



Common mistakes in ecological niche models

Neftalí Sillero & A. Márcia Barbosa

To cite this article: Neftalí Sillero & A. Márcia Barbosa (2021) Common mistakes in ecological niche models, International Journal of Geographical Information Science, 35:2, 213-226, DOI: [10.1080/13658816.2020.1798968](https://doi.org/10.1080/13658816.2020.1798968)

To link to this article: <https://doi.org/10.1080/13658816.2020.1798968>



Published online: 27 Jul 2020.



Submit your article to this journal [↗](#)



Article views: 5383



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 19 View citing articles [↗](#)



FORESIGHT PAPER



Common mistakes in ecological niche models

Neftalí Sillero  and A. Márcia Barbosa

CICGE-Centro de Investigação em Ciências Geo-
Espaciais, Faculdade de Ciências da Universidade do Porto, Vila Nova de Gaia, Portugal

ABSTRACT

Ecological niche models (ENMs) are widely used statistical methods to estimate various types of species niches. After lecturing several editions of introductory courses on ENMs and reviewing numerous manuscripts on this subject, we frequently faced some recurrent mistakes: 1) presence-background modelling methods, such as Maxent or ENFA, are used as if they were pseudo-absence methods; 2) spatial autocorrelation is confused with clustering of species records; 3) environmental variables are used with a higher spatial resolution than species records; 4) correlations between variables are not taken into account; 5) machine-learning models are not replicated; 6) topographical variables are calculated from unprojected coordinate systems, and; 7) environmental variables are downscaled by resampling. Some of these mistakes correspond to student misunderstandings and are corrected before publication. However, other errors can be found in published papers. We explain here why these approaches are erroneous and we propose ways to improve them.

ARTICLE HISTORY

Received 19 July 2019

Accepted 8 July 2020

KEYWORDS

Maxent; autocorrelation; filtering records; spatial resolution; topographical variables; resampling; species distribution models

Introduction

Ecological niche models (ENMs) are empirical or mathematical approaches to several types of species ecological niches (Booth *et al.* 1988, Peterson *et al.* 2011, Barbosa *et al.* 2012), relating respectively physiological (mechanistic models) or distribution data (correlative models) to environmental predictor variables (Hutchinson 1957, Sillero 2011). ENMs aim to identify the factors that limit and define the species niches, and thus, to forecast the distribution and suitable habitats of current (Chefaoui *et al.* 2015) or past (Nogues-Bravo *et al.* 2008) species. ENMs can be projected as well to other scenarios in time and space (Barbosa *et al.* 2009, Werkowska *et al.* 2017, Yates *et al.* 2018). ENMs have become popular tools for nature conservation and management (Franklin 2010).

Correlative models are the most frequently used, as they are more feasible. Indeed, sufficient physiological data are seldom available for mechanistic models. On the contrary, spatial environmental data (Hijmans *et al.* 2005, Kriticos *et al.* 2012, Lima-Ribeiro 2015, Fick and Hijmans 2017, Karger *et al.* 2017, Title and Bemmels 2018) and species occurrence data (Yesson *et al.* 2007, Beukema *et al.* 2013, Ficetola *et al.* 2014, Sillero *et al.* 2014, Bencatel *et al.* 2019) are currently very abundant and uncomplicated to obtain (e.g., from data papers and online sources such as GBIF, NatureServe, IUCN, Neotoma,

and others). Recently, global edaphic databases have also been developed for ENM studies (Booth 2018). Some correlative methods are easy to apply even for inexperienced modellers, such as Maxent (Phillips *et al.* 2017). This wide availability and ease of use can make niche modelling especially prone to beginner mistakes. Recent literature has provided sets of guidelines for niche or distribution modelling under different purposes (Araújo *et al.* 2019, Feng *et al.* 2019, Sofaer *et al.* 2019), but several common mistakes remain to be explicitly addressed and explained. Based on our research background on biogeography and spatial ecology, here we outline some common methodological mistakes made by students and modellers, and we explain how they can be amended.

After lecturing several editions of introductory courses on ecological niche models and reviewing numerous manuscripts, we realised that there is a set of recurrent mistakes, namely: 1) presence-background methods are used as pseudo-absence methods; 2) the term 'spatial autocorrelation' is used instead of spatial clustering of occurrence data, and is dealt with by filtering records; 3) environmental variables are used with a higher spatial resolution than species records; 4) correlations between variables are not taken into account; 5) machine learning models are not replicated; 6) topographical (distance-based) variables are calculated from geographical (unprojected) coordinate systems, and; 7) environmental variables are downscaled by resampling. Here we review these common mistakes in calculating correlative ENMs, explain why they are erroneous, and provide guidelines for more correct use.

Presence-background methods are not pseudo-absence methods

Presence-background methods such as Maxent (Phillips *et al.* 2017) or Ecological Niche Factor Analyses (ENFA; Hirzel *et al.* 2002) are frequently used as presence/pseudo-absence methods (Baek *et al.* 2019, Zhang *et al.* 2019). However, they actually are profile or presence-background methods (Guillera-Arroita *et al.* 2014). Background records do not mean species pseudo-absences, but rather a spectrum of the overall available conditions (Phillips *et al.* 2009, Iturbide *et al.* 2018, Hallgren *et al.* 2019). The background is the whole study area, including pixels with presences. Still, many students and authors mistake background with pseudo-absences (e.g. Cai *et al.* 2019, Chapman *et al.* 2019, Farrell *et al.* 2019, Mariani *et al.* 2019, Papier *et al.* 2019, Young *et al.* 2019, Yue *et al.* 2019). Pseudo-absences are artificial absence data (Barbet-Massin *et al.* 2012), i.e. places where the species is supposed (but not confirmed) to be absent. Presence-background methods simply distinguish between suitable and less suitable habitats within the overall analysed background, and not between occupied and unoccupied habitats like presence-absence methods do (Sillero 2011). Hence, presence-background methods require a set of presence-only (not presence/absence or presence/pseudo-absence) records plus a sample of the full background, including presence sites (Guillera-Arroita *et al.* 2014). In fact, better results may be obtained when restricting the background samples to areas near the presences (Phillips *et al.* 2009). Presence-absence methods require samples from non-occupied areas (absences): even if absences are artificial (i.e. pseudo-absences), they do not represent the overall conditions of the study area, but only the non-occupied habitats. Considering Maxent or ENFA 'presence-only methods' without acknowledging the importance of the background is another common mistake, as the chosen background is as

impactful as the chosen set of presences on the results of the model (Phillips *et al.* 2009, Anderson and Raza 2010).

Reduction of spatial autocorrelation between species records is not equivalent to filtering species records

Many students misunderstood reducing spatial autocorrelation with filtering (or thinning) species records. The distinction is also often unclear in published papers (e.g. Brito *et al.* 2011, Báez *et al.* 2019). Filtering is a procedure to reduce the clustering of species records (Boria *et al.* 2014) which frequently results from survey biases, as distribution data are collected more often from sites with easier access (Kadmon *et al.* 2004, Barbosa *et al.* 2010, 2013). Species presences can often be related to proximity to urban centres, roads or research institutions. These biases violate the assumption of independence among species records (Franklin 2010). The modelled distribution does not only correspond to the species' observed distribution, but to the distribution of sampling effort. Hence, it is customary to delete some records, normally using a distance rule, where any point below a distance threshold is deleted until the species dataset is considered as not clustered (Aiello-Lammens *et al.* 2015). Alternatively, Varela *et al.* (2014) showed that filtering by environmental criteria provides better results. However, filtering may provide worse results if performed without reliable information on the bias in species occurrence data (Gábor *et al.* 2019).

Autocorrelation actually is essential for modelling (Segurado *et al.* 2006, Dormann 2007, Dormann *et al.* 2007, De Marco *et al.* 2008). Spatial autocorrelation is a property of the First Law of Geography, Tobler's Law: 'Everything is related to everything else, but near things are more related than distant things' (Tobler 1970). This means that closer locations tend to have more similar values. Temperature and precipitation are generally more similar between two near locations than between two far away ones. This relationship between the values of the locations of the same variable is the spatial autocorrelation, a special case of correlation in space. In a more formal definition: 'Given a set S containing n geographical units, spatial autocorrelation refers to the relationship between some variable observed in each of the n localities and a measure of geographical proximity defined for all n ($n-1$) pairs chosen from n ' (Getis 2008). Therefore, spatial autocorrelation is an important and necessary property between the species' locations and the environmental variables. If the species' locations do not have spatial autocorrelation for a specific environmental variable, this variable is meaningless for the species, as it lacks any relationship dependent on distance. In other words, there is no environmental gradient: all values of that variable are similar across space, or they vary without any relationship to distance. And without an environmental gradient, we cannot produce a model: the stronger the gradient, the greater the explanatory or predictive power of the model (Seoane *et al.* 2005). However, autocorrelation causes biases when evaluating model performance, and it affects model coefficients and statistical inference (Oliveira *et al.* 2014).

Spatial autocorrelation must be avoided between training and test data, in order to guarantee independence between both datasets (Peterson and Soberón 2012). Test data include both presences and (pseudo)absences, thus autocorrelation must be also avoided between both types of species occurrence data (Oliveira *et al.* 2014). Similarly,

autocorrelation must be avoided among model residuals, as a result of biases in the presence records (Gaspard *et al.* 2019). Indeed, an incorrect filtering of records may lead to spatial autocorrelation in the residuals. Yet, spatial autocorrelation between training and test data, as well as in model residuals, is not frequently analysed. These problems can be reduced by splitting the data into training and test datasets using recently proposed block cross-validation procedures (Roberts *et al.* 2017, Valavi *et al.* 2019) or by selecting sampling sites independent in geographical and environmental spaces (Oliveira *et al.* 2014).

The spatial resolution of records and variables should be the same

Most ENM studies use gridded environmental data, such as raster maps, as predictor variables (see, however, Booth *et al.* 2014). In such cases, the pixel size of environmental variables should correspond to the spatial resolution of the species records (Barbosa *et al.* 2012, Araújo *et al.* 2019, Feng *et al.* 2019, Sofaer *et al.* 2019). This, which seems trivial, is ignored in numerous modelling works. If species records are collected, for instance, from a distribution atlas with a spatial resolution of $10 \times 10 \text{ km}^2$ (Sillero *et al.* 2014, Bencatel *et al.* 2019), the environmental variables should not be used at a finer resolution of e.g. $\sim 1 \text{ km}^2$, such as the original resolution of WorldClim variables (Hijmans *et al.* 2005, Fick and Hijmans 2017): the spatial resolution of the environmental variables must fit the spatial resolution of the species records.

This also applies to point occurrence records that may actually represent centroids of larger grid cells, such as many records available in online databases such as GBIF. Even when the spatial error is not explicitly reported, a visual inspection of the species occurrence map often reveals that many points are evenly spaced (e.g. currently on GBIF, many European mammals, amphibians and reptiles show regularly spaced points 10 km apart in both longitude and latitude). This indicates that those points result from surveys within grid cells of a fixed size, and their spatial error or resolution is the regular distance between those points.

The error in the species coordinates should be equal or inferior to the spatial resolution of the environmental variables (Sillero and Gonçalves-Seco 2014). Otherwise, each occurrence record will count as if the species was observed exactly at the central pixel of the corresponding grid cell, while the environmental values at that pixel may be far from those at the place where the species was actually observed within the cell (e.g. a central valley in a mountainous region or vice-versa). Therefore, if species coordinates correspond to a grid of $10 \times 10 \text{ km}^2$, then the environmental variables should be rescaled or aggregated (e.g. by calculating the mean, median, mode, minimum, or maximum) at this resolution. A related issue in aggregating variables is failure to account for the associated error, which typically varies in space and is further compounded by spatial aggregation. Many submitted and published manuscripts do not account for the error propagation arising from the predictor variables. In any case, the way in which aggregation is performed should be reported (Moudrý *et al.* 2019).

Also, it is not necessarily useful to filter species records with a distance threshold (e.g. 10 km) corresponding to the spatial resolution of the variables (e.g. 10 km^2). Filtering records at such level will not resolve the clustering degree of the species data. Clustering level must be analysed at the spatial resolution of the environmental variables (in this

case, 10 km²) and not at a higher resolution (e.g. 1 km²). A good solution is to record species positions at a high resolution whenever possible, e.g. with a GPS: in that case, the species records can be used with any spatial resolution higher than the error measurement, normally between 2–5 m. If models are built with a very high spatial resolution, i.e. finer than 1 m (Sillero and Gonçalves-Seco 2014), using a GPS with very high accuracy, e.g. ~10 cm, is necessary. Unfortunately, this is not always possible, and in that case, it may be better to lower the variables' resolution. A good trade-off should ideally be found between the spatial resolution of the environmental variables and the positional accuracy of the species occurrences.

The spatial resolution of variables is not increased by resampling

Related to the previous point, it is not possible to directly increase the spatial resolution of the environmental variables, without access to the interpolated relationships used to estimate them and an elevation model with adequate spatial resolution (Hijmans *et al.* 2005, Fick and Hijmans 2017). If the spatial resolution of the variables is 1 km², transforming them directly into rasters with pixels of e.g. 100 m² will not increase the spatial resolution of the variables. This mistake is generally harder to spot, and we usually only found this out after asking authors how exactly they had downscaled their variables from e.g. WorldClim data to the finer resolution (e.g. 100 meters, or 30 meters to match the resolution of a digital elevation model) that they claimed in their manuscripts. This mistake can thus be more prevalent than is commonly perceived.

Many authors collect species occurrence data on fine local scales, and several studies show that fine-scale models can indeed produce better results at such scales if positional uncertainty is low (Kaliontzopoulou *et al.* 2008, Gottschalk *et al.* 2011, Moudrý and Šimová 2012). However, these fine scales are often not directly matched in the available climatic variables. Resampling allows to recalculate and assign pixel values when adjusting pixel size or orientation of a raster grid. Therefore, resampling only modifies the properties of the raster. This means that, when downscaling a raster from 1 km² to 100 m² of pixel size by resampling with another raster of 100 m² as template, the values of the pixels of the original raster will continue to be same: the result will be a raster with the same spatial pattern as a raster of 1 km, but with pixels of 100 m. The size (as in number of pixels) of the raster increases, but the amount and resolution of the information does not. Even if a bilinear interpolation is used, the result will be nearly the same, because most pixels will have the same value, except those near the borders between grid cells. A more correct way of downscaling climate variables is to use dynamical downscaling for regional models (e.g. CORDEX framework, <http://www.cordex.org/>; Rummukainen 2010), although any downscaling can entail errors and should thus be avoided unless strictly necessary.

Machine learning models should be replicated

Maxent, Random Forests, and Boosted Regression Trees are examples of machine learning methods, where any time a model is calculated for the same dataset of species records the result is slightly different (Phillips *et al.* 2006, 2017). In contrast, the results provided by e.g. Generalized Linear Models (GLM), Ecological Niche Factor Analyses (ENFA) or BIOCLIM are always identical for the same dataset of species records (Hirzel *et al.* 2002, Booth *et al.*

2014). Consequently, if a machine learning method is used, it is necessary to calculate the model several times for the same species dataset and provide at least the average and standard deviation of the models. This is even more important if the selection of training and test records is random. The objective is to evaluate if the results are stable across the sample of models. Depending on the available computing time and storage, the replication of models can range from a minimum of 10 to 50, 100, or even more (Phillips *et al.* 2006). Maxent and other software packages incorporate a function to indicate the number of replicates to be calculated, providing the average, median, maximum, minimum, and standard deviation of models. However, modellers often do not replicate their models, and the results may consequently not be robust.

Highly correlated variables should be excluded from the model

Another common mistake is to include a full set of variables, such as all 19 BIOCLIM variables (Nix 1986, Booth *et al.* 1988) available from WorldClim (Hijmans *et al.* 2005, Fick and Hijmans 2017), without critically analysing them and excluding redundant ones. Building models with highly correlated variables can have several undesired effects (Franklin 2010, Field *et al.* 2012, De Marco and Nóbrega 2018): the null hypothesis can be wrongly rejected; coefficients can significantly change and even reverse their sign; insignificant variables can be selected; the model can suffer from over-fitting, being adjusted excessively to the data (possibly reflecting noise); and it may be not possible to disentangle correctly the response curves for each variable, as each variable will interact with others, hampering to obtain the actual response curves. Collinearity affects models trained on data from one region or time and projected to another with a different or unknown structure of collinearity (see Dormann *et al.* 2013 for a review). However, truly independent response curves can only be obtained when using orthogonal variables, such as the results of a principal components analysis (PCA). In the real world, as Tobler's Law (Tobler 1970) indicates, everything is related to everything else, and therefore it is impossible to have completely uncorrelated variables: precipitation will be always related to temperature, and both of them to elevation, and so on. So, a compromise is to select among the most correlated variables for the analysis. Traditionally, the correlation threshold is around 0.7 or higher (Dormann *et al.* 2013), because going below is very difficult as the number of variables might become too low. Other approaches for dealing with collinearity are latent variable methods, shrinkage and regularisation (Dormann *et al.* 2013). Some methods, such as BIOCLIM (Nix 1986), ENFA (Hirzel *et al.* 2002), and Mahalanobis distance (Clark *et al.* 1993) can use all variables without being affected by correlations among them. BIOCLIM is an envelope modelling method and it uses only the limits defined by the species records, without assuming any relationship function or interactions between variables (Nix 1986). ENFA does not depend on the correlation between variables, because it is itself a Principal Components Analysis (PCA), transforming all variables into different orthogonal components with ecological meaning (Hirzel *et al.* 2002). However, using the components of a PCA as variables for a model complicates the interpretation of the results, as we will obtain the contribution of each component, and not of each variable. When models are projected to other scenarios in space or time, PCA cannot be computed directly for those new scenarios, as the collinearity structure between variables can change with space and time. Individual PCAs may be not

equivalent with different dimensions of predictors when transforming the variables over independent orthogonal spaces. Thus, a PCA for the baseline climate must be computed; using its coefficients to compute the components from the other scenarios. As is required in PCAs, all variables must be standardized and centred by using the parameters from the baseline climate. The most important variable for the model will be the first component, as it accounts for the highest variability of the set of variables. It will be necessary to analyse the contribution of each variable inside each principal component. In the remaining methods, it is necessary to evaluate the correlations between variables and exclude redundant ones (e.g. Leroy *et al.* 2014, Báez *et al.* 2019, Pereira *et al.* 2020). For example, in Maxent, using five variables often provides reasonable results (Peterson and Cohoon 1999, Cumming 2000). Occasionally, the inclusion of correlated variables in the model could be considered, as long as it is justified that both variables can have determinant effects on the species' niche (e.g., when a species is affected by both maximum and minimum temperatures). R package ENMTML provides some useful tools to assist in the reduction of correlation between variables (Andrade *et al.* 2020). Also, the Variable Inflation Factor (VIF) can aid in the selection of variables to be included in the models: the R package usdm provides useful tools for this issue (Naimi *et al.* 2014). Note, however, that a high VIF is no sufficient motive to exclude a particular variable; it may be better to remove other, less meaningful variable(s) that are correlated with it and that cause its VIF to be high

Unprojected coordinate systems should not be used for calculating topographical variables

This mistake is composed of three. It is frequent to use topographical variables with geographical coordinate systems, such as WGS84, instead of a projected coordinate system. If elevation is the only topographical variable, there is no problem with that. However, it is not possible to correctly calculate slope, aspect, or other topographical indexes from a geographical coordinate system (Burrough and McDonnell 1998). All these variables are calculated using distances. All geographical coordinate systems are placed over a spheroid and any distance calculated over a sphere corresponds to an arc-distance rather than a Euclidean distance. Arc-distances are larger than Euclidean distances. Therefore, it is necessary to transform the elevation map to a projected coordinate system, and from there calculate all remaining topographical variables.

Another common mistake is related to the calculation of aspect. Aspect defines the cardinal position of any pixel: north, south, east, west, and intermediate combinations (Burrough and McDonnell 1998). Aspect is measured in degrees, and here comes the problem: 360° and 0° actually represent the same value, north. However, the modelling method does not understand that those two values are the same. One solution is to use aspect to the north, ranging from 0 to 180, and aspect to the east with the same range.

Finally, current and future suitable ranges for species are sometimes compared by counting the number of pixels that were gained or lost in maps with geographical coordinate systems. Because these systems are not planar, estimated areas are erroneous, as distance and areas change visibly with latitude. Thus, if a species is distributed across a range of latitudes, then the estimates of range size change will be biased (Budic *et al.* 2016).

Other (relatively) common mistakes

A list of commonly found mistakes is necessarily subjective and dependent on the authors' personal experiences. Other mistakes can be spotted, although these are either less frequent (in our experience) or do not yet have a clear solution.

For example, presence-only and presence-background models are often evaluated with discrimination performance metrics, which assume presence-(pseudo) absence information. Although this is erroneous without a true estimate of species prevalence (Guillera-Aroita *et al.* 2014, Leroy *et al.* 2018), discrimination metrics are still in high demand for these models, and alternatives are not widespread (but see e.g. Boyce *et al.* 2002, Hirzel *et al.* 2006, Liu *et al.* 2013, Báez *et al.* 2019).

Another frequent mistake is to use inadequate (usually excessive) numbers of pseudo-absences or of background points when surveyed absence data are not available. However, the recommendations are not clear on which is the adequate number of points to use (which probably depends on the modelling technique), or these recommendations are based on discrimination metrics applied to non-presence-absence data (Barbet-Massin *et al.* 2012), which tend to favour overfit models.

An additional mistake found sometimes in studies of climate change effects stems from the existence of many pixels with missing data, mainly along coastlines, in future climate projections. When comparing the extents of current and future suitable ranges, missing data need to be the same between the different time periods, or else the extents are not comparable and range size change will be miscalculated.

Discussion

The effects on model outputs caused by these mistakes are various, from no problem at all to invalidating the model. In some cases, the error is merely conceptual: modelling outputs may be identical whether you consider Maxent, for instance, as a presence/pseudo-absence or a presence/background method (Phillips *et al.* 2017), or whether you call filtering a reduction of autocorrelation (Franklin 2010, Aiello-Lammens *et al.* 2015). However, interpreting background as pseudo-absence can have important consequences, either for delimiting the study area (Anderson and Raza 2010, Barve *et al.* 2011) or for evaluating model performance (Phillips *et al.* 2009, Golicher *et al.* 2012): the background includes all habitats considered available, whether or not they are occupied or suitable. Study areas should normally exclude areas where the species cannot disperse (Anderson and Raza 2010). Downscaling variables by resampling can also have a small effect on model outputs: the spatial pattern will be similar to the original variables.

The other three mistakes, however, can strongly alter the results. Using a finer spatial resolution in environmental variables when the species occurrence records are coarser can make the model completely erroneous. Including highly correlated variables can make models appear better than they actually are (Field *et al.* 2012, De Marco and Nóbrega 2018). Again, the conclusions can be wrong, which can be especially dangerous when using them for conservation or management decisions (Addison *et al.* 2013, Guisan *et al.* 2013). Unfortunately, we do not know the exact effects of using topographical variables calculated from geographical coordinate systems. To our knowledge, no studies have addressed this problem. The difference between using a geographical (in degrees)

or a projected (in metres) coordinate system is that distances obtained from the former are larger and inconsistent across latitudes. This can modify the general pattern of the variable, and thus the output of the model as well. Using aspect with values of 360 and 0 degrees will affect the model output depending on the frequency of north values. If there are few pixels with the value of 360, the result may be poorly affected. However, if the species occurs mostly in northern slopes, the method will probably fail to recognise the correct importance of aspect in the species distribution.

Most of the mistakes presented here are previous to publication and are usually corrected on time by instructors, reviewers and editors. Nevertheless, highlighting such frequent mistakes here can help save valuable time to both students and their supervisors and reviewers. However, at least two of these mistakes can also be found in several published papers, namely considering Maxent as a pseudo-absence method and deriving topographical variables from geographical coordinate systems. The latter is sometimes difficult to track if the methodology is not described sufficiently. Articles frequently do not specify the coordinate reference system, with ellipsoid and datum, of their spatial data, and one can suspect that if authors are using the altitude map included in the WorldClim dataset, topographical variables were also obtained directly from the geographical unprojected map. The remaining mistakes are most frequently found among ENM students.

Why do these mistakes occur? This is a question difficult to answer. We suppose it is necessary to read more, namely methodological papers, but also reviews and textbooks. Although ENMs are not recent, many methods are relatively new (Fitzpatrick *et al.* 2013). For example, Maxent was published only 13 years ago (Phillips *et al.* 2004, 2006). It is also worthy highlighting here that we have only three textbooks on ecological niche modelling, all of which are relatively recent as well (Franklin 2010, Peterson *et al.* 2011, Guisan *et al.* 2017). There are some other works providing guidelines on modelling, which are just starting to be cited (Barbosa *et al.* 2012, Guillera-Aroita *et al.* 2015, Jarnevich *et al.* 2015, Araújo *et al.* 2019, Feng *et al.* 2019, Sofaer *et al.* 2019). Also, there are numerous training courses on ecological niche models, but probably they are attended by a small proportion of modelling beginners. Therefore, a stronger focus on ecological niche modelling reviews is necessary, as well as more studies analysing the effects of these mistakes and proposing clearer validation methods. Although ecological niche models are established and widely used statistical methods, users still continue to make some elemental mistakes.

Other mistakes may probably be found, and may even be more common in other fields that use ENMs. The mistakes presented here result from our combined experience while lecturing nearly 20 ENM courses and reviewing or editing over 150 manuscripts. Therefore, this list is necessarily subjective and based on our own background in biogeography and spatial ecology. Regardless of other common mistakes that other researchers and lecturers may come across, we need to increase our efforts to avoid the recurrence of methodological mistakes in the field of ENMs, in order to increase the reproducibility and reliability of ENM studies.

Acknowledgments

We thank six anonymous reviewers for useful and constructive comments on an earlier version of the manuscript. NS is supported by CEEC2017 contracts from FCT (CEECIND/02213/2017).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by CEEC 2017 contracts from FCT [CEECIND/02213/2017].

Notes on contributors

Neftalí Sillero works in the analysis and identification of biodiversity spatial patterns, from species to populations and individuals, applying new technologies on species' distributions atlases, ecological modelling of species' ranges, home ranges and road ecology.

A. Márcia Barbosa does research and training on biogeography, macroecology, species distribution and ecological niche modelling, biodiversity patterns, and their applications to conservation and management.

ORCID

Neftalí Sillero  <http://orcid.org/0000-0002-3490-3780>

References

- Addison, P.F.E., *et al.*, 2013. Practical solutions for making models indispensable in conservation decision-making. *Diversity and Distributions*, 19 (5–6), 490–502. doi:10.1111/ddi.12054.
- Aiello-Lammens, M.E., *et al.*, 2015. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, 38, 1–35. doi:10.1111/ecog.01132.
- Anderson, R.P. and Raza, A., 2010. The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography*, 37, 1378–1393. doi:10.1111/j.1365-2699.2010.02290.x.
- Andrade, A.F.A., Velazco, S.J.E., and De Marco Júnior, P., 2020. ENMTML: an R package for a straightforward construction of complex ecological niche models. *Environmental Modelling Software*, 125, 104615. doi:10.1016/j.envsoft.2019.104615.
- Araújo, M.B., *et al.*, 2019. Standards for distribution models in biodiversity assessments. *Science Advances*, 5, eaat4858. doi:10.1126/sciadv.aat4858.
- Baek, S., Kim, M., and Lee, J., 2019. Current and future distribution of *Ricania shantungensis* (Hemiptera: ricaniidae) in Korea: application of spatial analysis to select relevant environmental variables for MaxEnt and CLIMEX MODELING. *Forests*, 10 (490), 1–14. doi:10.3390/f10060490.
- Báez, J.C., *et al.*, 2019. Ensemble modeling of the potential distribution of the whale shark in the Atlantic Ocean. *Ecology and Evolution*, 10, 175–184.
- Barbet-Massin, M., *et al.*, 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, 3, 327–338. doi:10.1111/j.2041-210X.2011.00172.x.
- Barbosa, A.M., *et al.*, 2010. Positive regional species–people correlations: a sampling artefact or a key issue for sustainable development? *Animal Conservation*, 13, 446–447. doi:10.1111/j.1469-1795.2010.00402.x.
- Barbosa, A.M., Pautasso, M., and Figueiredo, D., 2013. Species–people correlations and the need to account for survey effort in biodiversity analyses. *Diversity and Distributions*, 19 (9), 1188–1197. doi:10.1111/ddi.12106.

- Barbosa, A.M., Real, R., and Vargas, J.M., 2009. Transferability of environmental favourability models in geographic space: the case of the Iberian desman (*Galemys pyrenaicus*) in Portugal and Spain. *Ecological Modelling*, 220, 747–754. doi:10.1016/j.ecolmodel.2008.12.004.
- Barbosa, M.A., et al., 2012. Ecological niche models in Mediterranean herpetology: past, present and future. In: W. Zhang, ed.. *Ecological modeling*, New York: Nova Science Publishers, Inc., 173–204.
- Barve, N., et al., 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, 222, 1810–1819. doi:10.1016/j.ecolmodel.2011.02.011.
- Bencatel, J., et al., 2019. *Atlas de Mamíferos de Portugal*. 2nd ed. Évora, Portugal: Universidade de Évora, 271.
- Beukema, W., et al., 2013. Review of the systematics, distribution, biogeography and natural history of Moroccan amphibians. *Zootaxa*, 3661, 1–60. doi:10.11646/zootaxa.3661.1.1.
- Booth, T.H., et al., 1988. Niche analysis and tree species introduction. *Forest Ecology and Management*, 23, 47–59. doi:10.1016/0378-1127(88)90013-8.
- Booth, T.H., et al., 2014. BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. *Diversity and Distributions*, 20 (1), 1–9. doi:10.1111/ddi.12144.
- Booth, T.H., 2018. Why understanding the pioneering and continuing contributions of BIOCLIM to species distribution modelling is important. *Austral Ecology*, 43 (8), 852–860.
- Boria, R.A., et al., 2014. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, 275, 73–77. doi:10.1016/j.ecolmodel.2013.12.012.
- Boyce, M.S., et al., 2002. Evaluating resource selection functions. *Ecological Modelling*, 157, 281–300. doi:10.1016/S0304-3800(02)00200-4.
- Brito, J.C., et al., 2011. Biogeography and conservation of viperids from North-West Africa: an application of ecological niche-based models and GIS. *Journal of Arid Environment*, 75, 1029–1037. doi:10.1016/j.jaridenv.2011.06.006.
- Budic, L., Didenko, G., and Dormann, C.F., 2016. Squares of different sizes: effect of geographical projection on model parameter estimates in species distribution modeling. *Ecology and Evolution*, 6, 202–211. doi:10.1002/ece3.1838.
- Burrough, P.A. and McDonnell, R., 1998. *Principles of geographical information systems for land resources assessment*. Oxford: Oxford University Press.
- Cai, T., et al., 2019. Analyzing stopover and wintering habitats of Hooded Cranes (*Grus monacha*): implications for conservation and species dispersion in the East Asia. *Pakistan Journal of Zoology*, 51 (4). doi:10.17582/journal.pjz/2019.51.4.1323.1333.
- Chapman, D., et al., 2019. Improving species distribution models for invasive non-native species with biologically informed pseudo-absence selection. *Journal of Biogeography*, 46 (5), 1029–1040. doi:10.1111/jbi.13555.
- Chefaoui, R.M., et al., 2015. Large-scale prediction of seagrass distribution integrating landscape metrics and environmental factors: the case of *Cymodocea nodosa* (Mediterranean–Atlantic). *Estuaries and Coasts*, 39, 123–137. doi:10.1007/s12237-015-9966-y.
- Clark, J.D., Dunn, J.E., and Smith, K.G., 1993. A multivariate model of female black bear habitat use for a geographic information system. *Journal of Wildlife Management*, 57 (3), 519–526. doi:10.2307/3809276.
- Cumming, G.S., 2000. Using between-model comparisons to fine-tune linear models of species ranges. *Journal of Biogeography*, 2, 441–455. doi:10.1046/j.1365-2699.2000.00408.x.
- De Marco, P., Diniz-Filho, J.A.F., and Bini, L.M., 2008. Spatial analysis improves species distribution modelling during range expansion. *Biological Letters*, 4 (5), 577–580. doi:10.1098/rsbl.2008.0210.
- De Marco, P. and Nóbrega, C.C., 2018. Evaluating collinearity effects on species distribution models: an approach based on virtual species simulation. *PLoS ONE*, 13 (9), e0202403. doi:10.1371/journal.pone.0202403.
- Dormann, C.F., 2007. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography*, 2, 129–138. doi:10.1111/j.1466-8238.2006.00279.x.

- Dormann, C.F., *et al.*, 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 5, 609–628. doi:[10.1111/j.2007.0906-7590.05171.x](https://doi.org/10.1111/j.2007.0906-7590.05171.x).
- Dormann, C.F., *et al.*, 2013. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36, 027–046. doi:[10.1111/j.1600-0587.2012.07348.x](https://doi.org/10.1111/j.1600-0587.2012.07348.x).
- Farrell, A., *et al.*, 2019. Machine learning of large-scale spatial distributions of wild turkeys with high-dimensional environmental data. *Ecology and Evolution*, 9 (10), 5938–5949. doi:[10.1002/ece3.5177](https://doi.org/10.1002/ece3.5177).
- Feng, X., *et al.*, 2019. A checklist for maximizing reproducibility of ecological niche models. *Nature Ecology & Evolution*, 3, 1382–1395. doi:[10.1038/s41559-019-0972-5](https://doi.org/10.1038/s41559-019-0972-5).
- Ficetola, G.F., *et al.*, 2014. An evaluation of the robustness of global amphibian range maps. *Journal of Biogeography*, 41, 211–221. doi:[10.1111/jbi.12206](https://doi.org/10.1111/jbi.12206).
- Fick, S. and Hijmans, R., 2017. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302–4315. doi:[10.1002/joc.5086](https://doi.org/10.1002/joc.5086).
- Field, A., Miles, J., and Field, Z., 2012. *Discovering statistics using R*. London: Sage Publications.
- Fitzpatrick, M.C., Gotelli, N.J., and Ellison, A.M., 2013. MaxEnt versus MaxLike: empirical comparisons with ant species distributions. *Ecosphere*, 4 (5), art55. Article 55. doi:[10.1890/ES13-00066.1](https://doi.org/10.1890/ES13-00066.1).
- Franklin, J., 2010. *Mapping Species Distributions*. New York: Cambridge University Press.
- Gábor, L., *et al.*, 2019. How do species and data characteristics affect species distribution models and when to use environmental filtering? *International Journal of Geographical Information Science*, 34 (8), 1–18.
- Gaspard, G., Kim, D., and Chun, Y., 2019. Residual spatial autocorrelation in macroecological and biogeographical modeling: a review. *Journal of Ecology and Environment*, 43, 19. doi:[10.1186/s41610-019-0118-3](https://doi.org/10.1186/s41610-019-0118-3).
- Getis, A., 2008. A history of the concept of spatial autocorrelation: a geographer's perspective. *Geographical Analysis*, 40, 297–309. doi:[10.1111/j.1538-4632.2008.00727.x](https://doi.org/10.1111/j.1538-4632.2008.00727.x).
- Golicher, D., *et al.*, 2012. Pseudo-absences, pseudo-models and pseudo-niches: pitfalls of model selection based on the area under the curve. *International Journal of Geographical Information Science*, 26 (11), 2049–2063. doi:[10.1080/13658816.2012.719626](https://doi.org/10.1080/13658816.2012.719626).
- Gottschalk, T.K., *et al.*, 2011. Influence of grain size on species–habitat models. *Ecological Modelling*, 222, 3403–3412. doi:[10.1016/j.ecolmodel.2011.07.008](https://doi.org/10.1016/j.ecolmodel.2011.07.008).
- Guillera-Aroita, G., *et al.*, 2015. Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, 24, 276–292. doi:[10.1111/geb.12268](https://doi.org/10.1111/geb.12268).
- Guillera-Aroita, G., Lahoz-Monfort, J.J., and Elith, J., 2014. Maxent is not a presence-absence method: A comment on Thibaud *et al.* *Methods in Ecology and Evolution*, 5 (11), 1192–1197. doi:[10.1111/2041-210X.12252](https://doi.org/10.1111/2041-210X.12252).
- Guisan, A., *et al.*, 2013. Predicting species distributions for conservation decisions. *Ecology Letters*, 16 (12), 1424–1435. doi:[10.1111/ele.12189](https://doi.org/10.1111/ele.12189).
- Guisan, A., Thuiller, W., and Zimmermann, N.E., 2017. *Habitat suitability and distribution models: with applications in R*. Cambridge: Cambridge University Press.
- Hallgren, W., *et al.*, 2019. Species distribution models can be highly sensitive to algorithm configuration. *Ecological Modelling*, 408, 108719. doi:[10.1016/j.ecolmodel.2019.108719](https://doi.org/10.1016/j.ecolmodel.2019.108719).
- Hijmans, R.J., *et al.*, 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25 (15), 1965–1978. doi:[10.1002/joc.1276](https://doi.org/10.1002/joc.1276).
- Hirzel, A.H., *et al.*, 2002. Ecological-niche factor analysis: how to compute habitat suitability maps without absence-data? *Ecology*, 7 (83), 2027–2036. doi:[10.1890/0012-9658\(2002\)083\[2027:ENFAHT\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[2027:ENFAHT]2.0.CO;2).
- Hirzel, A.H., *et al.*, 2006. Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, 199, 142–152. doi:[10.1016/j.ecolmodel.2006.05.017](https://doi.org/10.1016/j.ecolmodel.2006.05.017).
- Hutchinson, G.E., 1957. Concluding remarks. In: *Cold spring harbour symposium on quantitative biology*, (22), 415–427.
- Iturbide, M., Bedia, J., and Gutiérrez, J.M., 2018. Background sampling and transferability of species distribution model ensembles under climate change. *Global and Planetary Change*, 166, 19–29. doi:[10.1016/j.gloplacha.2018.03.008](https://doi.org/10.1016/j.gloplacha.2018.03.008).

- Jarnevich, C.S., *et al.*, 2015. Caveats for correlative species distribution modeling. *Ecological Informatics*, 29 (P1), 6–15. doi:[10.1016/j.ecoinf.2015.06.007](https://doi.org/10.1016/j.ecoinf.2015.06.007).
- Kadmon, R., Farber, O., and Danin, A., 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, 14 (2), 401–413. doi:[10.1890/02-5364](https://doi.org/10.1890/02-5364).
- Kaliontzopoulou, A., *et al.*, 2008. Modelling the partially unknown distribution of wall lizards (*Podarcis*) in North Africa: ecological affinities, potential areas of occurrence, and methodological constraints. *Canadian Journal of Zoology*, 86, 992–1001. doi:[10.1139/Z08-078](https://doi.org/10.1139/Z08-078).
- Karger, D.N., *et al.*, 2017. Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4, 1–20. doi:[10.1038/sdata.2017.122](https://doi.org/10.1038/sdata.2017.122).
- Kriticos, D.J., *et al.*, 2012. CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution*, 3 (1), 53–64. doi:[10.1111/j.2041-210X.2011.00134.x](https://doi.org/10.1111/j.2041-210X.2011.00134.x).
- Leroy, B., *et al.*, 2014. Forecasted climate and land use changes, and protected areas: the contrasting case of spiders. *Diversity and Distributions*, 20, 686–697. doi:[10.1111/ddi.12191](https://doi.org/10.1111/ddi.12191).
- Leroy, B. *et al.*, 2018. Without quality presence–absence data, discrimination metrics such as TSS can be misleading measures of model performance. *Journal of Biogeography*, 45 (9), 1994–2002.
- Lima-Ribeiro, M.S., 2015. EcoClimate: a database of climate data from multiple models for past, present, and future for macroecologists and biogeographers. *Biodiversity Informatics*, 10, 1–21. doi:[10.17161/bi.v10i04.4955](https://doi.org/10.17161/bi.v10i04.4955).
- Liu, C., White, M., and Newell, G., 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, 40, 778–789. doi:[10.1111/jbi.12058](https://doi.org/10.1111/jbi.12058).
- Mariani, M., *et al.*, 2019. Climate change reduces resilience to fire in subalpine rainforests. *Global Change Biology*, 25 (6), 2030–2042. doi:[10.1111/gcb.14609](https://doi.org/10.1111/gcb.14609).
- Moudrý, V., *et al.*, 2019. Potential pitfalls in rescaling digital terrain model-derived attributes for ecological studies. *Ecological Informatics*, 54, 100987. doi:[10.1016/j.ecoinf.2019.100987](https://doi.org/10.1016/j.ecoinf.2019.100987).
- Moudrý, V. and Šimová, P., 2012. Influence of positional accuracy, sample size and scale on modelling species distributions: A review. *International Journal of Geographical Information Science*, 26, 2083–2095. doi:[10.1080/13658816.2012.721553](https://doi.org/10.1080/13658816.2012.721553).
- Naimi, B., *et al.*, 2014. Where is positional uncertainty a problem for species distribution modelling? *Ecography*, 37, 191–203. doi:[10.1111/j.1600-0587.2013.00205.x](https://doi.org/10.1111/j.1600-0587.2013.00205.x).
- Nix, H.A., 1986. A biogeographic analysis of Australian Elapid Snakes. In: R. Longmore, ed. *Atlas of Elapid Snakes of Australia*. Canberra: Australian Flora and Fauna Series Number 7. Australian Government Publishing Service, 4–15.
- Nogues-Bravo, D., *et al.*, 2008. Climate change, humans, and the extinction of the Woolly Mammoth. *PLoS Biology*, 4 (6), e79. doi:[10.1371/journal.pbio.0060079](https://doi.org/10.1371/journal.pbio.0060079).
- Oliveira, G., *et al.*, 2014. Evaluating, partitioning, and mapping the spatial autocorrelation component in ecological niche modeling: A new approach based on environmentally equidistant records. *Ecography*, 37 (7), 637–647. doi:[10.1111/j.1600-0587.2013.00564.x](https://doi.org/10.1111/j.1600-0587.2013.00564.x).
- Papier, C.M., Poulos, H.M., and Kusch, A., 2019. Invasive species and carbon flux: the case of invasive beavers (*Castor canadensis*) in riparian *Nothofagus* forests of Tierra del Fuego, Chile. *Climatic Change*, 153 (1–2), 219–234. doi:[10.1007/s10584-019-02377-x](https://doi.org/10.1007/s10584-019-02377-x).
- Pereira, P.F., *et al.*, 2020. The spread of the red-billed leiothrix (*Leiothrix lutea*) in Europe: the conquest by an overlooked invader? *Biological Invasions*, 22, 709–722. doi:[10.1007/s10530-019-02123-5](https://doi.org/10.1007/s10530-019-02123-5).
- Peterson, A.T., *et al.*, 2011. *Ecological niche and geographical distributions*. New Jersey: Princeton University Press.
- Peterson, A.T. and Cohoon, K.P., 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecological Modelling*, 117, 159–164. doi:[10.1016/S0304-3800\(99\)00023-X](https://doi.org/10.1016/S0304-3800(99)00023-X).
- Peterson, A.T. and Soberón, J., 2012. Species distribution modeling and ecological niche modeling: getting the concepts right. *Natureza & Conservação*, 10 (2), 102–107. doi:[10.4322/natcon.2012.019](https://doi.org/10.4322/natcon.2012.019).

- Phillips, S.J., *et al.*, 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, 19 (1), 181–197. doi:10.1890/07-2153.1.
- Phillips, S.J., *et al.*, 2017. Opening the black box: an open-source release of Maxent. *Ecography*, 40 (7), 887–893. doi:10.1111/ecog.03049.
- Phillips, S.J., Anderson, R.P., and Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190 (3–4), 231–259. doi:10.1016/j.ecolmodel.2005.03.026.
- Phillips, S.J., Dudík, M., and Schapire, R.E., 2004. A maximum entropy approach to species distribution modeling. In: *Proceedings of the Twenty-First International Conference on Machine Learning*, Banff, Canada, 655–662.
- Roberts, D.R., *et al.*, 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40 (8), 913–929. doi:10.1111/ecog.02881.
- Rummukainen, M., 2010. State-of-the-art with regional models. *WIREs Climate Change*, 1, 82–96. doi:10.1002/wcc.8.
- Segurado, P., Araújo, M.B., and Kunin, W.E., 2006. Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology*, 43 (3), 433–444. doi:10.1111/j.1365-2664.2006.01162.x.
- Seoane, J., *et al.*, 2005. Species-specific traits associated to prediction errors in bird habitat suitability modelling. *Ecological Modelling*, 2–4 (185), 299–308. doi:10.1016/j.ecolmodel.2004.12.012.
- Sillero, N., 2011. What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling*, 222, 1343–1346. doi:10.1016/j.ecolmodel.2011.01.018.
- Sillero, N., *et al.*, 2014. Updated distribution and biogeography of amphibians and reptiles of Europe. *Amphibia-Reptilia*, 35 (1), 1–31. doi:10.1163/15685381-00002935.
- Sillero, N. and Gonçalves-Seco, L., 2014. Spatial structure analysis of a reptile community with airborne LiDAR data. *International Journal of Geographical Information Science*, 28 (8), 1709–1722. doi:10.1080/13658816.2014.902062.
- Sofaer, H.R., *et al.*, 2019. Development and delivery of species distribution models to inform decision-making. *BioScience*, 69 (7), 544–557. doi:10.1093/biosci/biz045.
- Title, P.O. and Bemmels, J.B., 2018. ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*, 41 (2), 291–307. doi:10.1111/ecog.02880.
- Tobler, A.W.R., 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46 (Supplement: Proceedings International Geographical Union), 234–240. doi:10.2307/143141.
- Valavi, R., *et al.*, 2019. blockCV: an r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods Ecology and Evolution*, 10, 225–232. doi:10.1111/2041-210X.13107.
- Varela, S., *et al.*, 2014. Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography*, 37, 1084–1091.
- Werkowska, W., *et al.*, 2017. A practical overview of transferability in species distribution modeling. *Environment Reviews*, 133 (May 2016), 127–133. doi:10.1139/er-2016-0045.
- Yates, K.L., *et al.*, 2018. Outstanding challenges in the transferability of ecological models. *Trends in Ecology and Evolution*, 33 (10), 790–802. doi:10.1016/j.tree.2018.08.001.
- Yesson, C., *et al.*, 2007. How global is the global biodiversity information facility? *PLoS ONE*, 11 (2), e1124. doi:10.1371/journal.pone.0001124.
- Young, N.E., *et al.*, 2019. Finding the needle in the haystack: iterative sampling and modeling for rare taxa. *Journal of Insect Conservation*, 23 (3), 589–595. doi:10.1007/s10841-019-00151-z.
- Yue, S., Bonebrake, T.C., and Gibson, L., 2019. Informing snake roadkill mitigation strategies in Taiwan using citizen science. *Journal of Wildlife Management*, 83 (1), 80–88. doi:10.1002/jwmg.21580.
- Zhang, Z., *et al.*, 2019. Using species distribution model to predict the impact of climate change on the potential distribution of Japanese whiting *Sillago japonica*. *Ecological Indicators*, 104, 333–340. doi:10.1016/j.ecolind.2019.05.023.