A checklist for maximizing reproducibility of ecological niche models

Xiao Feng [1,2,8]*, Daniel S. Park [3,8], Cassondra Walker[4], A. Townsend Peterson[5], Cory Merow[6] and Monica Papeș[7]

Reporting specific modelling methods and metadata is essential to the reproducibility of ecological studies, yet guidelines rarely exist regarding what information should be noted. Here, we address this issue for ecological niche modelling or species distribution modelling, a rapidly developing toolset in ecology used across many aspects of biodiversity science. Our quantitative review of the recent literature reveals a general lack of sufficient information to fully reproduce the work. Over two-thirds of the examined studies neglected to report the version or access date of the underlying data, and only half reported model parameters. To address this problem, we propose adopting a checklist to guide studies in reporting at least the minimum information necessary for ecological niche modelling reproducibility, offering a straightforward way to balance efficiency and accuracy. We encourage the ecological niche modelling community, as well as journal reviewers and editors, to utilize and further develop this framework to facilitate and improve the reproducibility of future work. The proposed checklist framework is generalizable to other areas of ecology, especially those utilizing biodiversity data, environmental data and statistical modelling, and could also be adopted by a broader array of disciplines.

Science is facing a reproducibility crisis. A recent *Nature* survey of 1,576 researchers from various disciplines found that more than 70% of researchers were unable to reproduce research by others, and 50% were not even able to reproduce their own results[1]. Indeed, the issue of reproducibility has been raised across many fields of science. For instance, the estimates of non-reproducible studies are as high as 89% in cancer research[2] and 65% in drug research[3], and even high-profile, 'landmark' studies are not free of reproducibility issues[4]. New scientific research builds on previous efforts, allowing methods for testing hypotheses to evolve continually[5]. Therefore, research results must be communicated with enough context, detail and circumstance to allow correct interpretation, understanding and, whenever possible, reproduction. Reproducibility is a cornerstone of the scientific process and must be emphasized in scientific reports and publications. Although best-practice guidelines have been published and adopted for areas such as computer science[6] and clinical research[7,8], for various reasons, guidelines for ensuring reproducibility are still largely absent in many (even large) research communities.

Along these lines, the issue of reproducibility may be especially difficult to address in ecology, given the less-controlled aspects of many studies (for example, natural community surveys, field experiments). The issue of reproducibility has been noted only recently in ecology[9,10], but is likely prominent[11,12]. Because ecological studies often encompass uncontrollable or unaccountable factors[13], it is especially important to report in detail the circumstances and methods that apply. Furthermore, ecological studies often depend on statistical models, such that reporting specific modelling methods and decisions and how they are intended to reflect biological knowledge or assumptions holds particular importance for reproducibility in ecology[14,15]. More than ever before, it has become critical to report

these aspects, as the data and analytical tools underlying ecological studies are accumulating and evolving at an unprecedented rate in the age of big data[16]; ecological niche modelling (ENM) is a prominent example.

Ecological niche modelling

Also known as species distribution modelling (SDM)[17–19], ENM uses associations between known occurrences of species and environmental conditions to estimate species' potential geographic distributions. Although ENM and SDM are often used interchangeably in the literature[20], ENM typically has a stronger focus on estimating parameters of fundamental ecological niches, whereas SDM is more focused on geographic distributions of species. ENM is widely applied across many aspects of ecology and evolution, and is increasingly incorporated in decision-making regarding land use and conservation[21]. ENM studies are proliferating rapidly; in particular, a popular ENM algorithm, Maxent[22], has been cited in tens of thousands of research papers in the past decade alone. Though methods and assumptions in these studies vary greatly, to our knowledge, no evaluation of reproducibility of ENM or SDM studies has been conducted to date (but see ref. [21] for scoring key model aspects for biodiversity assessments). Furthermore, no guidelines on reporting essential modelling parameters exist, hindering accurate evaluation (for example, scoring[21]) of model methodology and reuse of published research. It is concerning that such a fast-growing and fast-evolving body of literature lacks assessment and guidelines for reproducibility.

Typically, ENM analyses take biodiversity data and environmental data (such as point observations of a species and climate) as input and use correlative or machine-learning methods to quantify underlying relationships, which then are used in making spatial

[1]Institute of the Environment, University of Arizona, Tucson, AZ, USA. [2]School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, USA. [3]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA. [4]Department of Integrative Biology, Oklahoma State University, Stillwater, OK, USA. [5]Biodiversity Institute, University of Kansas, Lawrence, KS, USA. [6]Ecology and Evolutionary Biology, University of Connecticut, Storss, CT, USA. [7]Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN, USA. [8]These authors contributed equally: Xiao Feng, Daniel S. Park. *e-mail: fengxiao@e-mail.arizona.edu

predictions. This typical workflow of ENM — obtaining and processing data, model calibration, model transfer and evaluation — is shared widely across disciplines that rely on statistical models. Therefore, the fast development, broad use and application, and existence of a rather established workflow for ENM makes it an excellent and representative example to tackle the challenges of reproducibility. Here, we assess the reproducibility of ENM studies via a comprehensive literature review and introduce a checklist to facilitate reproducibility of ENMs that can be extended to other areas of ecological research or other disciplines.

A checklist for ecological niche modelling

Although the role of 'methods' sections of scientific publications is to provide information that makes the study replicable, they are often highly condensed and lacking details needed for reproducibility, owing in large part to space limitations in journals. What is needed is a standardized format for reporting the full suite of details that comprise the critical information to ensure reproducibility. Therefore, a compendium of crucial parameters and qualities — in effect a metadata standard for ecological niche models — would be highly useful. A metadata standard establishes a common use and understanding through defining a series of attributes and standardized terminology to describe them. Such standards have been applied in various fields, such as GeoTIFF for spatial rasters[23] and Darwin Core and Humboldt Core for biodiversity data[24,25]. A metadata standard can provide a straightforward way to balance efficiency and accuracy in facilitating research reproducibility[26] in ENM, as well as scientific studies in general[27–29].

Here we present a checklist for ENM, to demonstrate how to define general and flexible reproducibility standards that can be used across a wide range of sub-fields of ecology. We compiled a list of essential elements required to reproduce ENM results based on the literature to date, and organized the elements into four major topics: (A) occurrence data collection and processing, (B) environmental data collection and processing, (C) model calibration and (D) model transfer and evaluation (labels correspond to elements in Table 1). We justify the design of the checklist briefly, and provide detailed definitions, examples of reporting for each element, and related literature, in Table 1. We do not distinguish the relative importance among the checklist elements, as all are necessary to assure full reproducibility. We provide a template of the checklist for easier use (Supplementary Table 1). We envision this checklist as a dynamic entity that will continue to be developed and refined by the ENM/SDM community to keep pace with the state of the art in the field. We also provide access to the checklist on Github, as an open-source project where users can comment and suggest changes (https://github.com/shandongfx/ENMchecklist or https://doi.org/10.5281/zenodo.3257732).

Occurrence data (A). Across many fields, online databases are growing and changing rapidly[30], such that reporting data versions or providing complete datasets used in analyses is crucial to reproducibility. Occurrence data are increasingly available owing to mass digitization of museum specimens and increased interest and participation in observational data collection by citizen scientists[31]. Because the quality of occurrence data can vary significantly among data sources, data types and taxa[32–34], it is vital to record data curation details to assure consistent quality and accuracy. The first attribute to report is the source of the data (A1; labels correspond to elements in Table 1 hereafter). If the occurrence data were the result of an online database query, the Digital Object Identifier (DOI), query and download date, or the version of a database must also be reported (A2), as online biodiversity data are accumulating rapidly and these data are often edited, corrected, improved or excluded over time[35–37]. The final dataset (that is, after editing and quality control), with the exception of sensitive information (for example,

specific locations of endangered taxa), should be deposited in a data archive when reserving rights allow it, thereby assuring reproducibility in case of changes to the original data source.

Whenever available, the 'basis of record' (A3) as used in Darwin Core, a community-developed standard for sharing biodiversity data[24], should be reported. This field describes how records were originally collected, and thus can indicate different levels of quality and different auxiliary information available. For instance, 'MachineObservation' via automated identification may be more prone to error compared with a 'PreservedSpecimen' collected and identified by an expert and deposited in a museum. Further, with a deposited specimen and catalogue number, researchers have the opportunity to examine the specimen physically to verify the identification[38,39], whereas an observation may not be verifiable. Spatial uncertainty (see A6-3) can vary with the type of occurrence data, as well as the time when the data were collected. For example, coordinates associated with older 'PreservedSpecimens' are usually georeferenced from descriptions of administrative units (for example, township, county or country), thus involving higher spatial uncertainty, whereas coordinates linked to recent 'HumanObservations' may have been directly reported from GPS devices, making them more accurate. Information regarding the uncertainty of occurrences can also facilitate evaluation of whether the spatial resolution of environmental data utilized is appropriate (see B3). The spatial uncertainty in biodiversity data has long been recognized[40,41], though the quantification of such uncertainty has not been implemented systematically at large scale (thus A6-3 was excluded from our literature review; see below); this task could be facilitated by recently developed informatics tools[42,43].

Increasingly, ecological research uses data from large-scale data aggregators (for example, the Global Biodiversity Information Facility (GBIF)). As with many sciences relying on observational, rather than design-based data collection, biodiversity data used in ENM have generally not been collected explicitly for this purpose. Thus, the spatial and temporal attributes of occurrences, and how they have been parsed or filtered in preparation for modelling, are essential details required to model ecological niches adequately[17]. Checking the extent of occurrences (A4) against expert-defined distributions (for example, regional floras) may reduce errors in identification or data transcription. Underrepresentation of the known distribution may suggest inadequate or biased sampling of occurrences, whereas spatial outliers may represent recent range expansion[44,45], occasional or vagrant occurrences[46], sink populations[47], or errors of identification or georeferencing. The collection date of occurrence records may influence spatial accuracy; in general, records from before the 1980s will lack precise point location data (that is, GPS coordinates) and are often georeferenced by hand from locality descriptions and with less precision[33]. Also, because environments change over time (for example, seasonal change, climate change, land-use changes), the temporal range of the occurrence data (A5) must be specified to connect it appropriately to the temporal dimension of environmental conditions[48]. Often, occurrence data are processed before modelling (A6). Common procedures include removing duplicate coordinates, excluding spatial and/or environmental outliers, and eliminating records with high spatial uncertainty[43] or erroneous coordinates[49]. Additionally, scholars have proposed various ways to address the well-known issues of sampling bias[50–52] and spatial autocorrelation[53], often by imposing distance-based filters on occurrence data or incorporating spatial structure as a component in the modelling process[54] (A7).

Environmental data (B). Similar to occurrence data, sources for environmental data are numerous, and data often require processing before inclusion in ENM analyses. The source (B1), and database query/download date or version of the database must be reported (B2), as environmental data may be updated periodically (for example,

WorldClim[55,56]) or may accumulate new data regularly through time (for example, PRISM[57]). Such information is also important for environmental variables derived from remotely sensed data (such as MODIS, Landsat). For example, NASA conducts regular quality assessments of MODIS data products and reprocesses data that may have been influenced by algorithm or calibration issues[58].

The spatial resolution of the environmental variables used (B3) can affect ENM results, as different ecological processes occur at different spatial scales[59]. It has been hypothesized that at broad scales, abiotic conditions have a more dominant role in determining species' distributions than biotic conditions[60,61], though increasing numbers of reported exceptions suggest that this pattern is context dependent[62,63]. In practice, using different spatial resolutions of environmental variables can produce different results[64–66]. Reporting the spatial resolution of environmental variables can also facilitate checking the match or mismatch with the spatial uncertainty of occurrences, given that coordinates are at times georeferenced from county centroids at coarse resolution[33]. In addition to reporting the spatial resolution used for

modelling, aggregation or disaggregation methods used to align the spatial resolutions of variables (for example, if they came from different data providers) should also be reported.

Providing the temporal range covered by the environmental variables (B4) is important for two reasons[67,68]. First, shorter temporal ranges can capture finer variation of environments (for example, extremes of daily temperature[69]), whereas longer temporal ranges capture longer-term trends in environmental conditions (for example, temperature seasonality). Second, it is helpful to evaluate how the temporal range of environmental data relates to the temporal range of occurrence data. For instance, associating occurrence data with environmental data from completely different time periods (for example, Last Glacial Maximum versus present) could be problematic, though the environmental data may need to include time lags to correspond to the life history of particular species[70]. The same reporting should be applied to information on future or past environments, as appropriate (D9–12). Similarly, the details of methods for processing and resampling of environmental data in temporal dimensions should also be reported.

Table 1 | Details of the ENM checklist and representation of its elements (percentage) in a review of recent ecology and evolution literature (2017–2018; 163 papers)

Category	What to report	Why reporting this element is important	Exemplar papers reporting the element	Relevant references	Papers (%)
		(A) Obtaining and processing occurrence data			
Metadata	(A1) Source of occurrence data	Reporting occurrence data sources allows one to assess data quality and trace/correct any possible issues that may be detected.	"Species distribution records were collected from the Ocean Biogeographic Information System (OBIS; http://iobis.org, accessed February 2016), from the Global Biodiversity Information Facility (GBIF; http://gbif.org, accessed January 2016), the Reef Life Survey (RLS; http://reeflifesurvey.com, accessed February 2016) and for a few species via personal communications."[109]	NA	93
	(A2) Download date; version of data source	Databases and datasets change over time.	"Occurrences were downloaded from GBIF.org on 28 January 2016 (https://doi.org/10.15468/dl.iou7qq)."[110]	NA	22
	(A3) Basis of records	Biodiversity databases comprise many different types of data, each with specific uses and caveats. Relevant distinctions include whether data are collected opportunistically, as part of structured surveys, as part of repeated surveys, as part of comprehensive checklists of co-occurring species, by scientists, by citizen scientists and so on.	"Before fieldwork, we obtained locality information from *C. canescens* herbarium specimens and online biodiversity databases such as the Southwest Environmental Information Network and the Rocky Mountain Herbarium (University of Wyoming). In addition, the Rocky Mountain Herbarium and the Colorado State University Herbarium were visited to examine potentially misidentified specimens from outlying portions of the species' distribution."[111]	112–114	48
	(A4) Spatial extent	Spatial extent of occurrences is crucial for interpretation of model predictions, including whether potential sink populations are included, whether sampling is biased or whether records outside the native range are used.	"We integrated missing countries by obtaining occurrences from the literature — that is, for France, Italy and Switzerland. To increase the accuracy of the analysis, we excluded the following records: (1) localities for which we were not able to obtain precise coordinates; [...] (4) record of *M. bourneti* in the Canary Islands, due to taxonomical issues currently unresolved (C. Ribera, personal communication, 2016)."[115]	116,117	67
	(A5) Temporal range	Environments can change over time, thus the timestamp of occurrence records is crucial for linking them to the relevant environmental conditions experienced by the species, and hence correctly describing the niche.	"Although the sightings dataset extended over 257 years, 79% of sightings occurred between 2000 and 2015. Therefore only this subset of 5,419 sightings was retained for further analysis. These sightings were divided into each quarter of the year (Jan–Mar, Apr–May, Jun–Aug and Sep–Dec) and matched with recent climate data available through online data sharing platforms."[118]	70	26

Table 1 | Details of the ecological niche modelling checklist and representation of its elements (percentage) in a review of recent (2017–2018) ecology and evolution literature (163 papers) (continued)

Category	What to report	Why reporting this element is important	Exemplar papers reporting the element	Relevant references	Papers (%)
Processing	(A6-1) Duplicate coordinates	Duplicated coordinates can potentially bias model training. Also, different modelling algorithms may have different default options for handling duplicated coordinates, either at point level or at pixel level.	"We constructed potential distributions for each species in the program Maxent 3.3.3k (Phillips et al., 2006) using the default settings, including removing duplicate species records from the same grid square."[119]	120	23
	(A6-2) Spatial and environmental outliers; error	Outliers or errors may lead to model errors. Also, the model prediction may be sensitive to outliers or errors.	"Finally, we plotted all the points on maps and excluded any point falling far outside the proven distribution described in Krapovickas et al. (2007)."[121]	116	35
	(A6-3) Spatial and coordinate uncertainty	The coordinates of a record may not represent the exact location of collection. Coordinates are often recorded or processed to different degrees of specificity (for example, two decimal points versus four). Further, coordinates are often georeferenced from locality descriptions to, for example, centroids of political boundaries. The mismatch between the coordinate uncertainty and spatial resolution of environmental variables can significantly affect the results and interpretation of the model predictions. Thus, spatial uncertainty should be reported when adequate information is available.	"For this study, precise locality coordinates for *P. solenopsis* were not available, so the district-level occurrence data published by Nagrare et al. (2009) were used (*n* = 42 records). The centroid method may be acceptable if the target scale of prediction is global but may not be appropriate at national, state or finer scales; districts are not homogeneous, and some of them can be quite large. We calculated district-level averages of climatic variables in ArcMap (version 9.3, ESRI, Redlands, CA, USA) and used those as predictors. This is a relatively unconventional use of ENM/SDM, and the results may be useful for designing detailed surveys and making district-level state, regional or national pest management policies before more detailed, precise data for this species become available."[122] "With the GeoClean function from speciesgeocodeR R Package we also removed coordinates assigned to capital cities, coordinates with latitude equal to longitude, coordinates equal to exactly zero; coordinates based on centroids of provinces, and corrected country references (cleaned GBIF records)."[123]	33,40–43	NA
	(A7-1) Sampling bias	Biased sampling, unequal sampling of a species' distribution, may cause the model to overfit environmental conditions associated with such samples. Also, different algorithms may have different default methods for handling spatially biased occurrences.	"To reduce the effects of sampling bias, we spatially filtered the occurrence dataset to ensure that no two localities were within 10 km of one another."[124]	50–52, 120, 125–129	34
	(A7-2) Spatial autocorrelation	Spatial autocorrelation, here referring to the non-independent spatial distribution of occurrences, could violate the modelling assumption of independent and identical residuals, thus could bias estimations of model parameters.	"In order to account for autocorrelation in the observations, models were also fitted in which contagion (see below: spatial interpolators) was included as an autocovariate term in the initial variable set (AGLM). These models are termed autologistic (Smith, 1994; Augustin et al., 1996; Araújo & Williams, 2000). Measures of aggregation for point and lattice data, such as Kernel estimation and nearest neighbour measures (for example, Bailey & Gatrell, 1995), can be used to model species' probabilities of occurrence. This uses the idea of positive spatial autocorrelation (Legendre, 1993), in which the occurrence of a species in one area is expected to be more likely if the species occurs in many surrounding areas (Araújo & Williams, 2000, 2001; Araújo et al., 2002). A measure of contagion (CONT) for each cell, based on a two-order neighbourhood, was used to estimate a distance-based probability of occurrence."[130]	53,127, 131–133	18

Continued

Table 1 | Details of the ecological niche modelling checklist and representation of its elements (percentage) in a review of recent (2017–2018) ecology and evolution literature (163 papers) (continued)

Category	What to report	Why reporting this element is important	Exemplar papers reporting the element	Relevant references	Papers (%)
(B) Obtaining and processing environmental data					
Metadata and processing	(B1) Source	Reporting the source of environmental data enables the reader to access them and assess their relevance to study goals.	"Climate data comprised the 19 BIOCLIM variables available from WorldClim (Hijmans et al., 2005) at a resolution of 2.5 arc-min. Elevation data were obtained from the Digital Elevation Model at PRISM (Precipitation-elevation Regressions on Independent Slopes Model; Daly et al., 1994) at 2.5 arc-min resolution."[111]	NA	99
	(B2) Download date; version of data source	Data and databases are not static — they change over time. Thus reporting the access/query/download date or version of the dataset is necessary to ensure reproducibility.	"Out of the available WorldClim data (http://www.worldclim.org), we used the 19 bioclimatic variables, which express 11 temperature and 8 precipitation metrics at about 1-km resolution (WorldClim version 1.4; Hijmans, Cameron, Parra, Jones, & Jarvis, 2005)."[134]	NA	27
	(B3) Spatial resolution	Environmental data usually have various spatial resolutions that need to be reconciled for model training. Also, the decision of spatial resolution is both a technical and an ecological issue.	"Four static variables were derived from the digital elevation model (DEM) of the EMODnet Bathymetry portal: depth (the DEM); slope and curvature, calculated using DEM Surface Tools for ArcGIS 10.2; distance to the nearest 200 m bathymetric line, calculated using QGis 2.12. Curvature was used as a proxy of sea bottom roughness, providing an estimate of sea floor relief, which can influence some cetacean species (Lindsay et al., 2016). All static variables were calculated at a spatial resolution of 0.5 × 0.5 km."[68]	64–66,135	82
	(B4) Temporal range	The temporal range (time period across which the variable was measured and averaged) is needed to determine the temporal match or mismatch with species' occurrences.	"We summarized occurrence of passerine bird species at BBS routes in the conterminous U.S. during historical (1977–1979) and recent (2012–2014) periods. Land use covariates were the proportion of the buffer surrounding each route in developed and conservation or low human use classes based on the 1974 and 2012 versions of the U.S. conterminous wall-to-wall anthropogenic land use trends dataset (NWALT; Falcone, 2015)."[136]	68	42
(C) Model calibration					
Data input	(C1) Modelling domain	The geographic domain of a model has to be specified because it is associated with the underlying assumptions of the relationship between species' distribution and the environments, as well as background selection for some ENM algorithms.	"In the second approach, locality data were overlaid on terrain base maps in ArcGIS 10.2 (Environmental Systems Research Institute, 2011) together with a world ecoregions layer (World Wildlife Fund, 2011). These were used to identify breaks in habitat and ecological regions in topographically homogeneous areas. [...] Restricting calibration areas to regions bounded by significant abiotic barriers (for example, large rivers, mountain ranges) and known or hypothesized dispersal distances yielded more accurate models and reduced these errors (Barve et al., 2011; Owens et al., 2013; Royle, Chandler, Yackulic, & Nichols, 2012; Saupe et al., 2012). Thus, in our study, Ms were constrained by deep valleys (for example, the Maranon Valley), the crests of mountains (for example, the Andes) and ~other distinct features likely to act as barriers to species distributions (for example, the llanos of northern South America)."[137]	52,71, 72,74, 138–140	50
	(C2) Number of background data	Background data are assumed to represent the environmental composition of species' accessible area, thus the optimal number of background data may depend on the extent of the study area and resolution of environmental data, as well as computation capacity.	"For each geographical background we selected 10,000 random cells that did not hold a species presence record (or all available cells if fewer than 10,000 were available)."[141]	73,142	54

Table 1 | Details of the ecological niche modelling checklist and representation of its elements (percentage) in a review of recent (2017–2018) ecology and evolution literature (163 papers) (continued)

Category	What to report	Why reporting this element is important	Exemplar papers reporting the element	Relevant references	Papers (%)
	(C3) Sampling method for background data	Random selection of background data has been used as the default strategy in some algorithms, but new methods have been developed for different purposes.	"We used Maxent with default settings, except that we applied a targeted background sampling to reduce the influence of sample selection bias (Phillips et al., 2009) by using 666 vertebrate fossil site localities (excluding moa bones [Order Dinornithiformes] and swamp sites) throughout New Zealand as background points."[143]	72,73, 75,76, 142,144, 145–147	53
	(C4) Variable selection	Selection of variables is biologically and/or statistically relevant, thus criteria and justification are needed.	"Four 'bioclimatic' layers were used to calibrate models: mean temperature of the warmest quarter, mean temperature of the coldest quarter, precipitation of the wettest quarter, and precipitation of the driest quarter. These four layers were chosen because they represent the climatic extremes that often constrain species distributions and because most other bioclimatic layers are derived from different combinations of or are tightly correlated with these variables (Root, 1988)."[137]	77,87,148	70
Algorithm	(C5) Name	Reporting the name of modelling algorithm is the basis of reproducibility.	"ENM was performed using the maximum entropy approach as implemented in MAXENT 3.3.3k (Phillips, Anderson, & Schapire, 2006)."[149]	NA	100
	(C6) Version of algorithm and software	Modelling algorithm, default settings, and dependent libraries can change over time, so providing the version will enhance the reproducibility of a study.	"For BRTs, different combinations of learning rates (0.005, 0.01, 0.05) and tree complexity (1, 2, 3) were tested. Folds were set at random and other parameters were left as default in the gbm R package (version 2.1.1). Runs on R version 3.3.2."[150]	17,151,152	59
	(C7) Parameterization	Parameter or modelling settings can influence the resulting model, and default settings may not be appropriate for a study. Thus specific settings should be reported, including default ones.	"Selecting the best settings for the regularization multipliers and number of feature classes, which determine the model complexity, requires quantitative evaluation (Merow et al., 2013). The optimal model parameters were tuned using the function ENMevaluate in the package 'ENMeval' (Muscarella et al., 2014) for R. Within ENMevaluate, we jackknifed each species presence record and evaluated models with the following feature classes: linear, quadratic, and hinge, and the following values of regularization multipliers: 0.75, 1, 1.25, 1.5."[153]	72,81,83, 154–157	45
(D) Model transfer and evaluation					
Evaluation	(D1) Evaluation index	Proper understanding of model performance requires the use of model evaluation indices. Also, different evaluation indices may be informative of different aspects of a model.	"We evaluated the performance of the models by three different methods using an independent dataset of occurrences for model evaluation: (a) an omission error test [...] (b) the binomial cumulative probability [...] (c) the partial receiver operating characteristic [...]"[158]	89,90, 159–165	90
	(D2) Threshold for evaluation index	Calculation of some evaluation indices requires a threshold. The threshold will vary by study because there is no single, default method for choosing a threshold.	"A threshold to convert continuous predicted probabilities into a binomial output was estimated for each model run, using the threshold value that maximized specificity (true negative rate) and sensitivity (true positive rate) over the evaluation dataset predictions (Liu, Newell, & White, 2016). Using this threshold, two metrics of predictive performance were derived: the sensitivity of models when predicting ARGOS tracking locations ("sensitivity. ARGOS", in% correctly classified as presences), and the true statistic skill when predicting the evaluation datasets ("TSS"; Allouche, Tsoar, & Kadmon, 2006)."[150]	92,93,166	36
	(D3) Dataset used to evaluate models	Evaluation of a model is usually based on another independent dataset, or part of the dataset not used in model training. The choice of dataset can influence the evaluation results and the subsequent interpretations.	"All these methods used observed presences as input with a 70% random sample for model development and the remaining 30% sample for model evaluation."[167]	81,91,168	39

Table 1 | Details of the ecological niche modelling checklist and representation of its elements (percentage) in a review of recent (2017–2018) ecology and evolution literature (163 papers) (continued)
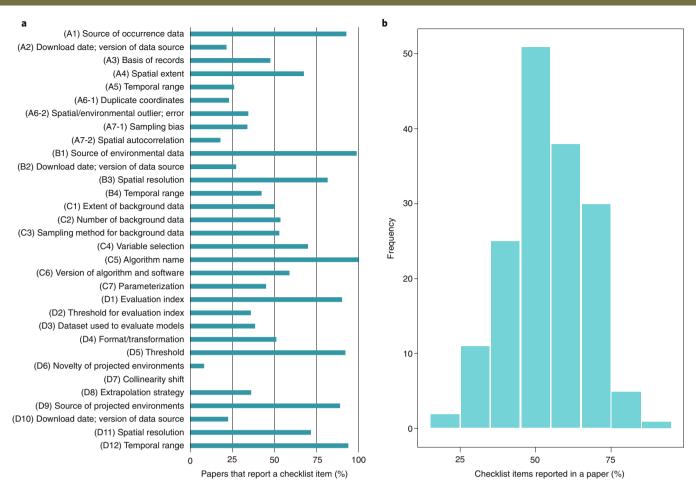
Category	What to report	Why reporting this element is important	Exemplar papers reporting the element	Relevant references	Papers (%)
Output	(D4) Format/ transformation	The raw model predictions are sometimes transformed (for example, logistic transformation) via different methods under different assumptions.	"[…] we used the logistic output format […]"[124]	80,169,170	51
	(D5) Threshold	Often, the model predictions are in continuous format, which is subsequently transformed into a binary prediction under a particular threshold. Researchers have proposed different ways of thresholding for different purposes and under varied assumptions.	"We repeated this procedure 20 times for each algorithm and used the Lowest Present Threshold values (Pearson et al., 2007) to transform each map in binary."[171]	92,93,172	92
Extrapolation	(D6) Novelty of projected environments relative to training environments	Transferring a model across space and/or time may lead to extrapolation if the projected environments are novel compared with training environments. Quantification of novel environments could help understand the uncertainties associated with model predictions.	"To assess the effect of model extrapolation on values of predictor variables lying outside the training range, that is, projecting models on non-analogous climates (cf. Nogues-Bravo, 2009), we conducted a multivariate environmental similarity surfaces (MESS) analysis, following Elith et al. (2011)."[173]	48,95,97, 174,175	8
	(D7) Collinearity shift between training and projected environments	Transferring a model outside training data may be affected by differences in collinearity structure between training and projection environments, which can lead to degraded prediction performance. Therefore, quantification of collinearity shift or any steps towards correcting for it should be specified.	"We compared the correlation matrix of the 6 variables in the training region to the average of the correlation matrices of present and future climate layers in the projected area (Tables S2 & S3 in Supplement 2). The highest absolute change of r was 0.3 for bio4 and bio17, and r increased above the 0.7 threshold for 2 pairs of variables (−0.78 for bio3 and bio4; 0.71 for bio16 and bio17; Supplement 2)."[176]	77,96,177	0
	(D8) Extrapolation strategy	Model extrapolation is statistically challenging. Different extrapolation strategies can lead to very different model predictions, therefore the choice of extrapolation, even the default setting of an algorithm, should be provided.	"Five replicates of each model were conducted with no clamping or extrapolation and with all the default 'features' used."[137]	94,97, 98,178	36
Metadata	(D9) Source	See (B1)	See (B1)	See (B1)	89
	(D10) Download date; version of data source	See (B2)	See (B2)	See (B2)	23
	(D11) Spatial resolution	See (B3)	See (B3)	See (B3)	72
	(D12) Temporal range	See (B4)	See (B4)	See (B4)	94

Model calibration (C). Typically, an ENM study first has to determine the geographic domain of interest (C1). Delimitation of the domain requires both ecological and practical justification, such as focusing on areas that have been accessible to a species[71,72], and areas that have been sampled. Many ENM algorithms make use of background points[22] that represent environmental conditions contrasting those known to be occupied by the taxa of interest. Several aspects of background point selection can influence model outcomes, including the number of points (C2)[73,74] and the algorithms used to select these points[75,76] (C3).

The suite of environmental predictors that are used in ENM should be directly relevant to a species' distributional ecology[19], and the rationale for selecting those variables should be transparent

(C4). However, as mechanistic relationships are often unknown, justification of variable selection procedures is necessary. Further, collinearity of environmental variables, a well-recognized issue in regression models, affects parameter estimation during model calibration[77]; one common strategy is to remove highly correlated environmental variable pairs following rule-of-thumb thresholds (for example, $|r| > 0.4$ or 0.7)[77,78]. Selecting one variable from a pair of variables can be subjective (for example, based on expert knowledge), objective (for example, using variable contribution to model fit[79]) or random; hence justification is required to ensure accurate interpretation and reproducibility of variable selection.

The version of the ENM software or algorithm used (C5 and C6) also needs to be provided, as these tools are often updated[80]

Fig. 1 | Completeness of checklist elements reported in the current literature. Assessments are based on 163 articles published in eight ecology and evolution journals during 2017–2018. **a**, Percentage of papers that report individual element of the checklist. **b**, Frequency of completeness (%) of checklist elements reported in all articles.

to include bug fixes or revised default settings. For instance, the default transformation method of Maxent raw output was changed from 'logistic' to 'cloglog' between versions 3.3 and 3.4[80]. Dependent libraries for coded algorithms may change over time as well.

Parameterizations or model settings and their justification (C7) are important to understanding how they may affect predictions. Examples of these settings include features and regularization values in Maxent[81,82], covariate formulas for regression-based models, link functions in generalized linear models (GLMs)[83], learning rate and maximum complexity in boosted regression trees (BRTs)[84], and optimizer values in generalized additive models [85]. In practice, authors often use the default settings provided by the software or algorithm utilized, which may or may not yield robust models[72,82,86], wheareas in other cases, authors fine-tune parameters to get best model performance[81,87].

Model transfer and evaluation (D). Understanding model performance requires model evaluation (D1). A first step is that of assessing model precision and significance — that is, whether the model can correctly predict independent presence (or absence) data and whether the model prediction is better than null expectations. Commonly used indices that measure model performance can be either threshold-independent (D2; for example, area under the receiver operating characteristic curve or ROC AUC[88]), or threshold-dependent (for example, partial ROC[89], true skill statistic or TSS, sensitivity and specificity[90]); the latter approaches require reporting of thresholds and how they were derived. In addition

to model accuracy, information criterion-based indices should be reported if they were used to select among competing models based on predictive performance and model complexity or used to generate ensembles of models. Authors should report whether and how data were partitioned to calculate the evaluation indices (D3), if genuinely independent testing data (that is, different sources and methods of collection) were not available. Common approaches include random partitioning of occurrence datasets into training and testing (for example, the default in Maxent); among other methods, partitioning based on structured blocks (for example, separating occurrences into spatial blocks) is expected to assess model transferability better[81,91]. Given the variety of options regarding data separation, it is important to specify methods used to ensure better reproducibility.

Once a model is calibrated, it may then be transferred or projected onto another landscape or time. Generally, these predictions are initially continuous (D4) and sometimes are subsequently transformed into binary predictions using a particular threshold (D5). Researchers have proposed different ways of thresholding[92,93] for different purposes and under varied assumptions, so these choices need to be reported.

Transferring a model across space and/or time may lead to extrapolation if the projected environments are novel relative to training environments. Several studies have found that environmental novelty[48,94,95] (D6) and collinearity shift (D7; changes of collinearity structure of covariates[77,96]) reduce predictive performance, and recommended quantifying the novelty of the projected environments and the collinearity shift between the calibrated and

projected environments[96,97]. Further, different algorithms use different strategies to extrapolate (clamping, truncation, extrapolation[94,98]); for example, the default clamping function in Maxent uses the marginal values in the calibration area as the prediction for more extreme conditions in transfer areas[22].

Assessing the state of reproducibility in ENM research

To assess the state of reproducibility in ENM research in the context of our proposed checklist, we reviewed current (2017–2018) ENM literature in eight widely read ecology and evolution journals: *Global Ecology and Biogeography*; *Diversity and Distributions*; *Journal of Biogeography*; *Evolution*; *Evolutionary Applications*; *Molecular Phylogenetics and Evolution*; *Molecular Biology and Evolution*; and *Systematic Biology*. Additional details of our review criteria are provided in Appendix 1, Supplementary Fig. 1, and Supplementary Tables 2 and 3.

Inclusion of elements of the checklist (32 in total) varied widely, ranging from fully reported (100%; C5 algorithm name) to not reported at all (0%; D7 collinearity shift), though documentation of the importance of this latter element is still limited in the literature[77,96]. Completeness of information across the checklist also varied among papers, ranging from 24% to 89%, averaging 54% (s.d. = 13%) of checklist elements reported in a given paper (Fig. 1).

Most studies (93%) fully reported sources of occurrence data (A1), but the date of access or version of the data source (A2) was included in only 22% of papers reviewed, and the basis of these records (A3) was described clearly in only 48% of papers. A relatively high number of papers (67%) reported the spatial extent (A4) of the occurrence data, but the temporal range (A5) was mentioned less frequently (26%). Few papers gave details of occurrence data processing, ranging between 18 and 35% in elements A6 and A7.

Although most papers we reviewed reported the source of environmental data (B1), they largely did not include download date or version of the data source: only 27% of papers reported such information for model training (B2) and only 23% for environmental data in model transfer (D10). The spatial resolution and the method of resampling layers with different spatial resolutions (B3) were generally reported (82%), although the temporal range (B4) was less frequently reported (42%). The pattern was opposite for environmental layers for model transfer: temporal range (D12) was almost always reported (94%) but spatial resolution (D11) was less frequently reported (72%).

Only 32% of papers fully reported information regarding modelling domain (C1–3). A high percentage of papers reported the variable selection procedure (C4; 70%). The ENM algorithm or software (C5) was always reported, though less frequently for the corresponding version (C6; 59%). In general, less than half of papers fully disclosed parameters for algorithms (C7; 45%).

Although model evaluation is critical for modelling studies, not all papers (90%) presented information pertaining to model evaluation (D1). Surprisingly, less than half adequately reported how the evaluation dataset was generated (D3; 39%) or mentioned specific values for threshold-dependent evaluation indices (D2; 36%). For model predictions generated, 51% of papers adequately specified output format or acknowledged that default settings were used. Among the papers that converted continuous predictions to binary, 92% specified the adopted threshold. When transferring model to different times and/or regions, few of the papers specified the extrapolation strategy (D8; 36%). The novelty of projected environments (D6) was rarely evaluated (8%).

Lessons from ecological niche modelling

Reproducibility of scientific studies has been under major scrutiny in recent years, and numerous high-profile studies have been found to be irreproducible, in large part because current reporting and publishing practices do not provide sufficient information regarding the methodologies, decisions and assumptions involved. Despite being based on a relatively recently developed toolset, ENM is no exception. For thorough evaluations of proper use of ENM applications (for example, use of ENMs in biodiversity assessment[21]), a detailed and standardized description of the methods must be provided. The checklist presented here includes the bare minimum of categories and elements necessary to evaluate and replicate ENM analyses. However, the details reported in recent publications varied greatly: on average, papers in our review included only 54% of checklist items, a generally incomplete set of information for reproducibility. This shortcoming may reflect a lack of community expectations on model reporting, or even unawareness of alternative options and underlying caveats in the modelling workflow. We highlight several key areas that were particularly deficient in reporting, and thus need attention to make ENM studies reproducible (Box 1).

Improving reproducibility with software solutions

The rapid development of ENM can be attributed at least in part to increased access to relevant data; with such development, informatics tools offer one route by which to improve reproducibility[99,100]. Such tools include data management plans[101], standardized metadata[102,103], programming language resources to record data analysis steps (for example, R and rmarkdown) and version-control tools (for example, GitHub). Open-source programming languages such as R have allowed for development of packages specifically designed for managing and processing large datasets in preparation for analysis. Exemplary packages include biogeo, which directly detects, corrects, and assesses occurrence data quality[42], and geoknife, a package designed specifically for United States Geological Survey gridded dataset management[104]. Other packages help users to create reproducible workflows, such as zoon[105], nicheA[106] kuenm[86], and Wallace[107]. In particular, the package Wallace provides a graphical user interface to build reproducible workflows, from data download to model output[107]. Borregaard and Hart[11] described how the use of these new software tools is facilitating ecological research that is both robust and transparent, and thus reproducible. The functionality of the software solutions, however, depends on developers monitoring changes in data, modelling algorithms and the software platforms (for example, R), to avoid incompatibility issues. As such, authors should report software versions for all such solutions to ensure reproducibility.

Implications for other fields

The design of the checklist presented here is based on a typical ENM workflow, involving steps of obtaining and processing data, and model calibration, transfer and evaluation. We emphasized reporting data origin and metadata; crucial steps in data processing, modelling decisions and model evaluation; and potential caveats in model transfer. Those concepts and principles are generalizable to other disciplines. Further, the specifics of the checklist that we have proposed for ENM studies could be readily generalized to be adopted by other fields, especially those that involve biological data, environmental data and statistical modelling.

Researchers have proposed similar solutions in other fields, such as climate change research[29]; however, to our knowledge, our checklist takes additional steps in refining the methodology workflow and is therefore more comprehensive. For example, information pertaining to occurrence data (data source, spatial and temporal range, and data cleaning procedures) can be generalized to other studies that rely on digitized biodiversity data and other categories of 'big data'. The information regarding environmental data necessary to reproduce studies is similar across biological research, such as in studies of relationships between species richness and environmental gradients[108]. The modelling algorithm details in the checklist are applicable to other studies that use statistical models, such

Box 1 | Improving the reproducibility of ENM studies and beyond

Reporting the use of default settings is better than not reporting anything. During our literature review, we came across many cases in which default software or algorithm settings were used. If the version of an algorithm or software is provided, others can infer and reproduce default settings. However, simple manipulations of parameter settings can lead to dramatic differences in model performance and predictions[72,81], and, since default settings can change with upgrades to software/algorithms, we recommend listing the actual default parameters employed instead of simply stating that default settings were used. This point applies to any scientific study that uses conventional software and algorithms.

Random may often be the default but should not be inferred to be so. Some modelling parameters and settings, such as data partitioning for training and testing or selecting background points (for example, in Maxent), are commonly set to 'random,' which is often the default setting. However, recent developments in ENM methodologies provide more options, such as 'block' partitioning of occurrences[81] and background selection from particular spatial extents or environmental conditions[51,75,144]. Although random may be the default setting in many modelling tools used in ENM and beyond[179–181], it needs to be reported explicitly to make the analysis reproducible.

Data and code change through time, so report the date and version. Data resources available online (for example, biodiversity records and environmental layers) are changing at an increasing pace, thus the reproducibility of studies based on these fast-evolving data, such as ENM studies, are particularly vulnerable to these changes. The status quo of biodiversity data continuously changes as more specimens or observations are collected, digitized, and mobilized online, and existing data are re-examined and updated. The accuracy of environmental data is also being improved, with advances in GIS and remote-sensing technologies. In addition to data input, existing algorithms, software and methodologies are updated and refined, and new ones are being developed and released. On a positive note, GBIF implemented a DOI system to track metadata of datasets from GBIF[182], to facilitate accurate reporting of metadata and thus increase reproducibility.

Models can be projected in both space and time. Modelling the ecological niche (via correlative methods[183]) is based on spatial locations of known occurrences of species and the corresponding environments. However, in reality species' geographic distributions are not fixed, but rather are dynamic, and environments can also change over time. Therefore, the temporal dimension is crucial in modelling niches accurately[67,68,70] and affects reproducibility. The temporal dimension of climate data was overlooked in 58% of climate research papers reviewed by Morueta-Holme et al.[29]. Our review also revealed that temporal dimensions of occurrence data and environmental data often go unreported in ENM studies (74% and 58%, respectively), although the temporal dimension tended to be better reported when the research involved model transfers (Table 1). Similar issues may plague many other studies that employ data that are spatiotemporally linked, regardless of discipline.

to these generally applicable elements, the checklist can easily be extended to incorporate information particular to a field.

Although the methods sections of most scientific publications lack the formal standardization needed for reproducibility and their length is frequently influenced by journal space limitations, the checklist approach can provide greater detail to ensure repeatability. The usual methods section, combined with a standardized checklist, will make papers easier to review and replicate. Other disciplines can and should design comparable checklists with similar concepts and levels of detail.

Closing remarks

ENM is increasingly used in ecological studies and incorporated into conservation decisions. Our literature review revealed numerous gaps that undermine reproducibility of these studies. We recommend researchers developing ENM studies in the future to consider our checklist, extend and adjust it to meet study needs, with particular focus on elements that are commonly neglected (Table 1), and include this more structured metadata in publications (see checklist template in Supplementary Table 1). This checklist provides an important tool for both understanding and replicating previous studies, and also provides editors and reviewers with an efficient way to gauge and promote ENM reproducibility[29]. As a general metadata framework linking observational data and statistical modelling, our checklist provides a starting point for adopting similar standards in other fields, both within and beyond ecology that rely on these methods.

Data availability

The checklist for ENM can be downloaded from Supplementary Table 1 and is available as an open-source project where users can comment and suggest changes (https://github.com/shandongfx/ENMchecklist or https://doi.org/10.5281/zenodo.3257732). Details of the literature review are available in the Supplementary Information.

References

1. Baker, M. 1,500 scientists lift the lid on reproducibility. *Nature* **533**, 452–454 (2016).
2. Begley, C. G. & Ellis, L. M. Drug development: raise standards for preclinical cancer research. *Nature* **483**, 531–533 (2012).
3. Mullard, A. Reliability of 'new drug target' claims called into question. *Nat. Rev. Drug Discov.* **10**, 643–644 (2011).
4. Problems with scientific research: how science goes wrong. *The Economist* https://www.economist.com/leaders/2013/10/21/how-science-goes-wrong (21 October 2013).
5. Popper, K. *Conjectures and Refutations: The Growth of Scientific Knowledge* (Routledge, 2014).
6. Wilson, G. et al. Best practices for scientific computing. *PLoS Biol.* **12**, e1001745 (2014).
7. Nichols, T. E. et al. Best practices in data analysis and sharing in neuroimaging using MRI. *Nat. Neurosci.* **20**, 299–303 (2017).
8. Tumor Analysis Best Practices Working Group. Expression profiling: best practices for data generation and interpretation in clinical trials. *Nat. Rev. Genet.* **5**, 229–237 (2004).
9. Cassey, P. & Blackburn, T. M. Reproducibility and repeatability in ecology. *Bioscience* **56**, 958–959 (2006).
10. Shapiro, J. T. & Báldi, A. Lost locations and the (ir)repeatability of ecological studies. *Front. Ecol. Environ.* **10**, 235–236 (2012).
11. Borregaard, M. K. & Hart, E. M. Towards a more reproducible ecology. *Ecography* **39**, 349–353 (2016).
12. Schnitzer, S. A. & Carson, W. P. Would ecology fail the repeatability test? *Bioscience* **66**, 98–99 (2016).
13. Milcu, A. et al. Genotypic variability enhances the reproducibility of an ecological study. *Nat. Ecol. Evol.* **2**, 279–287 (2018).
14. Nekrutenko, A. & Taylor, J. Next-generation sequencing data interpretation: enhancing reproducibility and accessibility. *Nat. Rev. Genet.* **13**, 667–672 (2012).

as linear regression models of abundance as a response to resource availability. The elements of model extrapolation (environmental novelty and collinearity shift) are also common issues for modelling practices that involve forecasting, for example, predictions of biodiversity or community changes under global change. In addition

15. Hunter, P. The reproducibility 'crisis': reaction to replication crisis should not stifle innovation. *EMBO Rep.* **18**, 1493–1496 (2017).

16. Hampton, S. E. et al. Big data and the future of ecology. *Front. Ecol. Environ.* **11**, 156–162 (2013).

17. Peterson, A. T. et al. *Ecological Niches and Geographic Distributions* (Princeton Univ. Press, 2011).

18. Franklin, J. *Mapping Species Distributions: Spatial Inference and Prediction* (Cambridge Univ. Press, 2010).

19. Guisan, A. & Zimmermann, N. E. Predictive habitat distribution models in ecology. *Ecol. Model.* **135**, 147–186 (2000).

20. Peterson, A. T. & Soberón, J. Species distribution modeling and ecological niche modeling: getting the concepts right. *Nat. Conservação* **10**, 102–107 (2012).

21. Araújo, M. B. et al. Standards for distribution models in biodiversity assessments. *Sci. Adv.* **5**, eaat4858 (2019).

22. Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **190**, 231–259 (2006).

23. Mahammad, S. S. & Ramakrishnan, R. GeoTIFF: a standard image file format for GIS applications. In *Map India Conf. 2003* 28–31 (2003).

24. Wieczorek, J. et al. Darwin Core: an evolving community-developed biodiversity data standard. *PLoS ONE* **7**, e29715 (2012).

25. Guralnick, R., Walls, R. & Jetz, W. Humboldt Core: toward a standardized capture of biological inventories for biodiversity monitoring, modeling and assessment. *Ecography* **41**, 713–725 (2018).

26. Gad-el-Hak, M. Publish or perish—an ailing enterprise? *Phys. Today* **57**, 61–62 (2004).

27. Munafò, M. R. et al. A manifesto for reproducible science. *Nat. Hum. Behav.* **1**, 0021 (2017).

28. Grimm, V. et al. A standard protocol for describing individual-based and agent-based models. *Ecol. Model.* **198**, 115–126 (2006).

29. Morueta-Holme, N. et al. Best practices for reporting climate data in ecology. *Nat. Clim. Change* **8**, 92–94 (2018).

30. Michener, W. K. et al. Participatory design of DataONE — enabling cyberinfrastructure for the biological and environmental sciences. *Ecol. Inform.* **11**, 5–15 (2012).

31. Bonney, R. et al. Citizen science: a developing tool for expanding science knowledge and scientific literacy. *Bioscience* **59**, 977–984 (2009).

32. Daru, B. H. et al. Widespread sampling biases in herbaria revealed from large-scale digitization. *New Phytol.* **217**, 939–955 (2018).

33. Park, D. S. & Davis, C. C. Implications and alternatives of assigning climate data to geographical centroids. *J. Biogeogr.* **44**, 2188–2198 (2017).

34. Meyer, C., Weigelt, P. & Kreft, H. Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecol. Lett.* **19**, 992–1006 (2016).

35. Enquist, B. J., Condit, R., Peet, R. K., Schildhauer, M. & Thiers, B. M. Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity. Preprint at https://doi.org/10.7287/peerj.preprints.2615v2 (2016).

36. Boyle, B. et al. The taxonomic name resolution service: an online tool for automated standardization of plant names. *BMC Bioinformatics* **14**, 16 (2013).

37. *iNaturalist Research-grade Observations* (iNaturalist.org, 2018); https://doi.org/10.15468/ab3s5x

38. Castro, M. C. et al. Reassessment of the hairy long-nosed armadillo '*Dasypus*' *pilosus* (Xenarthra, Dasypodidae) and revalidation of the genus *Cryptophractus* Fitzinger, 1856. *Zootaxa* **3947**, 30–48 (2015).

39. Park, D. S. & Potter, D. A reciprocal test of Darwin's naturalization hypothesis in two mediterranean-climate regions. *Glob. Ecol. Biogeogr.* **24**, 1049–1058 (2015).

40. Guralnick, R. P., Wieczorek, J., Beaman, R. & Hijmans, R. J. & the BioGeomancer Working Group BioGeomancer: automated georeferencing to map the world's biodiversity data. *PLoS Biol.* **4**, e381 (2006).

41. Wieczorek, J., Guo, Q. & Hijmans, R. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *Int. J. Geogr. Inf. Sci.* **18**, 745–767 (2004).

42. Robertson, M. P., Visser, V. & Hui, C. Biogeo: an R package for assessing and improving data quality of occurrence record datasets. *Ecography* **39**, 394–401 (2016).

43. Zizka, A. et al. CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. *Methods Ecol. Evol.* **10**, 744–751 (2019).

44. McCracken, G. F. et al. Rapid range expansion of the Brazilian free-tailed bat in the southeastern United States, 2008–2016. *J. Mammal.* **99**, 312–320 (2018).

45. Taulman, J. F. & Robbins, L. W. Range expansion and distributional limits of the nine-banded armadillo in the United States: an update of Taulman & Robbins (1996). *J. Biogeogr.* **41**, 1626–1630 (2014).

46. Feng, X. & Papeș, M. Ecological niche modelling confirms potential north-east range expansion of the nine-banded armadillo (*Dasypus novemcinctus*) in the USA. *J. Biogeogr.* **42**, 803–807 (2015).

47. Pulliam, H. R. On the relationship between niche and distribution. *Ecol. Lett.* **3**, 349–361 (2000).

48. Fitzpatrick, M. C. et al. How will climate novelty influence ecological forecasts? Using the Quaternary to assess future reliability. *Glob. Ecol. Biogeogr.* **24**, 3575–3586 (2018).

49. Belbin, L. et al. Data Quality Task Group 2: tests and assertions. *BISS* **2**, e25608 (2018).

50. Fourcade, Y., Engler, J. O., Rödder, D. & Secondi, J. Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PLoS ONE* **9**, e97122 (2014).

51. Phillips, S. J. et al. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol. Appl.* **19**, 181–197 (2009).

52. Merow, C., Allen, J. M., Aiello-Lammens, M. & Silander, J. A. Jr Improving niche and range estimates with Maxent and point process models by integrating spatially explicit information. *Glob. Ecol. Biogeogr.* **25**, 1022–1036 (2016).

53. Dormann, C. F. et al. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**, 609–628 (2007).

54. Latimer, A. M., Banerjee, S., Sang, H. Jr, Mosher, E. S. & Silander, J. A. Jr Hierarchical models facilitate spatial analysis of large data sets: a case study on invasive plant species in the northeastern United States. *Ecol. Lett.* **12**, 144–154 (2009).

55. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).

56. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).

57. *PRISM Gridded Climate Data* (PRISM Climate Group, accessed 1 July 2017); http://prism.oregonstate.edu

58. *QA Note: Case #PM_MOD16_17166* (LAADS and DAAC, 2017); https://go.nature.com/2lu5NCw

59. McGill, B. J. Matters of scale. *Science* **328**, 575–576 (2010).

60. Soberón, J. & Nakamura, M. Niches and distributional areas: concepts, methods, and assumptions. *Proc. Natl Acad. Sci. USA* **106**, 19644–19650 (2009). (suppl. 2).

61. Sunday, J. M., Bates, A. E. & Dulvy, N. K. Thermal tolerance and the global redistribution of animals. *Nat. Clim. Change* **2**, 686–690 (2012).

62. Wisz, M. S. et al. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biol. Rev.* **88**, 15–30 (2013).

63. Alexander, J. M., Diez, J. M. & Levine, J. M. Novel competitors shape species' responses to climate change. *Nature* **525**, 515–518 (2015).

64. Bradter, U., Kunin, W. E., Altringham, J. D., Thom, T. J. & Benton, T. G. Identifying appropriate spatial scales of predictors in species distribution models with the random forest algorithm. *Methods Ecol. Evol.* **4**, 167–174 (2012).

65. Song, W., Kim, E., Lee, D., Lee, M. & Jeon, S.-W. The sensitivity of species distribution modeling to scale differences. *Ecol. Model.* **248**, 113–118 (2013).

66. Connor, T. et al. Effects of grain size and niche breadth on species distribution modeling. *Ecography* **41**, 1270–1282 (2018).

67. Smeraldo, S. et al. Ignoring seasonal changes in the ecological niche of non-migratory species may lead to biases in potential distribution models: lessons from bats. *Biodiv. Conserv.* **27**, 2425–2441 (2018).

68. Fernandez, M., Yesson, C., Gannier, A., Miller, P. I. & Azevedo, J. M. N. The importance of temporal resolution for niche modelling in dynamic marine environments. *J. Biogeogr.* **44**, 2816–2827 (2017).

69. Barve, N., Martin, C., Brunsell, N. A. & Peterson, A. T. The role of physiological optima in shaping the geographic distribution of Spanish moss: physiological optima of Spanish moss. *Glob. Ecol. Biogeogr.* **23**, 633–645 (2014).

70. Williams, H. M., Willemoes, M. & Thorup, K. A temporally explicit species distribution model for a long distance avian migrant, the common cuckoo. *J. Avian Biol.* **48**, 1624–1636 (2017).

71. Barve, N. et al. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecol. Model.* **222**, 1810–1819 (2011).

72. Merow, C., Smith, M. J. & Silander, J. A. Jr. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* **36**, 1058–1069 (2013).

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

74. VanDerWal, J., Shoo, L. P., Graham, C. & Williams, S. E. Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecol. Model.* **220**, 589–594 (2009).

75. Senay, S. D., Worner, S. P. & Ikeda, T. Novel three-step pseudo-absence selection technique for improved species distribution modelling. *PLoS ONE* **8**, e71218 (2013).

76. Feng, X. & Papeş, M. Can incomplete knowledge of species' physiology facilitate ecological niche modelling? A case study with virtual species. *Divers. Distrib.* **23**, 1157–1168 (2017).

77. Dormann, C. F. et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27–46 (2013).

78. Suzuki, N., Olson, D. H. & Reilly, E. C. Developing landscape habitat models for rare amphibians with small geographic ranges: a case study of Siskiyou Mountains salamanders in the western USA. *Biodiv. Conserv.* **17**, 2197–2218 (2008).

79. Lee, D. N., Papeş, M. & Van den Bussche, R. A. Present and potential future distribution of common vampire bats in the Americas and the associated risk to cattle. *PLoS ONE* **7**, e42466 (2012).

80. Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E. & Blair, M. E. Opening the black box: an open-source release of Maxent. *Ecography* **40**, 887–893 (2017).

81. Muscarella, R. et al. ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods Ecol. Evol.* **5**, 1198–1205 (2014).

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

83. Elith, J. et al. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151 (2006).

84. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *J. Appl. Ecol.* **77**, 802–813 (2008).

85. Lehmann, A., Overton, J. M. & Leathwick, J. R. GRASP: generalized regression analysis and spatial prediction. *Ecol. Model.* **157**, 189–207 (2002).

86. Cobos, M. E., Peterson, A. T., Osorio-Olvera, L. & Narayani, B. kuenm: An R package for detailed development of Maxent ecological niche models. *PeerJ* **7**, e6281 (2019).

87. Moreno-Amat, E. et al. Impact of model complexity on cross-temporal transferability in Maxent species distribution models: an assessment using paleobotanical data. *Ecol. Model.* **312**, 308–317 (2015).

88. DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* **44**, 837–845 (1988).

89. Peterson, A. T., Papeş, M. & Soberón, J. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecol. Model.* **213**, 63–72 (2008).

90. Allouche, O., Tsoar, A. & Kadmon, R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS): assessing the accuracy of distribution models. *J. Appl. Ecol.* **43**, 1223–1232 (2006).

91. Roberts, D. R. et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **40**, 913–929 (2017).

92. Liu, C., Berry, P. M., Dawson, T. P. & Pearson, R. G. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28**, 385–393 (2005).

93. Liu, C., White, M. & Newell, G. Selecting thresholds for the prediction of species occurrence with presence-only data. *J. Biogeogr.* **40**, 778–789 (2013).

94. Owens, H. L. et al. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecol. Model.* **263**, 10–18 (2013).

95. Elith, J., Kearney, M. & Phillips, S. The art of modelling range-shifting species. *Methods Ecol. Evol.* **1**, 330–342 (2010).

96. Feng, X., Park, D. S., Pandey, R., Liang, Y. & Papeş, M. Collinearity in ecological niche modeling: confusions and challenges. *Ecol. Evol.* https://doi.org/10.1002/ece3.5555 (2019).

97. Qiao, H. et al. An evaluation of transferability of ecological niche models. *Ecography* **42**, 521–534 (2019).

98. Elith, J. & Graham, C. H. Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. *Ecography* **32**, 66–77 (2009).

99. Soberón, J. & Peterson, A. T. Biodiversity informatics: managing and applying primary biodiversity data. *Phil. Trans. R. Soc. Lond. B* **359**, 689–698 (2004).

100. Boyd, D. S. & Foody, G. M. An overview of recent remote sensing and GIS based research in ecological informatics. *Ecol. Inform.* **6**, 25–36 (2011).

101. Michener, W. K. in *Ecological Informatics* (eds Recknagel, F. & Michener, W.) 13–26 (Springer, 2018).

102. Borba, C. & Correa, P. L. P. in *Metadata and Semantics Research. MTSR 2015* (eds Garoufallou, E., Hartley, R. & Gaitanou, P.) 113–118 (Communications in Computer and Information Science Vol. 544, Springer, 2015).

103. Merow, C. et al. Species' range model metadata standards: RMMS. *Glob. Ecol. Biogeogr.* https://doi.org/10.1111/geb.12993 (2019).

104. Read, J. S. et al. geoknife: reproducible web-processing of large gridded datasets. *Ecography* **39**, 354–360 (2016).

105. Golding, N. et al. The zoon R package for reproducible and shareable species distribution modelling. *Methods Ecol. Evol.* **9**, 260–268 (2018).

106. Qiao, H. et al. NicheA: creating virtual species and ecological niches in multivariate environmental scenarios. *Ecography* **39**, 805–813 (2016).

107. Kass, J. M. et al. Wallace: A flexible platform for reproducible modeling of species niches and distributions built for community expansion. *Methods Ecol. Evol.* **9**, 1151–1156 (2018).

108. Sandel, B. et al. The influence of Late Quaternary climate-change velocity on species endemism. *Science* **334**, 660–664 (2011).

109. Bosch, S., Tyberghein, L., Deneudt, K., Hernandez, F. & De Clerck, O. In search of relevant predictors for marine species distribution modelling using the MarineSPEED benchmark dataset. *Divers. Distrib.* **24**, 144–157 (2018).

110. Franklin, J., Serra-Diaz, J. M., Syphard, A. D. & Regan, H. M. Big data for forecasting the impacts of global change on plant communities. *Glob. Ecol. Biogeogr.* **26**, 6–17 (2017).

111. McMinn, R. L., Russell, F. L. & Beck, J. B. Demographic structure and genetic variability throughout the distribution of Platte thistle (*Cirsium canescens* Asteraceae). *J. Biogeogr.* **44**, 375–385 (2017).

112. Graham, C. H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A. T. New developments in museum-based informatics and applications in biodiversity analysis. *Trends Ecol. Evol.* **19**, 497–503 (2004).

113. Jetz, W., McPherson, J. M. & Guralnick, R. P. Integrating biodiversity distribution knowledge: toward a global map of life. *Trends Ecol. Evol.* **27**, 151–159 (2012).

114. Edwards, T. C. Jr, Cutler, D. R., Zimmermann, N. E., Geiser, L. & Moisen, G. G. Effects of sample survey design on the accuracy of classification tree models in species distribution models. *Ecol. Model.* **199**, 132–141 (2006).

115. Mammola, S. & Isaia, M. Rapid poleward distributional shifts in the European cave-dwelling *Meta* spiders under the influence of competition dynamics. *J. Biogeogr.* **44**, 2789–2797 (2017).

116. Soley-Guardia, M., Radosavljevic, A., Rivera, J. L. & Anderson, R. P. The effect of spatially marginal localities in modelling species niches and distributions. *J. Biogeogr.* **41**, 1390–1401 (2014).

117. McPherson, J. M., Walter, J. & Rogers, D. J. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *J. Appl. Ecol.* **41**, 811–823 (2004).

118. Phillips, N. D. et al. Applying species distribution modelling to a data poor, pelagic fish complex: the ocean sunfishes. *J. Biogeogr.* **44**, 2176–2187 (2017).

119. Lee, T. R. C. et al. Ecological diversification of the Australian *Coptotermes* termites and the evolution of mound building. *J. Biogeogr.* **44**, 1405–1417 (2017).

120. Boria, R. A., Olson, L. E., Goodman, S. M. & Anderson, R. P. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecol. Model.* **275**, 73–77 (2014).

121. Ceolin, G. B. & Giehl, E. L. H. A little bit everyday: range size determinants in *Arachis* (Fabaceae), a dispersal-limited group. *J. Biogeogr.* **44**, 2798–2807 (2017).

122. Kumar, S., Graham, J., West, A. M. & Evangelista, P. H. Using district-level occurrences in MaxEnt for predicting the invasion potential of an exotic insect pest in India. *Comput. Electron. Agric.* **103**, 55–62 (2014).

123. Gomes, V. H. F. et al. Species distribution modelling: contrasting presence-only models with plot abundance data. *Sci. Rep.* **8**, 1003 (2018).

124. Boria, R. A., Olson, L. E., Goodman, S. M. & Anderson, R. P. A single-algorithm ensemble approach to estimating suitability and uncertainty: cross-time projections for four Malagasy tenrecs. *Divers. Distrib.* **23**, 196–208 (2017).

125. Varela, S., Anderson, R. P., García-Valdés, R. & Fernández-González, F. Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography* **33**, 1084–1091 (2014).

126. Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B. & Anderson, R. P. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **38**, 541–545 (2015).

127. Hijmans, R. J. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* **93**, 679–688 (2012).

128. Hortal, J., Valverde, A. J., Gómez, J. F. & Lobo, J. M. Historical bias in biodiversity inventories affects the observed environmental niche of the species. *Oikos* **117**, 847–858 (2008).

129. Royle, J. A., Nichols, J. D. & Kéry, M. Modelling occurrence and abundance of species when detection is imperfect. *Oikos* **110**, 353–359 (2005).

130. Segurado, P. & Araújo, M. B. An evaluation of methods for modelling species distributions. *J. Biogeogr.* **31**, 1555–1568 (2004).

131. Latimer, A. M., Wu, S., Gelfand, A. E. & Silander, J. A. Building statistical models to analyze species distributions. *Ecol. Appl.* **16**, 33–50 (2006).

132. Record, S., Fitzpatrick, M. C., Finley, A. O., Veloz, S. & Ellison, A. M. Should species distribution models account for spatial autocorrelation? A test of model projections across eight millennia of climate change: Projecting spatial species distribution models. *Glob. Ecol. Biogeogr.* **22**, 760–771 (2013).

133. Wintle, B. A. & Bardos, D. C. Modeling species-habitat relationships with spatially autocorrelated observation data. *Ecol. Appl.* **16**, 1945–1958 (2006).

134. Figueiredo, F. O. G. et al. Beyond climate control on species range: the importance of soil data to predict distribution of Amazonian plant species. *J. Biogeogr.* **45**, 190–200 (2018).

135. Guisan, A., Graham, C. H., Elith, J. & Huettmann, F. Sensitivity of predictive species distribution models to change in grain size. *Divers. Distrib.* **13**, 332–340 (2007).

136. Sofaer, H. R., Jarnevich, C. S. & Flather, C. H. Misleading prioritizations from modelling range shifts under climate change. *Glob. Ecol. Biogeogr.* **27**, 658–666 (2018).

137. Cooper, J. C. & Soberón, J. Creating individual accessible area hypotheses improves stacked species distribution model performance. *Glob. Ecol. Biogeogr.* **27**, 156–165 (2018).

138. Acevedo, P., Jiménez-Valverde, A., Lobo, J. M. & Real, R. Delimiting the geographical background in species distribution modelling. *J. Biogeogr.* **39**, 1383–1390 (2012).

139. Qiao, H., Escobar, L. E. & Peterson, A. T. Accessible areas in ecological niche comparisons of invasive species: recognized but still overlooked. *Sci. Rep.* **7**, 1213 (2017).

140. Saupe, E. E. et al. Variation in niche and distribution model performance: the need for *a priori* assessment of key causal factors. *Ecol. Model.* **237–238**, 11–22 (2012).

141. Hill, M. P., Gallardo, B. & Terblanche, J. S. A global assessment of climatic niche shifts and human influence in insect invasions. *Glob. Ecol. Biogeogr.* **26**, 679–689 (2017).

142. Renner, I. W. & Warton, D. I. Equivalence of Maxent and Poisson point process models for species distribution modeling in ecology. *Biometrics* **69**, 274–281 (2013).

143. Scofield, R. P. et al. The origin and phylogenetic relationships of the New Zealand ravens. *Mol. Phylogen. Evol.* **106**, 136–143 (2017).

144. Iturbide, M. et al. A framework for species distribution modelling with improved pseudo-absence generation. *Ecol. Model.* **312**, 166–174 (2015).

145. Hertzog, L. R., Besnard, A. & Jay-Robert, P. Field validation shows bias-corrected pseudo-absence selection is the best method for predictive species-distribution modelling. *Divers. Distrib.* **20**, 1403–1413 (2014).

146. Warton, D. I. & Shepherd, L. C. Poisson point process models solve the 'pseudo-absence problem' for presence-only data in ecology. *Ann. Appl. Stat.* **4**, 1383–1402 (2010).

147. Beyer, H. L. et al. The interpretation of habitat preference metrics under use-availability designs. *Phil. Trans. R. Soc. Lond. B* **365**, 2245–2254 (2010).

148. Petitpierre, B., Broennimann, O., Kueffer, C., Daehler, C. & Guisan, A. Selecting predictors to maximize the transferability of species distribution models: lessons from cross-continental plant invasions. *Glob. Ecol. Biogeogr.* **26**, 275–287 (2017).

149. Sochor, M., Šarhanová, P., Pfanzelt, S. & Trávníček, B. Is evolution of apomicts driven by the phylogeography of the sexual ancestor? Insights from European and caucasian brambles (*Rubus*, Rosaceae). *J. Biogeogr.* **44**, 2717–2728 (2017).

150. Derville, S., Torres, L. G., Iovan, C. & Garrigue, C. Finding the right fit: Comparative cetacean distribution models using multiple data sources and statistical approaches. *Divers. Distrib.* **24**, 1657–1673 (2018).

151. Guisan, A., Thuiller, W. & Zimmermann, N. E. *Habitat Suitability and Distribution Models: With Applications in R* (Cambridge Univ. Press, 2017).

152. Qiao, H., Soberón, J. & Peterson, A. T. No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. *Methods Ecol. Evol.* **6**, 1126–1136 (2015).

153. Herrera, J. P. et al. Estimating the population size of lemurs based on their mutualistic food trees. *J. Biogeogr.* **45**, 2546–2563 (2018).

154. Warren, D. L. & Seifert, S. N. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecol. Appl.* **21**, 335–342 (2011).

155. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **77**, 802–813 (2008).

156. Anderson, R. P. & Gonzalez, I. Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with Maxent. *Ecol. Model.* **222**, 2796–2811 (2011).

157. Guillera-Arroita, G. et al. Is my species distribution model fit for purpose? Matching data and models to applications. *Glob. Ecol. Biogeogr.* **24**, 276–292 (2015).

158. Martínez-Gutiérrez, P. G., Martínez-Meyer, E., Palomares, F. & Fernández, N. Niche centrality and human influence predict rangewide variation in population abundance of a widespread mammal: the collared peccary (*Pecari tajacu*). *Divers. Distrib.* **24**, 103–115 (2018).

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

160. Lobo, J. M., Jiménez-Valverde, A. & Real, R. AUC: a misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* **17**, 145–151 (2008).

161. Liu, C., White, M. & Newell, G. Measuring and comparing the accuracy of species distribution models with presence-absence data. *Ecography* **34**, 232–243 (2011).

162. Bahn, V. & Mcgill, B. J. Testing the predictive performance of distribution models. *Oikos* **122**, 321–331 (2012).

163. Boyce, M. S., Vernier, P. R., Nielsen, S. E. & Schmiegelow, F. Evaluating resource selection functions. *Ecol. Model.* **157**, 281–300 (2002).

164. Lawson, C. R., Hodgson, J. A., Wilson, R. J. & Richards, S. A. Prevalence, thresholds and the performance of presence-absence models. *Methods Ecol. Evol.* **5**, 54–64 (2013).

165. Fielding, A. H. & Bell, J. F. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* **24**, 38–49 (1997).

166. Liu, C., Newell, G. & White, M. On the selection of thresholds for predicting species occurrence with presence-only data. *Ecol. Evol.* **6**, 337–348 (2015).

167. Johnston, M. R., Elmore, A. J., Mokany, K., Lisk, M. & Fitzpatrick, M. C. Field-measured variables outperform derived alternatives in Maryland stream biodiversity models. *Divers. Distrib.* **23**, 1054–1066 (2017).

168. Hastie, T., Tibshirani, R. & Friedman, J. H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer, 2009).

169. Elith, J. et al. A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* **17**, 43–57 (2011).

170. Royle, J. A., Chandler, R. B., Yackulic, C. & Nichols, J. D. Likelihood analysis of species occurrence probability from presence only data for modelling species distributions. *Methods Ecol. Evol.* **3**, 545–554 (2012).

171. Bartoleti, L. F. M. et al. Phylogeography of the dry vegetation endemic species *Nephila sexpunctata* (Araneae: Araneidae) suggests recent expansion of the Neotropical Dry Diagonal. *J. Biogeogr.* **44**, 2007–2020 (2017).

172. Pearson, R. G., Raxworthy, C. J., Nakamura, M. & Peterson, A. T. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J. Biogeogr.* **34**, 102–117 (2006).

173. Di Febbraro, M. et al. Does the jack of all trades fare best? Survival and niche width in Late Pleistocene megafauna. *J. Biogeogr.* **44**, 2828–2838 (2017).

174. Peterson, A. T., Papeş, M. & Eaton, M. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography* **30**, 550–560 (2007).

175. Randin, C. F. et al. Are niche-based species distribution models transferable in space? *J. Biogeogr.* **33**, 1689–1703 (2006).

176. Feng, X., Lin, C., Qiao, H. & Ji, L. Assessment of climatically suitable area for *Syrmaticus reevesii* under climate change. *Endanger. Species Res.* **28**, 19–31 (2015).

177. Braunisch, V. et al. Selecting from correlated climate variables: a major source of uncertainty for predicting species distributions under climate change. *Ecography* **36**, 971–983 (2013).

178. Zurell, D., Elith, J. & Schröder, B. Predicting to new environments: tools for visualizing model behaviour and impacts on mapped distributions. *Divers. Distrib.* **18**, 628–634 (2012).

179. Matsumoto, M. & Nishimura, T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Sim.* **8**, 3–30 (1998).

180. Bak, P. & Sneppen, K. Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.* **71**, 4083–4086 (1993).

181. Newman, M. E. J., Watts, D. J. & Strogatz, S. H. Random graph models of social networks. *Proc. Natl Acad. Sci. USA* **99**, 2566–2572 (2002). (suppl. 1).

182. Citation guidelines. *GBIF* https://www.gbif.org/citation-guidelines (2018).

183. Peterson, A. T., Papeş, M. & Soberón, J. Mechanistic and correlative models of ecological niches. *Eur. J. Ecol.* **1**, 28–38 (2015).

Acknowledgements

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-019-0972-5.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence should be addressed to X.F.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.