéat et al. 2012, Mellin et al. 2012). Still others have called for more complex models that incorporate demography and other processes that impact abundance (Merow et al. 2014, Ehrlén and Morris 2015, Evans et al. 2016, Zurell et al. 2016). Overall, moving beyond occurrence-based models is expected to improve model performance when the goal is to predict abundance.

Population mean fitness and performance

We found 42 studies assessing the relationship between SDM predictions and independent estimates of population mean fitness or performance (Supporting information). Of these, only five studies compared SDM predictions to independent estimates of population growth rate (i.e. population mean fitness) and thus fully assessed the ability of populations to persist at a given location. These studies, all in plants, provide little support for the use of SDMs to predict population growth rates. Specifically, in contrast to expectations, Thuiller et al. (2014) found that relationships between SDM predictions and intrinsic rate of increase (r) tended to be negative in 108 temperate trees. Similarly, a transplant study of the scarlet monkeyflower *Erythranthe cardinalis* found a negative relationship between SDM predictions and finite rate of increase (λ ; Bayly and Angert 2019). Csörgő et al. (2017) found no relationship between SDM predictions and λ in a global meta-analysis of 34 plant species. More recently, Greiser et al. (2020) reported no relationship between SDM suitability and λ in the perennial forb *Lathyrus vernus*, while (Baer and Maron 2020) reported a j-shaped relationship in the northern half of the range of the perennial forb, *Astragalus utahensis* – with λ increasing with suitability only once a threshold of suitability had been met. Finally, two studies reporting derived indices of population growth rates (but not λ or r per se) found little support for a positive relationship between SDM predictions and population growth in roe deer (Acevedo et al. 2017) and North American trees (Le Squin et al. 2021) respectively.

Although measures of population growth rate (e.g. λ or r) represent the gold standard for assessing population mean fitness, for practical reasons, investigators often rely on functional traits or fitness components to measure population performance. Here, SDMs seem to do better, with 16 studies reporting a positive correlation between SDM predictions and at least one measure of performance (Supporting information; all support for this parameter in Fig. 2 is from studies testing performance). However, many studies reported negative or no relationships between SDM predictions and performance. This was especially true of studies testing predictions for a small number of populations, with no study with fewer than 12 sites finding the expected relationship between SDM predictions and population performance (Supporting

information). As with occurrence and abundance, studies examining results for a single species were more likely to provide clear support for the ability of SDMs to predict performance than studies of multiple species (Supporting information; Fig. 2), and notably, those multispecies studies finding a positive relationship between SDM predictions and performance examined just a single performance metric (Supporting information). Overall, although SDMs sometimes predict relative differences in population performance, in most cases the relationship is weak or dependent on the metric of performance used.

Limitations

Several factors may limit the ability of SDMs to predict population growth rates or performance (reviewed by Csergő et al. 2017). Importantly, different vital rates may respond differently to environmental gradients (Doak and Morris 2010). Demographic compensation, whereby some vital rates decrease while others increase across a gradient of habitat suitability, may cause λ to be unrelated to suitability (Villemas et al. 2015, Csergő et al. 2017). Although studies assessing multiple vital rates in relation to SDM predictions have not found evidence for demographic compensation (Bayly and Angert 2019, Baer and Maron 2020), more studies are needed to test the frequency with which such compensation occurs. Even in the absence of demographic compensation, SDM predictions and population growth rates may be decoupled for a number of reasons. For instance, time lags may mean long-lived species show positive growth rates in habitats that have only recently become unsuitable (Hylander and Ehrlén 2013). Populations may also experience disturbances that lead to transient perturbations to demographic rates (Stott et al. 2011) and thus to mismatches between estimates of suitability based on average environmental conditions and population growth rates at a particular time (Ureta et al. 2018). While demographic buffering may reduce the impact of environmental variability on population growth rates over time (Pfister 1998, Hilde et al. 2020), such mismatches may be particularly important in regions that only occasionally experience disturbance. Finally, we note that a positive relationship between the environment (i.e. habitat suitability) and intrinsic growth (r) is only expected when populations are at low densities. As a population approaches its equilibrium density at a location (i.e. the 'carrying capacity'), measured growth rates are expected to approach zero (ignoring migration; reviewed in Dhondt 1988, Holt 2020). Thus, at least in theory, it is possible for independent measures of population growth based on changes in abundance to be close to zero, even in suitable habitats. This may confound the relationship between SDM predictions and measured growth rates. Thus, if the goal is to determine where populations are likely to persist, it may be pertinent to also compare SDM predictions to estimates of population carrying capacities (Thuiller et al. 2014; see discussion in Holt 2020).

Conclusions and recommendations

To date there is little support for the use of SDMs to predict population mean fitness. However, few studies have measured r or λ and there is a strong taxonomic bias towards plants. SDMs sometimes predict other metrics of population performance. Even if performance does not translate into long-term population growth rates, being able to predict performance may be useful for understanding short-term population dynamics (e.g. changes in age-structure between generations) and the responses of populations to environmental change (Csergő et al. 2017). However, others have found that the strength of the relationship between SDM predictions and performance varies with scale, performance metric and species (Thuiller et al. 2010, Moore et al. 2011, Sheppard et al. 2014, Aizpurua et al. 2017). Future work is needed to assess whether SDMs predict certain performance metrics better than others (e.g. survival versus growth or reproduction). However, as with abundance, when the goal is to predict population mean fitness or performance, many have called for the outright replacement of simple SDMs with dynamic models that incorporate variation in demographic rates (Schurr et al. 2012, Pagel et al. 2020).

Genetic diversity

We found 18 studies that present data on the relationship between SDM predictions and genetic diversity within populations. Although most of these studies did not set out to validate the use of SDMs to predict genetic diversity per se (but see Diniz-Filho et al. 2015), all assume that habitat suitability influences genetic diversity (i.e. through effects on population size or persistence). Results from these studies allow us to examine the extent to which SDMs predict relative differences in genetic variation across space. Twelve studies examined the relationship between SDM predictions and range-wide patterns genetic diversity. Soares et al. (2015) and Jin et al. (2020) reported positive correlations between SDM predictions and genetic diversity across the range of a tropical tree and an arthropod species respectively. However, only two out of 12 studies (~17%) found clear support for a positive relationship between SDM predictions and range-wide genetic diversity, with most studies reporting either mixed or no support for this relationship (Supporting information). Thus range-wide patterns of genetic diversity are not consistently predicted by SDMs.

Seven studies compared SDM predictions and genetic diversity at local or regional scales within species ranges. All but one of these (Pitra et al. 2011) provided mixed (or no) support for a positive relationship between SDM predictions and genetic diversity (Supporting information). For example, Trigila et al. (2016) found that observed and expected heterozygosity were correlated with SDM predictions in river otters in Argentina, but that other measures of genetic diversity (number of alleles, allelic richness, inbreeding coefficient) showed no relationship with SDM predictions. Sinai et al. (2019) examined the relationship between SDM predictions

and allelic richness in three regions towards the edge of the range of an endangered salamander in Israel. Allelic richness was positively correlated with SDM predictions in one region but uncorrelated with SDM predictions in the other two regions studied (Sinai et al. 2019). Studies comparing models based on SDM predictions to models with predictors related to population connectivity have likewise found that SDM predictions have little explanatory power (i.e. are not retained in the ‘best’ model) when explaining genetic diversity (Ortego et al. 2015, Gaddis et al. 2016, Collevatti et al. 2020). Thus, there is little evidence to suggest that SDMs can be used to predict genetic diversity at any scale (Fig. 2).

Limitations

Although we might expect the effects of genetic drift to be more prominent in small populations occupying poor-quality habitat, other processes shape the amount and distribution of genetic diversity. First, the legacy of founder events during colonization may override any effect of habitat suitability on range-wide patterns of genetic variation (Eckert et al. 2008, Pironon et al. 2017). Indeed, range position (a common proxy for historical colonization) is confounded with SDM-predicted habitat suitability in the majority of studies reporting a positive relationship between genetic diversity and SDM predictions (see notes in the Supporting information). In addition to founder effects, historical conditions and environmental change can leave lasting signatures on

genetic diversity. Thus, static SDM estimates of habitat suitability based on current conditions may fail to predict genetic variation in many contexts (Carvalho et al. 2017, Jin et al. 2020). Selection can also influence genetic diversity across large parts of the genome (Ellegren and Galtier 2016), potentially obscuring the relationship between environmental conditions and genetic variation. Finally, gene flow increases genetic variation within recipient sites and thus well-connected populations may harbour high levels of genetic variation regardless of habitat suitability (Consuegra et al. 2005).

Conclusions and recommendations

As with other parameters, several processes other than habitat suitability can impact levels of genetic diversity within a site. Thus, even the best SDMs are likely poor stand-ins for direct measurements of genetic variation across space. Studies using SDMs to derive alternative measures of habitat suitability (i.e. ‘distance to niche centroid’) have similarly found mixed support for a relationship between suitability and genetic diversity. For example, in a study of 40 species, Lira-Noriega and Manthey (2014) found that range-wide patterns of genetic diversity were negatively correlated with distance to the centre of species’ climatic niches for 28 species (70%); However, correlations were only significant in nine cases and patterns were confounded with range position in most cases. Ultimately, when the goal is to use SDMs to identify sites or regions that may harbour relatively high

Support for the ability of SDMs to predict different population parameters

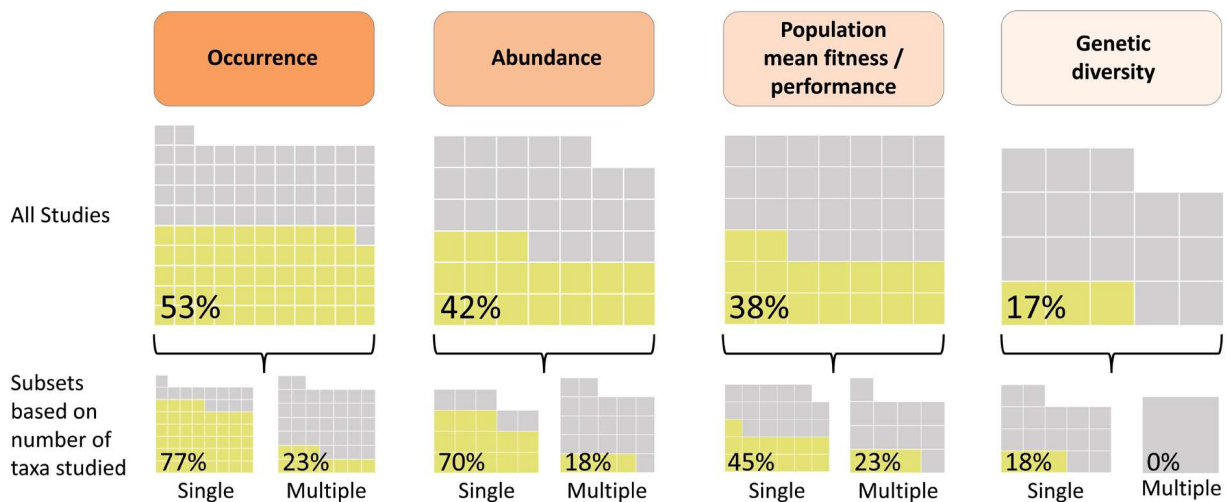


Figure 2. Support for the ability of species distribution models (SDMs) to predict independent data on occurrence, abundance, population mean fitness and performance, and genetic diversity. A positive relationship between SDM predictions and each of these population parameters is expected but support for this relationship declines with increasing distance away from occurrence. Each cell in the four panels represents one published study, with yellow cells denoting studies providing full support (i.e. across all models for a given species or all species for multi-species studies) for the expected relationship (‘yes’ in the ‘support for SDM predictions’ columns in the Supporting information). The top row of plots includes all studies, with the two plots below each panel distinguishing between studies testing single versus multiple species. Alternative breakdowns of support for the success of SDMs in each case (i.e. using less strict criteria for support) lead to slightly different percentage values; However, for no parameter does support exceed 65% and the relative differences in support across parameters and between studies of single versus multiple taxa remain qualitatively unchanged (for details see main text and tables in the Supporting information).

(or low) levels of genetic diversity, investigators may need to account for range position or the amount of population connectivity between sites (either as predictors in models or in post-hoc revisions to model; Ortego et al. 2015). Likewise, incorporating historic climatic conditions and/or the long-term environmental stability of sites in SDMs is likely to be important when using these models to predict genetic diversity (Duncan et al. 2015, Carvalho et al. 2017, Jin et al. 2020).

Discussion

How often do SDMs predict the population biology of species?

Across all parameters considered, we found mixed support for the ability of SDMs to accurately predict independent data (Fig. 2). Thus, our answer to the question of whether SDMs can be used to predict the population biology of species is a very tentative ‘sometimes’. Furthermore, empirical support for the use of SDMs varied considerably depending on the parameter examined (Fig. 2). Whereas SDMs often (but not always) predict the occurrence of species, there is less support for their ability to consistently predict population abundance, performance, or genetic diversity, and almost no support for the ability of these models to predict population growth rates.

However, we acknowledge several limitations to our ability to assess the accuracy of SDMs. Critically, few studies have tested models with independent data. At the time of writing, a less-targeted search of the literature excluding the terms we used to retrieve studies explicitly testing models with independent data retrieved > 21 000 studies – over ten-fold the number retrieved when targeting studies with independent data (~1800). Furthermore, after scrutiny, we were able to retain only 201 studies based on our criteria here. Although we excluded some studies for reasons other than not having independent data (e.g. studies focusing on invasive species or climate change; Supporting information), even if we assume all studies retrieved in our targeted search had independent data of some form (a very generous assumption given that we excluded dozens of studies for specifically failing to meet this criterion), relative to the widespread use of these models in ecology, studies validating models with independent data are rare.

Furthermore, our survey of the literature points to potential publication bias among studies that have conducted independent model evaluation. Support for the ability of SDMs to predict independent data was higher when considering single-species versus multi-species studies. For example, 77% of studies examining a single species found that SDMs successfully predicted occurrence, whereas only 23% of studies incorporating more than one species found unambiguous support for the ability of SDMs to predict independent occurrence (Fig. 2). Similar discrepancies were found for the other parameters (Fig. 2), suggesting that single-species

studies reporting non-significant results are less likely to be published (Greenwald 1975). Furthermore, authors of single-species studies may be more likely to discard models that perform poorly. Thus, the current literature may overestimate the success of SDMs and underestimate the variance in success for individual species. For these reasons, even our tentative answer of ‘sometimes’ may be optimistic when it comes to answering the question as to whether SDMs can be used to predict different population parameters.

Issues with reporting also limit our ability to fully assess the performance of SDMs. In many cases, we had to rely on author conclusions because results for individual tests or species were not (clearly) presented. However, we noted substantial variation in interpretation among those studies presenting statistical results. For example, Chardon et al. (2020) found a weak but significant positive relationship between SDM predictions and plant size in *Silene acaulis* and concluded that SDMs were unreliable for predicting population performance. In contrast, Sheppard et al. (2014) concluded that support for a positive relationship between performance and SDM suitability was ‘generally high’, despite a number of non-significant results. Furthermore, sample sizes within studies were often small, not only limiting statistical power within studies, but precluding exploration of more complex relationships (i.e. ‘wedge-shape’ or ‘J-shaped’ relationships) between SDM predictions and parameters of interest. Formal meta-analyses allowing for calculation of overall effect sizes for each parameter would be ideal, but are currently limited by the wide variation in methods used and metrics reported by different studies. Thus, additional studies, larger studies and greater standardization in reporting (Zurell et al. 2020) is needed to more quantitatively assess the fit of SDM predictions to independent data.

Finally, in their discussion of best practices for model evaluation, Araújo et al. (2019) distinguish between model accuracy (the ability of models to correctly predict events in the same region and timeframe) and model generality (the ability of models to correctly predict events in other regions or timeframes). Although it was not possible to quantify whether conditions at the independent test sites used in each study were within the range of conditions used to generate models, we excluded studies involving extreme model transfer across time and space (i.e. invasive species; Supporting information). Thus, our conclusions speak more to model accuracy than generality, and suggest that many models are inaccurate when it comes to predicting different population parameters. Whether SDMs are effective tools for predicting population parameters in other times (e.g. under past or future climates) or places (e.g. invasive species) is a different question; but we suspect the same challenges and limitations apply.

Limitations to the use of SDMs

Like other studies examining more than one population parameter (Gogol-Prokurat 2011, Nagaraju et al. 2013, Thuiller et al. 2014), we found that the predictive performance of SDMs declined as population parameters more

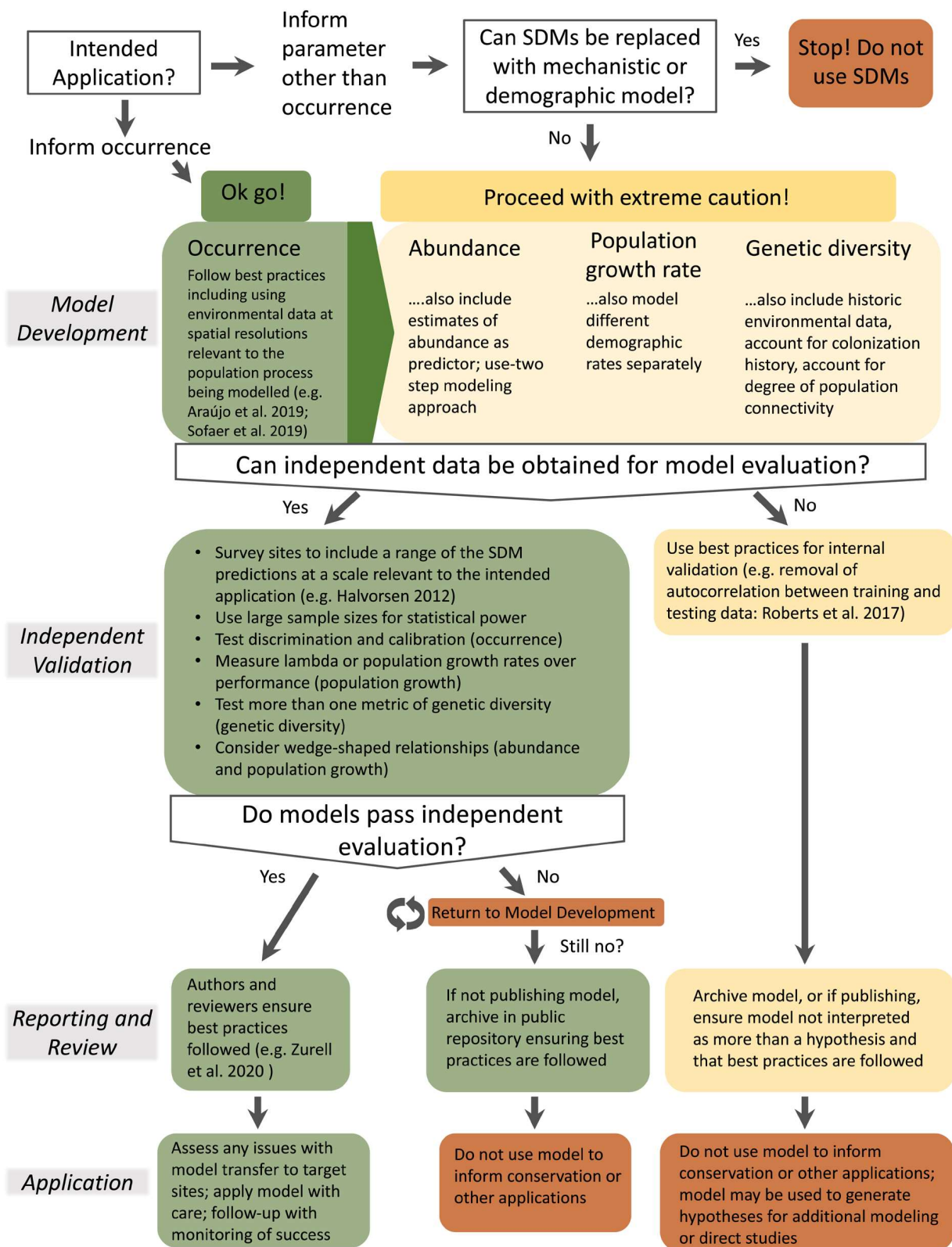


Figure 3. Recommendations for the use of species distribution models (SDMs) in population biology. Results from our review suggest that for parameters other than occurrence, these models should be used with extreme caution or replaced altogether. Regardless of the intended application, SDMs should be evaluated with independent data whenever possible. Models that do not pass independent evaluation should be rejected for use in applications requiring model accuracy but publicly archived to reduce publication bias and to improve our understanding of when and where SDMs perform well versus poorly. Models that cannot be tested with independent data should be regarded as hypotheses only and not used for applications where model accuracy is critical, especially when the goal is to model something other than occurrence.

distantly linked to occurrence were considered (Fig. 2). That SDMs do best at predicting independent occurrences may reflect the fact that SDMs use occurrence data as input and thus implicitly model occurrence. The use of SDMs to infer other population parameters assumes that the conditions governing occurrence also govern other parameters, which may not be the case (Potts and Elith 2006, Bijleveld et al. 2018, Pironon et al. 2018). Even when the distribution of the occurrences of a species is well-aligned with the distribution of suitable habitat (but see discussion below and Pagel et al. 2020), occurrence in a closed population only requires that conditions are such that deaths are matched by births. Past this threshold for persistence, abundance, performance and population mean fitness are expected to scale with the degree to which conditions are favourable (Brown 1984, Brown et al. 1995) and SDMs based on occurrence data may not adequately capture this variation. Put simply, probability of occurrence may not scale with habitat suitability.

Mismatches in spatial or temporal scale may also limit the ability of SDMs to predict different population parameters. SDMs are often based on coarse-scale abiotic variables (Peterson 2011). However, realized growth rates at a given site may depend on microhabitat conditions (Dullinger et al. 2012, Hylander and Ehrlén 2013) and biotic interactions and/or be subject to regulatory processes (discussed in Thuiller et al. 2014), time lags (Hylander and Ehrlén 2013) or stochastic effects (Stott et al. 2011). In other words, SDMs based on coarse-scale abiotic variables may, at best, predict the upper limits of population growth rates. Thus, the wedge-shape relationship proposed by VanDerWal et al. (2009) to describe the relationship between SDM predictions and abundance, might be more broadly applicable to the other parameters considered here (Csergő et al. 2017, Chang and Bourque 2020).

Finally, input occurrences can be misleading with respect to what constitutes (even minimally) suitable habitat. Apart from any issues with sampling bias (Phillips et al. 2009), the distribution of species across space may be decoupled from the suitability of sites for several reasons (see Limitations under Occurrence). Some have even put forth that it may be rare for populations to occur in truly suitable habitats and that sink populations may be prevalent (Van Horne 1983, Pulliam 1988). This would severely impact our ability to draw conclusions about habitat suitability from occurrence data, limiting the use of SDMs for many of the parameters considered here.

Implications for conservation

SDMs are used to make conservation decisions (Franklin 2010, Guillera-Arroita et al. 2015, Sofaer et al. 2019); However, our findings call for caution in this regard. Although SDMs can sometimes predict the upper limits of abundance (see discussion for Abundance) and thus may be useful for applications where there is a need to identify places where population sizes and growth rates could be high (e.g.

choosing sites for translocation or restoration: Osborne and Seddon 2012, Zellmer et al. 2019), the limited ability of these models to predict relative differences in performance (let alone population mean fitness), suggests they may be less useful for prioritizing existing populations for protection. Likewise, because SDMs do not consistently predict genetic diversity, they cannot inform the evolutionary potential of populations, and thus population resilience to environmental change. Thus, careful demographic and genetic studies are needed to complement SDMs when the goal is to identify populations with the best chance of long-term persistence.

More optimistically, we found that SDMs often do a reasonable job of predicting independent occurrences. Thus, these models may be most useful for applications where the end-goal is to identify undetected populations (Boetsch et al. 2003, Rosner-Katz et al. 2020) or to rule out locations where conditions are unable to support even minimally viable populations (Heinrichs et al. 2010). However, we note that care is needed when parameterizing models (Araújo et al. 2019, Sofaer et al. 2019) and interpreting their output (Guillera-Arroita et al. 2015) for these purposes. Most importantly, we reiterate that models should be validated with independent data before they are used in this regard (Elith and Burgman 2002, Guillera-Arroita et al. 2015).

Conclusions

Our review shows that for a number of applications in population biology, SDMs are frequently inaccurate. We have made several recommendations that may allow for the development of better models going forward (summarized in Fig. 3). However, achieving accurate models for many of the parameters considered here may ultimately require moving beyond simple, occurrence-based SDMs and using approaches that specifically incorporate demographic information (Schurr et al. 2012, Diez et al. 2014, Ehrlén and Morris 2015, Evans et al. 2016, Pagel et al. 2020). However, SDMs are not useless. Our survey of the literature suggests that SDMs often make reasonably accurate predictions about occurrence. But we propose that the most appropriate use of these models is as a tool for generating initial hypotheses about where a species may be and how it may fare in different places. From there, as Elith and Burgman (2002) have noted: ‘if (the SDM) has not been tested with independent data at the grain and extent that is relevant to the application, the results need to be assessed with care’. Indeed, given the limited ability of these models to predict the demographics of real populations, without such independent validation, the use of these models for many applications in ecology, and especially conservation biology, should be viewed with extreme skepticism.

Acknowledgements—We appreciate the time and careful consideration received from anonymous reviewers and the subject editor – their comments helped to greatly improve the paper.

Funding – JAL and JLM were supported by Natural Sciences and Engineering Research Council of Canada Discovery Grants. SP was funded by a Future Leader Fellowship from Royal Botanic Gardens, Kew. SNS was supported in this work by the USDA National Inst. of Food and Agriculture Hatch project no. 1016272.

Author contributions

Julie Lee-Yaw: Conceptualization (lead); Data curation (equal); Formal analysis (equal); Methodology (equal); Project administration (lead); Visualization (equal); Writing – original draft (lead); Writing – review and editing (lead). **Jenny L. McCune:** Data curation (equal); Formal analysis (equal); Methodology (equal); Visualization (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Samuel Pironon:** Data curation (equal); Formal analysis (equal); Methodology (equal); Project administration (equal); Visualization (equal); Writing – original draft (supporting); Writing – review and editing (supporting). **Seema N. Sheth:** Data curation (equal); Formal analysis (equal); Methodology (equal); Visualization (supporting); Writing – original draft (supporting); Writing – review and editing (supporting).

Transparent Peer Review

The peer review history for this article is available at <<https://publons.com/publon/10.1111/ecog.05877>>.

Data availability statement

The studies reviewed and data extracted from these studies are included as supplementary material.

Supporting information

Any supporting information associated with this article is available from the online version.

References

- Acevedo, P. et al. 2017. Population dynamics affect the capacity of species distribution models to predict species abundance on a local scale. – *Divers. Distrib.* 23: 1008–1017.
- Agashe, D. 2009. The stabilizing effect of intraspecific genetic variation on population dynamics in novel and ancestral habitats. – *Am. Nat.* 174: 255–267.
- Aitken, M. et al. 2007. Modeling distributions of rare plants in the Great Basin, western North America. – *West. N. Am. Nat.* 67: 26–38.
- Aizpurua, O. et al. 2017. Evaluating the reliability of species distribution models with an indirect measure of bird reproductive performance. – *J. Avian Biol.* 48: 1575–1582.
- Araújo, M. B. et al. 2019. Standards for distribution models in biodiversity assessments. – *Sci. Adv.* 5: eaat4858.
- Baer, K. C. and Maron, J. L. 2020. Ecological niche models display nonlinear relationships with abundance and demographic performance across the latitudinal distribution of *Astragalus utahensis* (Fabaceae). – *Ecol. Evol.* 10: 8251–8264.
- Bayly, M. J. and Angert, A. L. 2019. Niche models do not predict experimental demography but both suggest dispersal limitation across the northern range limit of the scarlet monkeyflower *Erythranthe cardinalis*. – *J. Biogeogr.* 46: 1316–1328.
- Bijleveld, A. I. et al. 2018. Presence–absence of marine macrozoobenthos does not generally predict abundance and biomass. – *Sci. Rep.* 8: 3039.
- Boetsch, J. R. et al. 2003. Predicting rare plant occurrence in Great Smoky Mountains National Park, USA. – *Nat. Areas J.* 23: 229–237.
- Boiffin, J. et al. 2017. Species distribution models may misdirect assisted migration: insights from the introduction of Douglas-fir to Europe. – *Ecol. Appl.* 27: 446–457.
- Boulangeat, I. et al. 2012. Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances: the role of dispersal and biotic interactions in explaining species distributions and abundances. – *Ecol. Lett.* 15: 584–593.
- Box, G. E. P. 1976. Science and statistics. – *J. Am. Stat. Assoc.* 71: 791–799.
- Braz, A. G. et al. 2020. Interspecific competition constrains local abundance in highly suitable areas. – *Ecography* 43: 1560–1570.
- Brown, J. H. 1984. On the relationship between abundance and distribution of species. – *Am. Nat.* 124: 255–279.
- Brown, J. H. et al. 1995. Spatial variation in abundance. – *Ecology* 76: 2028–2043.
- Burgman, M. A. and Fox, J. C. 2003. Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. – *Anim. Conserv.* 6: 19–28.
- Calixto-Pérez, E. et al. 2018. Integrating expert knowledge and ecological niche models to estimate Mexican primates' distribution. – *Primates* 59: 451–467.
- Carvalho, C. da S. et al. 2017. Climatic stability and contemporary human impacts affect the genetic diversity and conservation status of a tropical palm in the Atlantic Forest of Brazil. – *Conserv. Genet.* 18: 467–478.
- Chang, Y. and Bourque, C. P.-A. 2020. Relating modelled habitat suitability for *Abies balsamea* to on-the-ground species structural characteristics in naturally growing forests. – *Ecol. Indic.* 111: 105981.
- Chardon, N. I. et al. 2020. Incorporating intraspecific variation into species distribution models improves distribution predictions, but cannot predict species traits for a wide-spread plant species. – *Ecography* 43: 60–74.
- Chase, J. M. and Leibold, M. A. 2003. Ecological niches: linking classic and contemporary approaches. – Univ. of Chicago Press.
- Chatfield, C. 1995. Model uncertainty, data mining and statistical inference. – *J. R. Stat. Soc. Ser. A* 158: 419.
- Collevatti, R. G. et al. 2020. Landscape changes decrease genetic diversity in the Pallas' long-tongued bat. – *Perspect. Ecol. Conserv.* 18: 169–177.
- Consuegra, S. et al. 2005. Asymmetric gene flow and the evolutionary maintenance of genetic diversity in small, peripheral Atlantic salmon populations. – *Conserv. Genet.* 6: 823–842.
- Csergő, A. M. et al. 2017. Less favourable climates constrain demographic strategies in plants. – *Ecol. Lett.* 20: 969–980.
- Cunningham, H. R. et al. 2016. Abiotic and biotic constraints across reptile and amphibian ranges. – *Ecography* 39: 1–8.

- Dallas, T. A. and Hastings, A. 2018. Habitat suitability estimated by niche models is largely unrelated to species abundance. – *Global Ecol. Biogeogr.* 27: 1448–1456.
- Dhondt, A. 1988. Carry capacity: a confusing concept. – *Acta Oecol.* 9: 337–346.
- Diez, J. M. et al. 2014. Probabilistic and spatially variable niches inferred from demography. – *J. Ecol.* 102: 544–554.
- Diniz-Filho, J. A. F. et al. 2015. Correlation between genetic diversity and environmental suitability: taking uncertainty from ecological niche models into account. – *Mol. Ecol. Resour.* 15: 1059–1066.
- Doak, D. F. and Morris, W. F. 2010. Demographic compensation and tipping points in climate-induced range shifts. – *Nature* 467: 959–962.
- Domisch, S. et al. 2016. Model-based integration of observed and expert-based information for assessing the geographic and environmental distribution of freshwater species. – *Ecography* 39: 1078–1088.
- Duan, R.-Y. et al. 2014. The predictive performance and stability of six species distribution models. – *PLoS One* 9: e112764.
- Dullinger, S. et al. 2012. Post-glacial migration lag restricts range filling of plants in the European Alps: range filling of Alpine plants. – *Global Ecol. Biogeogr.* 21: 829–840.
- Duncan, S. I. et al. 2015. History matters more when explaining genetic diversity within the context of the core-periphery hypothesis. – *Mol. Ecol.* 24: 4323–4336.
- Eckert, C. G. et al. 2008. Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. – *Mol. Ecol.* 17: 1170–1188.
- Edwardsen, A. et al. 2011. A fine-grained spatial prediction model for the red-listed vascular plant *Scorzoneria humilis*. – *Nord. J. Bot.* 29: 495–504.
- Ehrlén, J. and Morris, W. F. 2015. Predicting changes in the distribution and abundance of species under environmental change. – *Ecol. Lett.* 18: 303–314.
- Elith, J. and Burgman, M. 2002. Predictions and their validation: rare plants in the Central Highlands, Victoria, Australia. – In: Scott, J. M. et al. (eds), *Predicting species occurrences: issues of accuracy and scale*. Island Press, pp. 303–313.
- Elith, J. and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. – *Annu. Rev. Ecol. Evol. Syst.* 40: 677–697.
- Elith, J. et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. – *Ecography* 29: 129–151.
- Elith, J. et al. 2011. A statistical explanation of MaxEnt for ecologists: statistical explanation of MaxEnt. – *Divers. Distrib.* 17: 43–57.
- Elith, J. et al. 2020. Presence-only and presence–absence data for comparing species distribution modeling methods. – *Biodivers. Inform.* 15: 69–80.
- Ellegren, H. and Galtier, N. 2016. Determinants of genetic diversity. – *Nat. Rev. Genet.* 17: 422–433.
- Evans, M. E. K. et al. 2016. Towards process-based range modeling of many species. – *Trends Ecol. Evol.* 31: 860–871.
- Fois, M. et al. 2018. Using species distribution models at local scale to guide the search of poorly known species: review, methodological issues and future directions. – *Ecol. Model.* 385: 124–132.
- Fourcade, Y. 2016. Comparing species distributions modelled from occurrence data and from expert-based range maps. Implication for predicting range shifts with climate change. – *Ecol. Inform.* 36: 8–14.
- Frankham, R. 1996. Relationship of genetic variation to population size in wildlife. – *Conserv. Biol.* 10: 1500–1508.
- Franklin, J. 2010. Moving beyond static species distribution models in support of conservation biogeography: moving beyond static species distribution models. – *Divers. Distrib.* 16: 321–330.
- Gaddis, K. D. et al. 2016. Dry-washes determine gene flow and genetic diversity in a common desert shrub. – *Landscape Ecol.* 31: 2215–2229.
- Gaston, K. J. 2003. *The structure and dynamics of geographic ranges*. – Oxford Univ. Press.
- Gogol-Prokurat, M. 2011. Predicting habitat suitability for rare plants at local spatial scales using a species distribution model. – *Ecol. Appl.* 21: 33–47.
- Greenwald, A. G. 1975. Consequences of prejudice against the null hypothesis. – *Psychol. Bull.* 82: 1–20.
- Gregg, E. J. et al. 2019. Why less complexity produces better forecasts: an independent data evaluation of kelp habitat models. – *Ecography* 42: 428–443.
- Greiser, C. et al. 2020. Climate limitation at the cold edge: contrasting perspectives from species distribution modelling and a transplant experiment. – *Ecography* 43: 637–647.
- Guillera-Aroita, G. et al. 2015. Is my species distribution model fit for purpose? Matching data and models to applications: matching distribution models to applications. – *Global Ecol. Biogeogr.* 24: 276–292.
- Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – *Ecol. Lett.* 8: 993–1009.
- Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. – *Ecol. Model.* 135: 147–186.
- Hallgren, W. et al. 2019. Species distribution models can be highly sensitive to algorithm configuration. – *Ecol. Model.* 408: 108719.
- Halvorsen, R. 2012. A gradient analytic perspective on distribution modelling. – *Sommerfeltia* 35: 1–165.
- Heikkinen, R. K. et al. 2007. Biotic interactions improve prediction of boreal bird distributions at macro-scales. – *Global Ecol. Biogeogr.* 16: 754–763.
- Heinrichs, J. A. et al. 2010. Assessing critical habitat: evaluating the relative contribution of habitats to population persistence. – *Biol. Conserv.* 143: 2229–2237.
- Hilde, C. H. et al. 2020. The demographic buffering hypothesis: evidence and challenges. – *Trends Ecol. Evol.* 35: 523–538.
- Holt, R. D. 2020. Reflections on niches and numbers. – *Ecography* 43: 387–390.
- Hosmer, D. W. and Lemeshow, S. 2000. *Applied logistic regression*, 2nd ed. – Wiley.
- Hutchinson, G. E. 1957. Concluding remarks. Population studies: animal ecology and demography. – *Cold Springs Harbor Symp. Quant. Biol.* 22: 415–427.
- Hylander, K. and Ehrlén, J. 2013. The mechanisms causing extinction debts. – *Trends Ecol. Evol.* 28: 341–346.
- Jenkins, C. N. et al. 2003. Why sparrow distributions do not match model predictions. – *Anim. Conserv.* 6: 39–46.
- Jetz, W. et al. 2012. Integrating biodiversity distribution knowledge: toward a global map of life. – *Trends Ecol. Evol.* 27: 151–159.
- Jin, P.-Y. et al. 2020. Geography alone cannot explain *Tetranychus truncatus* (Acari: Tetranychidae) population abundance and genetic diversity in the context of the center–periphery hypothesis. – *Heredity* 124: 383–396.
- Kozak, K. H. and Wiens, J. J. 2006. Does niche conservatism promote speciation? A case study in North American salamanders. – *Evolution* 60: 2604.

- Lawson, C. R. et al. 2014. Prevalence, thresholds and the performance of presence-absence models. – *Methods Ecol. Evol.* 5: 54–64.
- Le Squin, A. et al. 2021. Climate-induced variation in the demography of 14 tree species is not sufficient to explain their distribution in eastern North America. – *Global Ecol. Biogeogr.* 30: 352–369.
- Lee-Yaw, J. A. et al. 2016. A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. – *Ecol. Lett.* 19: 710–722.
- Lee-Yaw, J. A. et al. 2018. Environmental marginality and geographic range limits: a case study with *Arabidopsis lyrata* ssp. *lyrata*. – *Ecography* 41: 622–634.
- Leimu, R. et al. 2006. How general are positive relationships between plant population size, fitness and genetic variation? – *J. Ecol.* 94: 942–952.
- Lira-Noriega, A. and Manthey, J. D. 2014. Relationship of genetic diversity and niche centrality: a survey and analysis. – *Evolution* 68: 1082–1093.
- Lobo, J. M. et al. 2008. AUC: a misleading measure of the performance of predictive distribution models. – *Global Ecol. Biogeogr.* 17: 145–151.
- Loiselle, B. A. et al. 2007. Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? – *J. Biogeogr.* 35: 105–116.
- Loyn, R. H. et al. 2001. Modelling landscape distributions of large forest owls as applied to managing forests in north-east Victoria, Australia. – *Biol. Conserv.* 97: 361–376.
- Maguire, B. 1973. Niche response structure and the analytical potentials of its relationship to the habitat. – *Am. Nat.* 107: 213–246.
- Mainali, K. et al. 2020. Matching expert range maps with species distribution model predictions. – *Conserv. Biol.* 34: 1292–1304.
- Marcere, A. et al. 2013. Using species distribution modelling to disentangle realised versus potential distributions for rare species conservation. – *Biol. Conserv.* 166: 221–230.
- McCune, J. L. 2016. Species distribution models predict rare species occurrences despite significant effects of landscape context. – *J. Appl. Ecol.* 53: 1871–1879.
- McCune, J. L. et al. 2020. Do traits of plant species predict the efficacy of species distribution models for finding new occurrences? – *Ecol. Evol.* 10: 5001–5014.
- Mccusker, M. R. and Bentzen, P. 2010. Positive relationships between genetic diversity and abundance in fishes: genetic diversity and abundance relationships. – *Mol. Ecol.* 19: 4852–4862.
- Mellert, K. H. et al. 2011. Hypothesis-driven species distribution models for tree species in the Bavarian Alps. – *J. Veg. Sci.* 22: 635–646.
- Mellin, C. et al. 2012. Geographic range determinants of two commercially important marine molluscs: modelling the distribution of abalone. – *Divers. Distrib.* 18: 133–146.
- Merow, C. et al. 2014. What do we gain from simplicity versus complexity in species distribution models? – *Ecography* 37: 1267–1281.
- Merow, C. et al. 2017. Integrating occurrence data and expert maps for improved species range predictions: expert maps and point process models. – *Global Ecol. Biogeogr.* 26: 243–258.
- Moore, K. A. et al. 2011. Can spatial isolation help predict dispersal-limited sites for native species restoration? – *Ecol. Appl.* 21: 2119–2128.
- Muñoz, A.-R. et al. 2015. Environmental favourability as a cost-efficient tool to estimate carrying capacity. – *Divers. Distrib.* 21: 1388–1400.
- Nagaraju, S. K. et al. 2013. Do ecological niche model predictions reflect the adaptive landscape of species? A test using *Myristica malabarica* Lam., an endemic tree in the Western Ghats, India. – *PLoS One* 8: e82066.
- O’Hara, C. C. et al. 2017. Aligning marine species range data to better serve science and conservation. – *PLoS One* 15: e0175739.
- Ortega-Huerta, M. A. and Vega-Rivera, J. H. 2017. Validating distribution models for twelve endemic bird species of tropical dry forest in western Mexico. – *Ecol. Evol.* 7: 7672–7686.
- Ortego, J. et al. 2015. Living on the edge: the role of geography and environment in structuring genetic variation in the southernmost populations of a tropical oak. – *Plant Biol.* 17: 676–683.
- Osborne, P. E. and Seddon, J. P. 2012. Selecting suitable habitat for reintroductions: variation, change and the role of species distribution modelling. – In: Ewen, J. G. et al. (eds), *Reintroduction biology: integrating science and management*. Wiley Blackwell, pp. 73–104.
- Osorio-Olvera, L. et al. 2020. Relationships between population densities and niche-centroid distances in North American birds. – *Ecol. Lett.* 23: 555–564.
- Page, J. et al. 2020. Mismatches between demographic niches and geographic distributions are strongest in poorly dispersed and highly persistent plant species. – *Proc. Natl Acad. Sci. USA* 117: 3663–3669.
- Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – *Global Ecol. Biogeogr.* 12: 361–371.
- Peterson, A. T. 2006. Uses and requirements of ecological niche models and related distributional models. – *Biodivers. Inform.* 3: 59–72.
- Peterson, A. T. and Soberón, J. 2012. Species distribution modeling and ecological niche modeling: getting the concepts right. – *Nat. Conserv.* 10: 102–107.
- Peterson, A. T. et al. 2011. *Ecological niches and geographic distributions*. – Princeton Univ. Press.
- Peterson, A. T. et al. 2018. Assumption- versus data-based approaches to summarizing species’ ranges: summarizing species’ ranges. – *Conserv. Biol.* 32: 568–575.
- Pfister, C. A. 1998. Patterns of variance in stage-structured populations: evolutionary predictions and ecological implications. – *Proc. Natl Acad. Sci. USA* 95: 213–218.
- Phillips, S. J. et al. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. – *Ecol. Appl.* 19: 181–197.
- Pinto, C. et al. 2016. Using individual tracking data to validate the predictions of species distribution models. – *Divers. Distrib.* 22: 682–693.
- Pironon, S. et al. 2017. Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. – *Biol. Rev.* 92: 1877–1909.
- Pironon, S. et al. 2018. The ‘Hutchinsonian niche’ as an assemblage of demographic niches: implications for species geographic ranges. – *Ecography* 41: 1103–1113.
- Pitra, C. et al. 2011. Linking habitat quality with genetic diversity: a lesson from great bustards in Spain. – *Eur. J. Wildl. Res.* 57: 411–419.
- Potts, J. M. and Elith, J. 2006. Comparing species abundance models. – *Ecol. Model.* 199: 153–163.
- Pulliam, H. R. 1988. Sources, sinks and population regulation. – *Am. Nat.* 132: 652–661.

- Pulliam, H. R. 2000. On the relationship between niche and distribution. – *Ecol. Lett.* 3: 349–361.
- Raedig, C. and Kreft, H. 2011. Influence of different species range types on the perception of macroecological patterns. – *Syst. Biodivers.* 9: 159–170.
- Randin, C. F. et al. 2006. Are niche-based species distribution models transferable in space? – *J. Biogeogr.* 33: 1689–1703.
- Roberts, D. R. et al. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical or phylogenetic structure. – *Ecography* 40: 913–929.
- Rosner-Katz, H. et al. 2020. Using stacked SDMs with accuracy and rarity weighting to optimize surveys for rare plant species. – *Biodivers. Conserv.* 29: 3209–3225.
- Rotenberry, J. T. and Balasubramaniam, P. 2020. Connecting species' geographical distributions to environmental variables: range maps versus observed points of occurrence. – *Ecography* 43: 897–913.
- Sangermano, F. and Eastman, J. R. 2012. A GIS framework for the refinement of species geographic ranges. – *Int. J. Geograph. Inform. Sci.* 26: 39–55.
- Santini, L. et al. 2019. Addressing common pitfalls does not provide more support to geographical and ecological abundant-centre hypotheses. – *Ecography* 42: 696–705.
- Sarquis, J. A. et al. 2018. Species distribution models and empirical test: comparing predictions with well-understood geographical distribution of *Bothrops alternatus* in Argentina. – *Ecol. Evol.* 8: 10497–10509.
- Schurr, F. M. et al. 2012. How to understand species' niches and range dynamics: a demographic research agenda for biogeography. – *J. Biogeogr.* 39: 2146–2162.
- Sheppard, C. S. et al. 2014. Predicting plant invasions under climate change: are species distribution models validated by field trials? – *Global Change Biol.* 20: 2800–2814.
- Sillero, N. 2011. What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. – *Ecol. Model.* 222: 1343–1346.
- Sinai, I. et al. 2019. The role of landscape and history on the genetic structure of peripheral populations of the Near Eastern fire salamander, *Salamandra infraimmaculata*, in northern Israel. – *Conserv. Genet.* 20: 875–889.
- Soares, T. N. et al. 2015. Patterns of genetic variability in central and peripheral populations of *Dipteryx alata* (Fabaceae) in the Brazilian Cerrado. – *Plant Syst. Evol.* 301: 1315–1324.
- Soberón, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. – *Ecol. Lett.* 10: 1115–1123.
- Sofaer, H. R. et al. 2019. Development and delivery of species distribution models to inform decision-making. – *BioScience* 69: 544–557.
- Sporbert, M. et al. 2020. Testing macroecological abundance patterns: the relationship between local abundance and range size, range position and climatic suitability among European vascular plants. – *J. Biogeogr.* 47: 2210–2222.
- Stevens, R. D. and Amarilla-Stevens, H. N. 2012. Seasonal environments, episodic density compensation and dynamics of structure of chiropteran frugivore guilds in Paraguayan Atlantic forest. – *Biodivers. Conserv.* 21: 267–279.
- Stott, I. et al. 2011. A framework for studying transient dynamics of population projection matrix models: a synthesis of transient demography. – *Ecol. Lett.* 14: 959–970.
- Svenning, J.-C. et al. 2008. Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. – *Ecography* 31: 316–326.
- Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. – *Science* 240: 1285–1293.
- Thuiller, W. et al. 2010. Variation in habitat suitability does not always relate to variation in species' plant functional traits. – *Biol. Lett.* 6: 120–123.
- Thuiller, W. et al. 2014. Does probability of occurrence relate to population dynamics? – *Ecography* 37: 1155–1166.
- Trigila, A. P. et al. 2016. Genetic diversity in the neotropical river otter, *Lontra longicaudis* (Mammalia, Mustelidae), in the Lower Delta of Parana River, Argentina and its relation with habitat suitability. – *Hydrobiologia* 768: 287–298.
- Ureta, C. et al. 2018. Inferring space from time: on the relationship between demography and environmental suitability in the desert plant *O. rastrera*. – *PLoS One* 13: e0201543.
- Van Horne, B. 1983. Density as a misleading indicator of habitat quality. – *J. Wildl. Manage.* 47: 893–901.
- van Manen, F. T. et al. 2002. A model to predict the occurrence of surviving butternut trees in the southern Blue Ridge mountains. – In: Scott, J. M. et al. (eds), *Predicting species occurrences: issues of accuracy and scale*. Island Press, pp. 491–497.
- van Manen, F. T. et al. 2005. Habitat models to assist plant protection efforts in Shenandoah National Park, Virginia, USA. – *Nat. Areas J.* 25: 339–350.
- VanDerWal, J. et al. 2009. Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? – *Ecol. Model.* 220: 589–594.
- Vaughan, I. P. and Ormerod, S. J. 2005. The continuing challenges of testing species distribution models: testing distribution models. – *J. Appl. Ecol.* 42: 720–730.
- Vaz, U. L. et al. 2015. Trends and biases in global scientific literature about ecological niche models. – *Braz. J. Biol.* 75: 17–24.
- Villellas, J. et al. 2015. Demographic compensation among populations: what is it, how does it arise and what are its implications? – *Ecol. Lett.* 18: 1139–1152.
- Wauchope-Drumm, M. et al. 2020. Using a species distribution model to guide NSW surveys of the long-footed potoroo *Potorous longipes*: SDM to guide survey of long-footed potoroo. – *Austral Ecol.* 45: 15–26.
- Weber, M. M. et al. 2017. Is there a correlation between abundance and environmental suitability derived from ecological niche modelling? A meta-analysis. – *Ecography* 40: 817–828.
- Willems, E. P. and Hill, R. A. 2009. A critical assessment of two species distribution models: a case study of the vervet monkey *Cercopithecus aethiops*. – *J. Biogeogr.* 36: 2300–2312.
- Zellmer, A. J. et al. 2019. Predicting optimal sites for ecosystem restoration using stacked-species distribution modeling. – *Front. Mar. Sci.* 6: 3.
- Zurell, D. et al. 2016. Benchmarking novel approaches for modelling species range dynamics. – *Global Change Biol.* 22: 2651–2664.
- Zurell, D. et al. 2020. A standard protocol for reporting species distribution models. – *Ecography* 43: 1261–1277.