ECOGRAPHY

Review and synthesis

A standard protocol for reporting species distribution models

Damaris Zurell, Janet Franklin, Christian König, Phil J. Bouchet, Carsten F. Dormann, Jane Elith, Guillermo Fandos, Xiao Feng, Gurutzeta Guillera-Arroita, Antoine Guisan, José J. Lahoz-Monfort, Pedro J. Leitão, Daniel S. Park, A. Townsend Peterson, Giovanni Rapacciuolo, Dirk R. Schmatz, Boris Schröder, Josep M. Serra-Diaz, Wilfried Thuiller, Katherine L. Yates, Niklaus E. Zimmermann and Cory Merow

D. Zurell (https://orcid.org/0000-0002-4628-3558) ✉ (damaris@zurell.de), C. König (https://orcid.org/0000-0003-0585-5983) and G. Fandos (https://orcid.org/0000-0003-1579-9444), Geography Dept, Humboldt-Univ. Berlin, Berlin, Germany, and Inst. for Biochemistry and Biology, Univ. Potsdam, Potsdam, Germany. DZ also at: Land Change Science, Swiss Federal Research Inst. WSL, Birmensdorf, Switzerland. – J. Franklin (https://orcid.org/0000-0003-0314-4598), Dept Botany and Plant Sciences, Univ. California, Riverside, USA. – P. J. Bouchet (https://orcid.org/0000-0002-2144-2049), Centre for Research into Ecological and Environmental Modelling, School of Mathematics and Statistics, Univ. St Andrews, St Andrews, Scotland. – C. F. Dormann (https://orcid.org/0000-0002-9835-1794), Biometry and Environmental Analysis, Univ. Freiburg, Freiburg, Germany. – J. Elith (https://orcid.org/0000-0002-8706-0326), G. Guillera-Arroita (https://orcid.org/0000-0002-8387-5739) and J. J. Lahoz-Monfort (https://orcid.org/0000-0002-0845-7035), School of BioSciences, Univ. Melbourne, Melbourne, Australia. – X. Feng (http://orcid.org/0000-0003-4638-3927), Inst. Environment, Univ. Arizona, Tucson, USA. – A. Guisan (https://orcid.org/0000-0002-3998-4815), Univ. Lausanne, DEE & IDYST, Lausanne, Switzerland. – P. J. Leitão (https://orcid.org/0000-0003-3038-9531), Landscape Ecology and Environmental Systems Analysis, Technische Univ. Braunschweig, Braunschweig, Germany, and Geography Dept, Humboldt-Univ. Berlin, Berlin, Germany. – D. S. Park (https://orcid.org/0000-0003-2783-530X), Dept Organismic and Evolutionary Biology and Harvard Univ. Herbaria, Harvard Univ., Cambridge, USA. – A. T. Peterson (https://orcid.org/0000-0003-0243-2379), Biodiversity Inst., Univ. Kansas, Lawrence, USA. – G. Rapacciuolo (http://orcid.org/0000-0003-1494-9017), Inst. Biodiversity Science and Sustainability, California Academy of Sciences, San Francisco, USA. – D. R. Schmatz (https://orcid.org/0000-0002-8198-6136) and N. E. Zimmermann (https://orcid.org/0000-0003-3099-9604), Land Change Science, Swiss Federal Research Inst. WSL, Birmensdorf, Switzerland. NEZ also at: Swiss Federal Inst. of Technology ETH, Dept Environmental Systems Science, Zürich, Switzerland. – B. Schröder (https://orcid.org/0000-0002-8577-7980), Landscape Ecology and Environmental Systems Analysis, Technische Univ. Braunschweig, Braunschweig, Germany, and Berlin-Brandenburg Inst. of Advanced Biodiversity Research (BBIB), Berlin, Germany. – J. M. Serra-Diaz (https://orcid.org/0000-0003-1988-1154), Univ. Lorraine, AgroParisTech, INRAE, Nancy, France. – W. Thuiller (https://orcid.org/0000-0002-5388-5274), Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LECA, Grenoble, France. – K. L. Yates (http://orcid.org/0000-0001-8429-2941), School of Science, Engineering and Environment, Univ. Salford, UK. – C. Merow, Ecology and Evolutionary Biology, Univ. Connecticut, Storrs, USA.

Ecography
43: 1–17, 2020
doi: 10.1111/ecog.04960

Subject Editor: David Nogués-Bravo
Editor-in-Chief: Miguel Araújo
Accepted 30 April 2020

www.ecography.org

Species distribution models (SDMs) constitute the most common class of models across ecology, evolution and conservation. The advent of ready-to-use software packages and increasing availability of digital geoinformation have considerably assisted the application of SDMs in the past decade, greatly enabling their broader use for informing conservation and management, and for quantifying impacts from global change. However, models must be fit for purpose, with all important aspects of their development and applications properly considered. Despite the widespread use of SDMs, standardisation and documentation of modelling protocols remain limited, which makes it hard to assess whether development steps are appropriate for end use. To address these issues, we propose a standard protocol for reporting SDMs, with an emphasis on describing how a study's objective is achieved through a series of modeling decisions. We call this the ODMAP (Overview, Data, Model, Assessment and Prediction) protocol, as its components reflect the main steps involved in building SDMs and other empirically-based biodiversity models. The ODMAP protocol serves two main purposes. First, it provides a checklist for authors, detailing key steps for

model building and analyses, and thus represents a quick guide and generic workflow for modern SDMs. Second, it introduces a structured format for documenting and communicating the models, ensuring transparency and reproducibility, facilitating peer review and expert evaluation of model quality, as well as meta-analyses. We detail all elements of ODMAP, and explain how it can be used for different model objectives and applications, and how it complements efforts to store associated metadata and define modelling standards. We illustrate its utility by revisiting nine previously published case studies, and provide an interactive web-based application to facilitate its use. We plan to advance ODMAP by encouraging its further refinement and adoption by the scientific community.

Keywords: biodiversity assessment, ecological niche model, habitat suitability model, reproducibility, Shiny, transparency

Introduction

Modelling species' environmental requirements and mapping their distributions through space and time constitute important aspects of many biological analyses, particularly in support of conservation and management interventions (Franklin 2010). Species distribution models (SDMs) represent a set of popular techniques for interpolating and extrapolating species distributions based on quantitative or rule-based models, with several review papers (Franklin 1995, Guisan and Zimmermann 2000, Guisan and Thuiller 2005, Elith and Leathwick 2009) and textbooks describing their application in detail (Franklin 2010, Peterson et al. 2011, Guisan et al. 2017). The number of studies employing SDMs has increased tremendously over recent decades (Sequeira et al. 2018, Araújo et al. 2019), with > 1000 publications related to SDMs being released every year (Peterson and Soberón 2012), including many receiving > 1000 citations each (Barbosa and Schneck 2015). Today, SDMs present the most widely used modelling tool for forecasting global change impacts on biodiversity (Guisan et al. 2013, Ehrlén and Morris 2015, Ferrier et al. 2016). This boom in SDM studies is likely related to the increasing availability of digital data (Jetz et al. 2012, Franklin et al. 2017, Wüest et al. 2020) and easy-to-use software packages (Phillips et al. 2006, Thuiller et al. 2009, Brown 2014, Naimi and Araújo 2016, Golding et al. 2018, Kass et al. 2018) accompanied by detailed guides, manuals and textbooks (Elith et al. 2008, Merow et al. 2013, Guisan et al. 2017). Despite their widespread use, SDM methods and results are often limited in their reproducibility because of a lack of reporting standards (Rodríguez-Castañeda et al. 2012, Araújo et al. 2019, Feng et al. 2019, Hao et al. 2019). In the the lexicon of research reproducibility (Goodman et al. 2016), methods reproducibility means that sufficient details are provided on data and methods in order to independently repeat the study, while results reproducibility means that the same results can be obtained from an independent study (Plesser 2018). Here, we propose a standard protocol for reporting SDMs to improve their methods reproducibility, ensuring transparency and consistency in their development and application.

We here use the term SDM to refer to any empirically-based biodiversity model obtained from statistical and machine learning methods that associate geographic biodiversity records (i.e. typically in the form of expert-derived or observed presences, and sometimes absences/non-detections, or measured counts) with the abiotic and/or biotic characteristics at those locations (following Elith and Leathwick 2009). Common terms used synonymously for SDMs or closely related models include ecological niche models (ENM), species range models, environmental or climate envelopes, habitat suitability and habitat distribution models, occupancy models, resource selection functions, abundance and N-mixture models. Often, these names emphasise different aspects of the entities being modelled: the niche, the distribution or the habitat preferences of species, or the data type used (Elith and Leathwick 2009, Peterson and Soberón 2012, Guisan et al. 2017).

Generally, information on both the data and methods used should be provided in sufficient detail to allow anyone to reproduce the findings of a given study – provided data are also available – and to maximise transparency and allow robust quality control (Feng et al. 2019, Merow et al. 2019). Transparency and reproducibility are especially important for models intended as quantitative tools for ecological impact assessments, conservation planning and decision making, and biodiversity analyses (Golding et al. 2018, Araújo et al. 2019, Rapacciuolo 2019). Key to this is communicating sufficient detail about the input data, the model implementation, its evaluation and validation, and output processing such that end-users (e.g. conservationist, evaluator) has enough information at hand to judge the model's reliability and relevance without personal communication with the authors (Araújo et al. 2019, García-Díaz et al. 2019, Rapacciuolo 2019).

Methods reproducibility is crucial for ensuring adherence to minimum standards and supporting the delivery of adequate outputs for policy decisions. Indeed, poor or inconsistent modelling practices can lead to inappropriate inference and misguided conservation actions (García-Díaz et al. 2019). Recognizing the necessity for reproducibility and transparency, the recent IPBES (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services) methodological assessment report acknowledged the need for agreed-upon standards in biodiversity assessments (Ferrier et al. 2016). Similarly, the IUCN (International Union for Conservation of Nature) also defined preliminary standards that should be adhered to for assessing the threat status of species based on SDMs (IUCN Standards and Petitions Subcommittee 2017); if these standards are not

adequately met by a scientific study, then the results cannot be used as input for conservation assessments or decision making. More recently, Araújo et al. (2019) proposed best-practice standards for biodiversity assessments using SDMs, and suggested scoring SDM studies into gold (aspirational), silver (current best practice), bronze (acceptable practice) and deficient categories based on the combined quality of the input data and the modelling, evaluation and predictions approaches employed. When scoring a random subset of 400 SDM studies, Araújo et al. (2019) found that 46% of the studies were deficient in at least one aspect. In particular, many studies did not test the effects of uncertainty in predictor variables, structural and parameter uncertainty in the models, or robustness of model assumptions.

Best practice standards in modelling cannot be achieved unless standard procedures for reporting exist. A standard protocol for reporting individual-based and agent-based models (IBM/ABM) was introduced more than a decade ago: the ODD protocol (Overview, Design concepts, Details; Grimm et al. 2006). A review of the first five years of the ODD protocol showed that it not only improved the transparency of IBM/ABM studies but also facilitated a more rigorous formulation of models by providing a checklist of critical modelling steps to consider (Grimm et al. 2010). Similarly, shared data standards like the Darwin Core Standard (DwC; Wieczorek et al. 2012) and metadata standards like the Ecological Metadata Language (EML; based on Michener et al. 1997) have proved essential to compiling primary biodiversity data records in repositories such as GBIF (<www.gbif.org/>; Anderson et al. 2016). Recently, Merow et al. (2019) defined a range modelling metadata framework to report the modelling steps and results from SDMs, and Feng et al. (2019) suggested a checklist with essential elements needed to ensure SDM reproducibility. Both author groups emphasised that the proposed frameworks provide only starting points that require further development through community efforts. With this in mind, we engaged in such a community effort to refine these initial metadata standards and merge them within a standard protocol for reporting SDM methods from scientific studies.

Standardised approaches not only benefit beginners in the field, but also authors, expert referees and journal editors. Specifically, for authors, standard protocols encourage defining and reporting the modelling steps in a structured way. For reviewers and editors, they provide an efficient way of judging whether appropriate modelling decisions were made with respect to the study objectives and whether the modelling study is reproducible. For evaluators and policy makers, standard protocols will help form expectations of which information will be found where (Rapacciuolo 2019), thus simplifying meta-analyses and facilitating scoring the various model elements according to best-practice standards (Araújo et al. 2019).

Here, we propose an adaptation of the ODD protocol to SDM studies. Our aim is not to define best practice in data and methods (Araújo et al. 2019), but rather to support best practice in reporting data and modelling choices. In particular, the standard protocol provides a quick guide to the main steps of fitting SDMs and a checklist of all the information necessary to evaluate the validity and reproducibility of an SDM study for a particular application. This complements and integrates recent work defining range model metadata standards (RMMS; Merow et al. 2019). Importantly, we provide a web-based application to fill in the protocol and which relies on and extends the metadata dictionary defined by Merow et al. (2019). Methodologies and data types evolve over time, and will require redefining best practices with respect to intended objectives. By harnessing the RMMS dictionary (Merow et al. 2019), ODMAP provides a guide for developing and a language for documenting SDMs based on the RMMS dictionary. Although we acknowledge that the protocol will require some time investment and may seem cumbersome at the start, we believe that, in the long run, it should ease the burden on authors and reviewers by providing a generic workflow and clear reporting guidelines that are understandable and easy to follow. Overall, the protocol should not increase the length of publications because much of the description can be provided as Supplementary material.

A standard protocol for species distribution models

We propose a standard protocol that follows the five basic modelling steps of SDMs (described in e.g. Guisan and Zimmermann 2000, Elith and Leathwick 2009, Franklin 2010, Peterson et al. 2011, Guisan et al. 2017, Araújo et al. 2019): Overview/Conceptualisation, Data, Model fitting, Assessment and Prediction (ODMAP; Fig. 1). We set it up in an easy to follow checklist format (Table 1). In principle, this protocol should work for any empirically-based biodiversity model beyond single species distribution models, including e.g. community-level models (D'Amen et al. 2017, Norberg et al. 2019, Zurell et al. 2020) and models of functional composition (Wüest et al. 2018). Often, SDMs constitute only one part of the methods of a study and are supplemented by further analyses. Here, we argue that any scientific application of SDMs should include the entire ODMAP protocol (Table 1), but in most publications it will be sufficient to include the Overview section of ODMAP (Fig. 1) as prose in the methods of the main text, while moving the entire ODMAP checklist to the Supplementary material (also see example case studies in Supplementary material Appendix 1–9). In the following, we first give a brief overview of the different ODMAP sections before providing further details on each of these (Fig. 1, Table 1).

Any SDM or biodiversity analysis starts with the conceptualisation of one or several underlying questions and related hypotheses. These conceptual considerations should be summarised in the Overview section, which captures the skeleton of the analyses, providing enough information for readers to understand the model setup and workflow

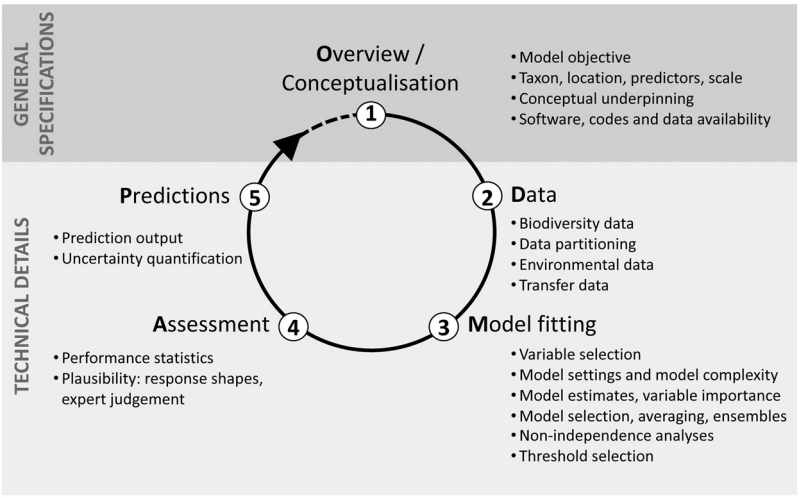

Figure 1. The five main modelling steps in the species distribution modelling cycle also constitute the five main sections of the ODMAP (Overview, Data, Model, Assessment, Prediction) protocol. Each section contains unique information that is detailed in Table 1.

(Guisan and Zimmermann 2000, Austin 2002, 2007). In particular, Overview specifies the model objectives, the focal organism(s), the type of biodiversity data, the type of environmental predictor variables, the spatiotemporal scale of the analyses, the underlying hypotheses about the biodiversity–environment relationship, the critical model assumptions, the chosen SDM algorithms and desired model complexity, and, lastly, the software used. Overview thus provides a brief but informative summary of the basic modelling decisions and the modelling pipeline (Table 1). Including Overview in the methods section of a publication will thus ensure that all key aspects of the SDM are specified in the main text of the scientific article or report while details could be relegated as Supplementary material (Fig. 1).

The Data, Model, Assessment and Prediction sections in ODMAP summarise the technical details needed to reproduce the methods (Feng et al. 2019, Merow et al. 2019) and to assess their appropriateness for different purposes (e.g. biodiversity assessments, Araújo et al. 2019). The Data section details the data and their preparation, including potential sampling bias and/or imperfect detection, any data cleaning and processing steps, as well as any (re-)scaling or transformation of data (spatial, temporal, taxonomic scaling). Model fitting is the central step where species–environment relationships are estimated using the selected algorithms. In the protocol, details should be provided about model settings, model tuning/selection, and whether and how potential sampling bias and/or imperfect detection have been dealt with. The section on Assessment (of models) describes both how the estimated species-environmental relationship was assessed for plausibility and how the model's predictive ability was quantified using appropriate goodness-of-fit measures and performance statistics. The Predictions section outlines the methods used to generate the spatial and/or temporal outputs of the model (e.g. transfers/projections in space and time) as well as any procedures for addressing uncertainty in those predictions. If pure explanation is the goal of the SDM study and no predictions are being made, then the protocol can be reduced to the first four sections.

ODMAP sections and elements

Each ODMAP section is divided into several subsections that consist of different elements. A checklist of these is provided in Table 1 (and more detail provided in Supplementary material Table A1). We distinguish sections, subsections and elements that are mandatory and should always be reported, from those that are only needed for specific model purposes or are optional (Table 1). Filling in all mandatory, and potentially the optional, fields of ODMAP will ensure methods reproducibility and transparency for peers and evaluators.

Overview

We identified eleven obligatory subsections that should be specified in the Overview section. These are the modelling objective, five data-related subsections (focal taxon/taxa, location, biodiversity data overview, predictor type, spatial and temporal scale), two conceptual subsections (hypotheses, underlying assumptions) and three technical subsections (SDM algorithms, model workflow and the software and data used; Fig. 1, Table 1). The Overview section thus briefly summarises the key information relating to the analyses. In practice, the Overview section may appear twice in scientific publications, once as flow text in the methods section of the manuscript (cf. case study 9 in Supplementary material Appendix 9) and once as part of the full ODMAP checklist (Table 1) that should always be provided in SDM studies, preferably in the appendix. To make this checklist a self-contained document, the author list and title of the study should also be specified in the Overview section of the checklist.

Table 1. The five main ODMAP sections and list of ODMAP elements. The full ODMAP v1.0 checklist is available in Supplementary material Table A1.

ODMAP section	ODMAP subsection	ODMAP elements
Overview	Authorship	Authors, contact email, title, doi
	Model objective/model purpose	SDM objective/purpose (inference, mapping, transfer), main target output
	Taxon	Focal taxon
	Location	Location of study area
	Scale of analysis	Spatial extent (lon/lat), spatial resolution, temporal extent/time period, temporal resolution, type of extent boundary (e.g. rectangular, natural, political)
	Biodiversity data overview	Observation type, response/data type
	Type of predictors	Climatic, topographic, edaphic, habitat, etc.
	Conceptual model/hypotheses	Hypotheses about biodiversity-environment relationships
	Assumptions	State critical model assumptions (cf. Table 2)
	SDM algorithms	Model algorithms, justification of model complexity, is model averaging/ensemble modelling used?
	Model workflow	Brief description of modelling steps
	Software, codes and data	Specify software, availability of codes, availability of data
Data	Biodiversity data	Taxon names, taxonomic reference system, ecological level, biodiversity data sources, sampling design, sample size per taxon, country/region mask, details on scaling, data cleaning/filtering, absence data collection, pseudo-absence and background data, potential errors and biases in data
	Data partitioning	Selection of training data (for model fitting), validation data and test (truly independent) data
	Predictor variables	State predictor variables used, data sources, spatial resolution and extent of raw data, map projection, temporal resolution and extent of raw data, data processing and scaling, measurement errors and bias, dimension reduction
	Transfer data for projection	Data sources, spatial resolution and extent, temporal resolution and extent, models and scenarios used, data processing and scaling, quantification of novel environments
Model	Variable pre-selection	Details on pre-selection of variables
	Multicollinearity	Methods for identifying and dealing with multicollinearity
	Model settings/model complexity	Models settings for all selected algorithms and for extrapolation beyond sample range
	Model estimates	Model coefficients, variable importance
	Model selection/model averaging/ensembles	Model selection strategy, method for model averaging, ensemble method
	Non-independence correction/analyses	Spatial autocorrelation in residuals, temporal autocorrelation in residuals, nested data
	Threshold selection	Details on threshold selection
Assessment	Performance statistics	Performance statistics estimated on training data, on validation data and on test (truly independent) data
	Plausibility check	Response plots; expert judgements (e.g. map display)
Prediction	Prediction output	Prediction unit; post-processing steps
	Uncertainty quantification	Uncertainty through algorithms, input data, parameters, scenarios; visualisation/treatment of novel environments

Obligatory; Objective: mapping/interpolation; Objective: forecast/transfer; Optional/context dependent.

Model objective

The Overview section should always start by specifying the modelling objective/purpose. Please note that this does not refer to the overall study objective but rather describes the specific use of the model. Following Araújo et al. (2019) we suggest clearly distinguishing between three potential purposes of modelling: 1) explanation, 2) mapping and 3) transfer. If several or all of these purposes apply, we suggest to regard these as nested (explanation < mapping < transfer). Nesting accounts for the fact that transfer should not be attempted without first having a thorough understanding of the model (inference) (Araújo et al. 2019). Importantly, several aspects of the modelling process, and with this several elements of ODMAP, vary depending on the modelling objective (Table 1). 'Explanation' (also termed inference) regards detailed analyses of species–environment relationships and aims to provide or test specific hypotheses about the main factors driving the species distributions. 'Mapping' (also termed interpolation) means that the estimated species–environment relationships are used to map (or interpolate) the species distributions in the same geographic area and time period in which the model was calibrated. 'Transfer' (also termed forecast or projection; but these terms are less precise) means that the estimated species–environment relationships are transferred to a different geographic region or time period – future or past (Yates et al. 2018). If mapping or transfer is

the goal, the main target output (prediction unit) should also be specified in Overview as this will affect other ODMAP elements.

Taxon, location, data and scale

The data-related subsections of Overview should specify the focal taxon/taxa, the location of the study, the type of biodiversity data, the type of predictors and the spatial and temporal scales of analysis (Table 1). If the study focusses on multiple species, it will be sufficient to specify the main taxon/taxa or higher category here, e.g. birds or passerines. Then, the authors should specify the type of biodiversity observation (e.g. standardized monitoring, field survey, range map, citizen science, GPS tracking) and the data/response type used (e.g. presence-only, presence–absence, counts). In addition, the type of predictor variables should be indicated (e.g. climatic, topographic, edaphic). Finally, information should be provided regarding the spatial and temporal resolution and extent of the study system. Here, we refer to the target scales of analyses while details on data processing and scaling are given in the Data section. Where relevant, the type of boundary should be indicated (e.g. rectangular within specified spatial extent, natural, political). In all cases multiple answers are possible, to allow for studies across multiple regions, taxa and data types.

Conceptual underpinning

Authors should clearly present their hypotheses about the expected biodiversity–environment relationship, meaning that they should justify what abiotic and biotic factors are taken into account to model the focal taxon, and the rationale behind these choices. Occasionally, studies may not seem to build on a priori hypotheses but may be rather exploratory, particularly when modelling many species in an automated way. Nevertheless, we encourage authors to be explicit about these conceptual considerations (Mod et al. 2016). For instance, authors could argue that they are using climatic layers in an exploratory way because climate was known to be an important driver of species distributions at a continental scale. In models that account for imperfect detection, hypotheses regarding the ecological model predictors (biodiversity–environment relationship) should be clearly separated from hypotheses regarding the observation (detectability–environment relationship) model predictors (Guillera-Arroita 2017).

Underlying model assumptions are often overlooked or unreported in SDM studies. Table 2 lists a number of typical assumptions that are often made in SDMs. We encourage authors to be specific about such underlying assumptions, because this helps reviewers or users assess the validity of the chosen approach for a given application. For example, when transferring SDMs under scenarios of global change, critical assumptions are that 1) all relevant environmental drivers are included in the model, 2) the species' observed distribution is in pseudo-equilibrium with the environment, 3) the entire realised niche is encompassed by data, 4) the correlation structure between predictors does not change between source and target landscape and 5) if extrapolation is involved, that the model extrapolates in a biologically sensible manner (Dormann et al. 2013, Elith 2017, Guisan et al. 2017, Feng et al. 2019). Other critical assumptions related to the observation process are also frequently ignored: when fitting SDMs, it is important to consider the issue of imperfect detection (Kéry 2011, Lahoz-Monfort et al. 2014). We recommend that authors be explicit about potential biases in the data (Guillera-Arroita 2017).

Technical aspects

Important technical aspects include the SDM algorithms being used, along with a verbal description of model complexity (especially if DMAP sections are moved to the Supplementary material). The choice and number of SDM algorithms contained in any study may vary depending on modelling objectives and personal experience, and as new algorithms appear. For example, when SDMs are used for transfer under scenarios of global change, some scientists advocate using ensembles of SDM algorithms to account for algorithmic uncertainty (IUCN Standards and Petitions Subcommittee 2017, Araújo et al. 2019, Thuiller et al. 2019). In contrast, if explanation or mapping is the goal, many users

Table 2. Typical model assumptions in species distribution models (Franklin 2010, Peterson et al. 2011, Guisan et al. 2017). Some of these assumptions can be relaxed by extending models accordingly. For instance, models can be built to capture occurrence dynamics, including spatial dependence, therefore relaxing the species–environment equilibrium assumption. Similarly, methods exist to address issues such as sampling bias, imperfect detection or spatial autocorrelation.

Assumption	Description
Species–environment equilibrium	Species fill their niche and do not occur elsewhere
No observation bias issues	Species data are free from observational bias (sampling bias, imperfect detection), or it is accounted for in the model
Independence of species observations	Each species record represents new information (e.g. not the same individual reported twice)
Availability of all important predictors	Key explanatory variables are available and incorporated in the model; ideally these should be proximal predictors, particularly when the objective is model transfer
Predictors are free of error	Predictors are measured (or estimated) without error
Niche stability/constancy, niche conservatism	Species retain their niches across space and time; particularly relevant when transferring predictions
No other extrapolation issues	Relationship fitted under current conditions apply when transferring predictions, even when projected beyond the range of the training data; no change in correlation structure of environmental variables; no change in key limiting processes (e.g. biotic interactions)

select only one algorithm; for example, MaxEnt (Phillips et al. 2006, Elith et al. 2011, Merow et al. 2013) has been particularly prominent in recent years. Authors should also indicate whether they use model averaging or ensemble modelling (Hao et al. 2019). Here, we do not give any recommendation as to which approach or algorithm may be more appropriate for a given application (Araújo et al. 2019). Rather, we emphasise that the modelling decisions need to be clearly described and justified, and model complexity needs to be aligned with the model objective. We define model complexity as the flexibility of the fitted biodiversity–environment relationship (cf. Merow et al. 2014, Muscarella et al. 2014, Cobos et al. 2019). Models can be more or less complex depending on the algorithm but complexity also depends on several parameter settings that determine the flexibility of the response surface (Merow et al. 2014). Specific model settings should be detailed in the Model section. Nevertheless, we encourage authors to provide a general description of model complexity as part of the Overview section, for example 'the model settings were chosen to yield simple, smooth response surfaces because we attempt extrapolation and the species may not be at equilibrium with the environment' or 'the model settings were chosen to yield complex response surfaces because our goal is to accurately map the potential species' distribution in the region and our model is based on a large enough sample size for calibrating such complex response surfaces'.

Lastly, the Overview section should contain a brief description of the overall model workflow (or point to a flowchart in main text or appendix) and information on the software packages and version, and software environment used for modelling. Importantly, the availability of codes and data needs to be specified. Here, we want to emphasise that while ODMAP supports methods reproducibility, results reproducibility can only be achieved if access to the exact data and codes are provided.

Data

The Data section provides details about the species and environmental data, and about data processing. We have identified two mandatory subsections that should always be described: biodiversity data, and environmental data (Table 1). Two other subsections are optional: data partitioning (for model assessment/evaluation), particularly important for mapping and transfer (Table 1), and transfer data (Table 1).

Biodiversity data

This subsection should contain all relevant information on the biodiversity data (Table 1). First, authors should provide the taxon name(s) and information on the taxonomic reference system (e.g. APG IV, GBIF Backbone Taxonomy), the latter being of particular importance if multiple species or taxa are being modelled (e.g. all known migratory birds, Zurell et al. 2018). Then, the focal taxonomic units being modelled should be defined. Although species are the most often used focal taxonomic unit, biodiversity models could also focus on: populations, demographic traits, supraspecific taxa, operationally defined taxa (e.g. OTUs or ASVs from barecoding), functional types, functional traits, ecological communities, community traits or species richness, among others. Likewise, studies modelling community-level properties need to specify how the community is being defined (e.g. trophic levels). Next, the data source needs to be described. If the data do not stem from one's own field surveys, then proper reference to the data source needs to be given. If the data stem from online data repositories such as GBIF (<http://data.gbif.org>) or OBIS (<http://iobis.org>), then information on accession date and/or of the source should be provided. Generally, authors should follow good data citation practices, for example as laid out by GBIF (<www.gbif.org/citation-guidelines>). Authors should also describe the underlying (spatial and temporal) sampling design and any details regarding temporal replications or nestedness of the data. This point applies to all types of biodiversity data, not only on one's own field data. If the biodiversity data stem from a standardised monitoring programme, then authors should detail here how the monitoring was carried out, e.g. how often observations were repeated and by whom (volunteers, trained volunteers, experts). If the data stem from online data repositories such as GBIF and OBIS, information should be supplied regarding the type of observations used (Anderson et al. 2016) as these databases may include mixed data from museum specimens, opportunistic observations and monitoring data. It is crucial that authors report the sample size for the focal taxa, as well as prevalence in the case of presence–absence data.

Absence, pseudo-absence and background data are an important issue for most SDM applications, and thus crucial to report in ODMAP. This is often also relevant when the response variable is abundance or species richness. It is crucial to report how these absence data were obtained and how accurate they are. Low detection probability will inevitably yield false absences (Guillera-Arroita 2017). Many SDM studies are based on presence-only data. Most algorithms then require background data (also called 'pseudo-absences' or 'quadrature points') against which the observed presences are compared (Renner et al. 2015). For example, when presence records are spatially biased, one could sample the background data such that they reflect the same spatial bias (Phillips et al. 2009, Kramer-Schadt et al. 2013). Or if GPS tracking data are used, then the background data (use versus availability) derived for each logged GPS location could be drawn dependent on empirically observed distributions of movement distances and directions (Fortin et al. 2005, Thurfjell et al. 2014). Thus, authors need to specify the geographic region from which background data are drawn, any biases induced in the background data, the number of background data points, and whether different strategies for background data derivation are used for different algorithms (Barbet-Massin et al. 2012).

Many of the remaining elements in the biodiversity data subsection are designated optional in Table 1 (i.e. context-dependent), because their necessity depends on the context of the study. We encourage authors to consider any potential

errors and biases in the data. For example, data may vary in terms of spatial and temporal precision (Meyer et al. 2016), which could significantly affect model accuracy (Park and Davis 2017). Also, any steps taken to clean and scale the data, both spatially and temporally, should be detailed here (Table 1; Daru et al. 2018). Common data cleaning steps include the removal of outliers, duplicates, records pre-dating a specified year, and records with insufficient accuracy or associated information (Serra-Diaz et al. 2017).

Data partitioning

In most SDM studies, one will assess model performance using data independent from those used for model fitting. This is most important when the model objective is mapping or transfer, although we do not wish to imply that it may not be important for explanation as well. Ideally, a modeller would use truly independent data from different sources or methods of collection to assess model performance (Araújo et al. 2005); if these are not available, it is common practice to partition data into training and testing sets. Any such data partitioning should be clearly specified in the Data section (Table 1). Following Hastie et al. (2009), we suggest to clearly distinguish 1) training data that are being used for model fitting, 2) validation data that are withheld from model fitting and are used for estimating prediction errors for model selection, model averaging and ensemble building and 3) independent test data that are used to assess the generalisation error of the final model. Here, the strategy for partitioning the data should be detailed. For example, validation data could be obtained by splitting the data randomly, or into spatially or environmentally stratified blocks for cross-validation (Roberts et al. 2017, Valavi et al. 2018). The protocol here demands description of the way the data are partitioned, and appropriate justification based on strategies to remove different types of biases in evaluation.

Environmental data

Although the overall types of environmental predictor variables (e.g. climatic, topographic) should be mentioned in the Overview section, the Environmental data subsection should clearly specify the individual environmental variables used. All relevant information should also be given about the data sources, the original spatial and temporal resolution and extent of the data as well as potential measurement errors and biases (Morueta-Holme et al. 2018). Furthermore, all data processing steps need to be described in detail, for example spatial and temporal scaling, thematic scaling (e.g. collapsing of categories), transformations and normalisations, among others (Table 1). Spatial and temporal coverage, resolution, and/or coordinate reference systems are likely to differ amongst predictor variables obtained from multiple data sources. Thus, any data harmonisation steps need to be clearly described in the Data section. In recent years, we have seen an upsurge in the availability of digitally available geoinformation relevant for biodiversity modelling, spanning climatologies (Hijmans et al. 2005, Karger et al. 2017), land cover data (Fritz et al. 2017), remote sensing data (Cord et al. 2013, Kennedy et al. 2014, Leitão and Santos 2019) and human impact data (Venter et al. 2016, Di Marco et al. 2018), among others. As these databases are under constant development, it is crucial to provide information on accession date and versioning. Importantly, these data come with different uncertainties that need to be addressed. For example, cloud cover may affect the accuracy of the remote sensing derived products such as vegetation indices. Also, when using remote sensing time series for e.g. extracting phenological metrics, different data densities along the time series imply different levels of certainty in the derived metrics. It is therefore important for authors to report how they have dealt with this problem when using remote sensing variables (Schwieder et al. 2018).

Environmental data are often subject to some form of dimension reduction, e.g. in the case of multi-collinearity (Dormann et al. 2013), meaning that not all available environmental predictors will enter the model fitting step. We suggest that if the dimension reduction is done without taking into account the response variable, then it should be part of the Data section (cf. Table 1). For example, this could be the case if principal component analysis is used to identify the main environmental axes, which are then used for modelling instead of the original environmental predictors, or where all but one of a set of highly correlated variables are dropped.

Transfer data

If the main modelling objective is to transfer the model to different geographic regions and/or different time periods (Yates et al. 2018), then authors need to report information on the data used in the model transfer, i.e. the environmental data to which the model is projected (Table 1). Analagous to the environmental data for model fitting, information should be included about the transfer data source (including accession date, version, etc.), spatiotemporal resolution and extent, data uncertainties or errors and any data processing and scaling steps. If the transfer data stem from scenario modelling, for example future or past climate scenarios (IPCC 2013) and land cover scenarios (van Vuuren and Carter 2013), it should be explicitly specified and justified which underlying models (e.g. global circulation model, regional circulation model, global vegetation model) and scenarios (e.g. representative concentration pathway, shared socioeconomic pathways) have been used.

When transferring models, we advocate for assessing environmental novelty because the transfer data may include conditions not present in the calibration data (i.e. in non-analogue situations) and, thus, the calibrated model may be forced to extrapolate (Sequeira et al. 2018). We suggest reporting how environmental novelty (Fitzpatrick and Hargrove 2009) was quantified as part of the transfer data in the Data section. Authors should specify exactly how novelty was defined and quantified; for example novel environments along single environmental gradients (Elith et al. 2010) or novel combinations of environments (Zurell et al. 2012,

Mesgaran et al. 2014). In addition to environmental novelty, modeling algorithms may also be sensitive to differences in collinearity structures of training and projection environments (Dormann et al. 2013), thus assessments of collinearity shifts can help evaluate the accuracy of model projections (Feng et al. 2019 and references therein).

Model

The Model section reports all the information necessary to repeat the model building. We have identified six subsections (see checklist in Table 1); three are mandatory (multicollinearity, model settings/model complexity, analysis of non-independence of data), two are context-dependent (variable pre-selection, model selection/model averaging/ensembles) and one is relevant for mapping and transfers only (threshold selection).

Multicollinearity and variable selection

Highly collinear variables allow alternative model structures to yield very similar model fits. The uncertainty around which environmental predictor represents the true causal mechanism may propagate into 'inflated' standard errors (Morrissey and Ruxton 2018). Different strategies exist to deal with multicollinearity, some of which will reduce the number of environmental variables to a set of reasonably correlated predictors (Dormann et al. 2013). In that sense, dealing with multicollinearity could also be seen as a data processing step. However, some strategies also involve the response variable and some preliminary model fitting, or deal with multicollinearity as part of the model building process (e.g. regularization). We thus regard it as a mandatory part of the Model section to report how multicollinearity was approached.

There may be other reasons to pre-select a specific set of predictor variables additional to reducing multicollinearity problems. For example, when attempting transfers authors may choose to limit the number of variables to avoid overfitting and achieve simpler, more transferable models (Elith et al. 2010). The strategy and rationale for selecting the final set of predictors should be clearly described here.

Model settings and model complexity

Detailing the choice of algorithms, specific model settings and model complexity is key to ensure methods reproducibility. We encourage authors to explicitly report the default settings of specific software packages (rather than just noting 'default settings were used'), as these may change based on the software version. For algorithms and estimation frameworks that rely on prior information (e.g. offsets in GLMs), prior distributions (Bayesian models) or weights, these need to be specified and justified. For Bayesian model fitting via Markov Chain Monte Carlo (MCMC) sampling, the number of MCMC samples discarded (burn-in) and kept, number of chains and convergence criterion need to be reported. When the model is being transferred, authors also need to report model settings relevant for making such spatiotemporal predictions (e.g. clamping in MaxEnt).

We recommend the use of the 'range model metadata standards' (RMMS) dictionary (Merow et al. 2019) for reporting model settings, although we acknowledge that not all potential algorithms and settings are currently included.

Model selection, model averaging and ensembles

Often, authors do not simply fit one model but consider a set of different candidate models or model algorithms, applying additional steps including model selection, model averaging or ensemble modelling. Model selection refers to situations where different model structures are compared in order to choose a single 'best' model or 'best' model set, either to improve prediction accuracy by reducing the variance of predicted values or to facilitate interpretation (Hastie et al. 2009). Different approaches such as information criterion-based variable selection and shrinkage of parameters fall into this topic. Model averaging refers to situations where different models are fit and then combined into a single prediction, which could be desirable when several candidate models are similarly plausible or because several alternative modelling approaches are available (Hastie et al. 2009, Dormann et al. 2018). Models might be averaged using an unweighted consensus method or using weighted averaging following information–theoretic, cross-validation or resampling approaches (Dormann et al. 2018). It is crucial that authors detail exactly how model selection or model averaging was carried out, including what data were used to choose amongst models. The term ensemble modelling is often used interchangeably with model averaging, but could more specifically refer to cases where, in addition to using different modelling algorithms, the initial and boundary conditions are also varied (Araújo and New 2007). Ensemble modelling is most often used in the context of making transfers (forecasting) and useful for exploring the range of predictions given the different uncertainties (initial conditions, model classes, model parameters, boundary conditions, Araújo and New 2007, Thuiller et al. 2019), but it is also increasingly used for mapping, e.g. to model rare species (Breiner et al. 2015, 2018, Hao et al. 2019). Initial conditions refer to different input data, for example when alternative species data sources are available or alternative climatologies. Boundary conditions refer to assumptions being made about changes in predictor variables in the transfer data, for example the different climate or land use scenarios as mentioned in the Data section. Similar to model averaging, different weighted and unweighted methods exist to combine the predictions (Hao et al. 2019), which should be reported.

Model estimates

We encourage authors to report whether and how model coefficients were extracted from the models and analysed, and how variable importance was determined (e.g. through permutation, Strobl et al. 2007). Further, identifying and assessing parameter uncertainty is important for guiding future work, e.g. monitoring efforts for improving the model and reducing uncertainty, and for attributing confidence to a certain model (Beale and Lennon 2012). It is thus crucial to

report how parameter uncertainty in SDMs was quantified (e.g. using asymptotic approximations based on statistical theory, or approaches based on resampling such as bootstrapping, Kéry et al. 2013).

Non-independence analysis/correction

Most standard statistical methods, and most SDM techniques, assume that the response data are random samples and, thus, that errors are independent and identically distributed. However, three common kinds of non-independence could occur in SDMs: spatial autocorrelation, temporal autocorrelation and nesting (Table 1). As a result, we mandate that authors clearly describe how non-independence in data and residuals was analysed and corrected in this subsection of the Model section of ODMAP. Spatial autocorrelation in model residuals means that predicted values at nearby locations are not independent from each other (Dormann et al. 2007) (but see Diniz-Filho et al. 2003). Analogously, temporal autocorrelation in residuals may occur when consecutive time steps are not independent from each other, which might be an issue when analysing e.g. GPS locations from animal movement data. Lastly, the assumption of independence is violated if the data contain repeated observations of the same subject or are grouped or nested in some way. For example, radio-tracking animals will yield multiple, non-independent GPS locations per individual and the locations of the same individual are likely to be more related to each other than to locations of other individuals. If several individuals have been radio-tracked in different regions, then individuals from the same region may show a more similar habitat preference than individuals from different regions. If the data are grouped in such a way, then the model needs to account for this relatedness, for example by means of random effects (Zuur et al. 2009).

Threshold selection

This subsection is important in presence-only and presence–absence models, and in particular for mapping or transfer. In the case of a binary response variable, most SDM approaches produce continuous outputs such as habitat suitability indices or probabilities of occurrence. Whilst there are good arguments for retaining predictions on a continuous scale (Lawson et al. 2014, Guillera-Arroita et al. 2015), some users prefer to threshold them for certain applications. To do so, they need to define an adequate threshold to transform the data. Several different thresholds have been proposed depending on whether the presence–absence or presence-only data are being used for modelling (Liu et al. 2005, 2013) or when modelling communities (Scherrer et al. 2018). Here, authors need to specify which threshold is used and explain why thresholding is deemed necessary.

Assessment

After model building, typically a series of analysis steps are aimed at assessing whether the modelled biodiversity–environment relationships are fit for purpose. We have designated two mandatory subsections to report in this section: performance statistics, and plausibility checks (Table 1). Irrespective of the model aim, assessing predictive performance on (semi-) independent data informs us of generalisability and overfitting (Hastie et al. 2009). Constructing partial plots (= effect plots, response curves, marginal responses, Elith et al. 2005) provides an intuitive way to evaluate the ecological plausibility of the fitted model. Plausibility could also be checked by inspecting the spatial (and/or temporal) predictions. Both plausibility checks are a form of expert judgement.

Performance statistics

Performance statistics are important for assessing the validity of a model for a specific goal and for comparing models, and different statistics are under constant development and testing. We do not wish to give advice on which performance measures should be used but rather emphasise the need to report on these performance measures and any additional information necessary to interpret them. Generally, performance should be assessed with respect to the aim of the application and to the response variable. For most response variables, e.g. for abundance and presence–absence data alike, distance measures between hold-out data and prediction are potentially suitable (Sequeira et al. 2018). These include the root–mean–square–error (RMSE), log-likelihood, various variations of R^2 (Nash and Sutcliffe 1970, Nagelkerke 1991), the percentage of deviance explained (Hosmer and Lemeshow 2013) or calibration curve estimates (with the intercept quantifying bias and the slope depicting overconfidence, Harrell 2006). If predictions were re-calibrated (Guisan et al. 2017), this should be noted as well.

For presence–absence data, we may distinguish threshold-independent measures such as the AUC (area under receiver-operating characteristic curve ROC, Swets 1988), explained deviance and log-likelihood, and threshold-dependent indices (Guisan et al. 2017). The latter are typically based on the confusion matrix (e.g. correct classification rate, sensitivity and specificity, precision, Fielding and Bell 1997), are sensitive to the prevalence and cannot be interpreted without it. When reporting thresholded indices, such as the true-skill statistic (Allouche et al. 2006) or kappa (Cohen 1960), authors must report the threshold selected and the rationale for selecting it (cf. Model section, Table 1) or whether any threshold-optimisation approach was applied (e.g. maxTSS, Guisan et al. 2017). Lastly, for presence-only methods alternative performance measures have been introduced that avoid using a confusion matrix, such as the Boyce index (Boyce et al. 2002, Hirzel et al. 2006) or POC plot (Phillips and Elith 2010).

Plausibility checks

The ecological plausibility of the model and model predictions can be checked by inspecting the response shape of the fitted biodiversity–environment relationship and by inspecting the mapped predictions. Response shapes are one of the most important outputs of SDMs as they summarise the estimated species–environment relationship and can thus be directly

subjected to plausibility checks against available biological knowledge. For example, when the input data were selected to approximate drivers known to be ecologically important, we can determine whether the model represents plausible relationships between the drivers and the species' occurrence. Checking the plausibility of the functional relationships in a model is also particularly important when the model is used to transfer the species–environment relationship to new time periods and regions (Thuiller et al. 2004). Generally, response shapes can be visualised using more traditional partial dependence plots, evaluation strips (Elith et al. 2005) or inflated response curves, which also help to identify extrapolation (Zurell et al. 2012). Furthermore, such plots indicate the range of predictor values present in the calibration data, beyond which predictions would rely on modelling assumptions (Qiao et al. 2018) and become less reliable. Ideally, such partial plots should also include a 95% confidence or credible interval. Simply plotting the predictions against the environmental predictors used in model fitting can provide a first approximation of response shapes. Additionally, visual inspection of the mapped prediction can constitute an important plausibility check for spatial models. We encourage modellers to describe any such evaluations here.

Prediction

The Prediction section of ODMAP only bears relevance if models are used to make spatial (or temporal) predictions to new sites including mapping (interpolating) and/or transferring (extrapolating). It comprises two main subsections: 1) prediction output, and 2) uncertainty quantification. Although this section deals primarily with spatial predictions, note that the final product may not necessarily be a map but could also be a data table containing the predictions at specific locations with specific environmental conditions.

Prediction output

First, prediction unit(s) should be clearly stated in ODMAP, for example continuous occurrence probabilities or potential presence derived by thresholding. Also, for some SDM algorithms there may exist alternative interpretations of outputs, e.g. MaxEnt and point process models where predictions could be interpreted as relative occurrence rates or relative densities, depending on assumptions about the data (samples of species versus samples of individuals, respectively). Second, any post-processing steps undertaken after predicting are detailed here. This could include clipping the predictions to a specific region or land cover map, e.g. clipping predicted butterfly occurrences to where the host plant is predicted to occur.

Uncertainty quantification

Studies applying SDMs for mapping and/or transferring should always address how uncertainty in model predictions was quantified. We can distinguish between uncertainties in the input data, model structure (e.g. between model algorithms), parameters, residual uncertainty (irreducible, aleatory uncertainty) and in boundary conditions (e.g. scenario uncertainty). In the Prediction section, it is important to report how any sources of uncertainty were dealt with when deriving the final prediction(s), such that maps of potential species distributions are accompanied by equivalent 'maps of ignorance' that convey how and where reliable predictions are (i.e. magnitude and extent of prediction uncertainty), thereby supporting their correct and honest interpretation (Rocchini et al. 2011). We note that suitable tools for uncertainty estimation are now readily available for all stages of the modelling process (Beale and Lennon 2012). Error propagation, for instance, is possible via bootstrapping or within Bayesian frameworks. García-Díaz et al. (2019) recommend plotting (posterior) distributions of model outputs to give a measure of the likelihood of different values that can be readily interpreted in an ecological risk assessment context.

Implementations of ODMAP involving model transfers should specify how environmental novelty was accounted for in predictions. We are aware that some overlap and confusion with the Data section could occur, which demands details on how environmental novelty was quantified (Table 1). In the Prediction section, we particularly recommend to focus on reporting any post-processing steps related to predictions, such as masking or highlighting predictions to novel environments (Zurell et al. 2012).

Applying ODMAP

Template and web application

Table 1 provides the basic template for the ODMAP (ver. 1.0) protocol (for the detailed template see Supplementary material Table A1). As indicated previously, we distinguish fields that are mandatory and fields that are optional. The mandatory fields also vary depending on the model objective (inference, mapping or transfer). That way, the ODMAP table can be filled in step by step.

To simplify use of ODMAP, we provide an interactive Shiny web application as an online resource (<https://odmap.wsl.ch>; ODMAP v1.0). This allows filling in the different ODMAP elements through a browser interface (Fig. 2). The resulting ODMAP table can be downloaded, and also uploaded again for resuming work on the ODMAP protocol. We call this version ODMAP v1.0. The ODMAP Shiny app interacts with rangeModelMetaData R-package (RMMS, Merow et al. 2019) and uses the RMMS dictionary to make auto-suggestions, for example, concerning algorithms and model settings. The app also allows existing RMMS objects to be loaded to fill in the ODMAP table. An important difference between RMMS and ODMAP is that RMMS is meant to store metadata for each model object, which could mean that several RMMS objects are needed for a single study, and RMMS also stores important results to ensure results reproducibility. By contrast, ODMAP is meant to contain the methodological descriptions for the entire SDM component of a study and is dedicated to method

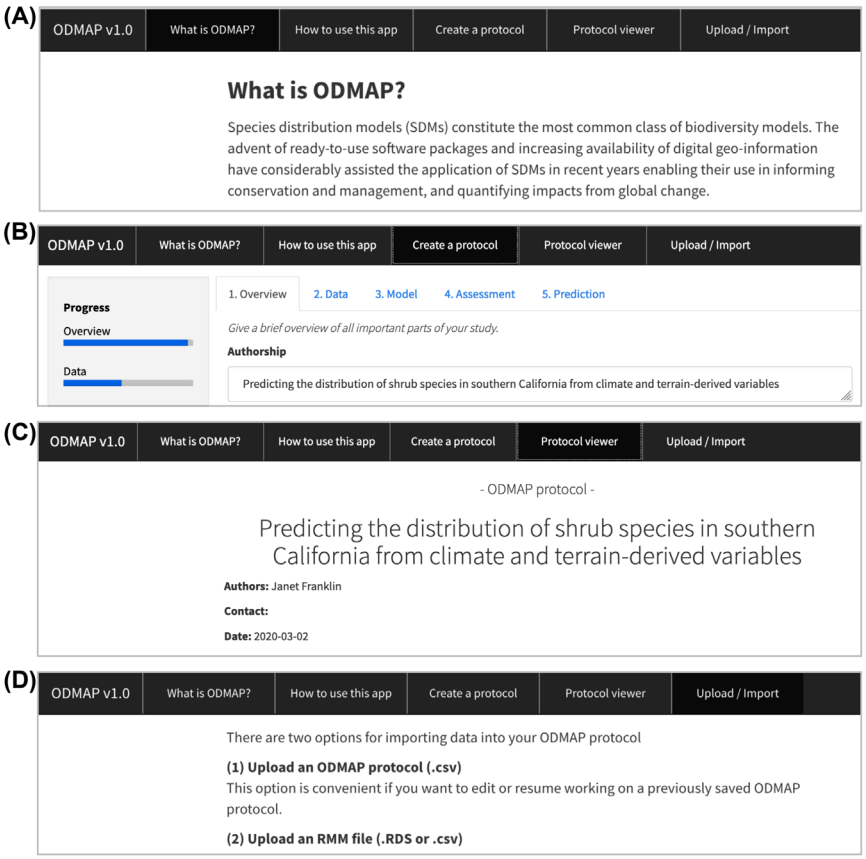

Figure 2. Screenshots of the interactive Shiny web application of ODMAP. The browser interface shows several tabs. (A) Describes the main features of ODMAP and provides the reference. (B) Contains the ODMAP core and allows entering the relevant information into the different ODMAP fields. Optional fields (cf. Table 1) can be hidden. Preliminary or finished ODMAP protocols can be downloaded as word document or as csv file. (C) The progress of ODMAP can also be assessed using the Protocol viewer. (D) Previous ODMAP protocols can be uploaded to continue protocolling or revising.

reproducibility. To accommodate these differences, we introduced an ODMAP family into the RMMS package to allow reporting for an entire study rather than single model objects.

As Merow et al. (2019) pointed out, the RMMS dictionary will need to grow through a community effort. Here, we attempted a first such effort and updated the dictionary by adding more algorithms and model settings to report. Any further updates to the dictionary will also be automatically accommodated in the ODMAP Shiny app. Similar to the ODD protocol, we anticipate that ODMAP will require regular and systematic evaluation by the scientific community to identify elements that are not being used or interpreted consistently and may potentially need updating (Grimm et al. 2010). Any future ODMAP versions will be published in the web application, with changes and updates clearly specified to ensure that older and newer ODMAP applications will remain comparable and compatible.

We recommend that the entire ODMAP checklist (e.g. obtained from filling in the template based on Table 1, or by filling in the ODMAP fields in the web application) should be provided as Supplementary material in SDM studies, indicating the ODMAP version. Additionally, we suggest that the general specifications from the Overview section should be formulated as flow text for the methods section of the main text following the structure of the Overview section in Table 1.

Case studies

The Supplementary material Appendix 1–9 includes nine example applications of ODMAP. All of these examples are taken from previously published studies, and we revised the associated model descriptions according to ODMAP. Most examples relate to terrestrial plants, birds and butterflies (Franklin 1998, Dormann et al. 2008, Schröder et al. 2009, Leitão et al. 2010, Rapacciuolo et al. 2012, Fandos and Tellería 2017, Zurell et al. 2020) but we also included a marine (Bouchet and Meeuwig 2015) and an epidemiological example (Peterson and Samy 2016). Examples cover all model objectives (inference/explanation, mapping/interpolation, forecasts/transfers), single and multiple species, different SDM algorithms as well as JSDMs. All case studies are presented as ODMAP tables (Table 1), which we would generally advise to include in the appendices of publications. In one case study (Zurell et al. 2020), we also provide an example version of the flow text that could be included in the corresponding manuscripts and reports as part of the Overview section.

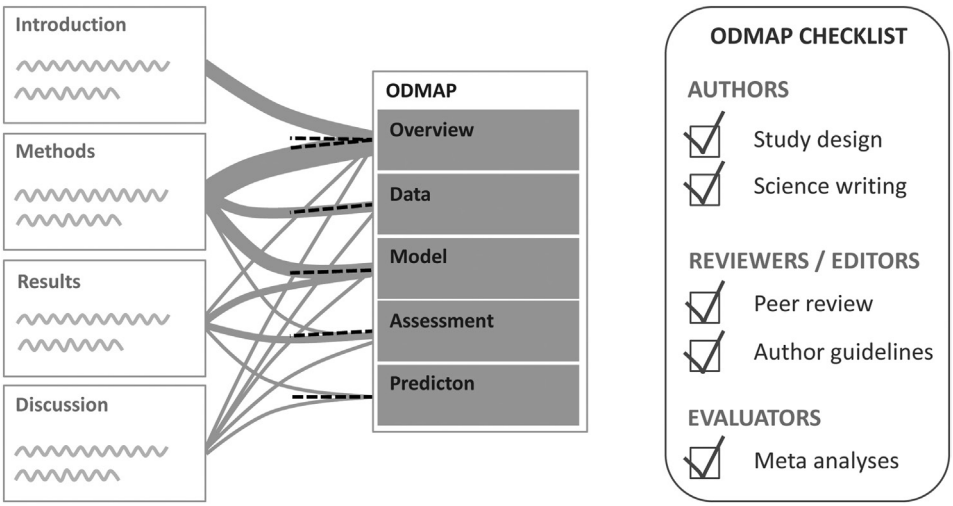

Figure 3. Schematic representation how ODMAP compiles relevant information about the SDM modelling process. Left: application of ODMAP to the case study by Franklin (1998) showed that relevant information has previously been scattered in scientific publications (grey lines) or missing (black dotted lines). Also see corresponding ODMAP protocol in Supplementary material Appendix 4. Right: ODMAP provides an easy-to-follow checklist for authors, reviewers, editors and evaluators.

In most of the case studies, we found that there had been a great deal of detail provided for the biodiversity and environmental data, and also data processing and potential biases were described in depth. Often missing were details about software versions, packages and parameter settings that would be required for reproducibility. Most of the information specified in ODMAP elements was provided in the main text of the original publications. Therefore, ODMAP relevant information was sometimes scattered across the entire publication rather than just in Methods sections (Fig. 3). When applying ODMAP, most test authors found that the protocol considerably helped identifying and structuring relevant information for model descriptions. Nevertheless, test authors also indicated that retrieving the single ODMAP elements and contents from the original publications was sometimes difficult. This emphasises that the method descriptions of SDM studies have not, to date, followed any standard structure or operating procedures to this date (Feng et al. 2019), which hampers reproducibility and peer review as well as literature reviews, expert assessments and meta-analyses (Araújo et al. 2019). It also means that ODMAP will take some time getting used to, but the overall benefits should outweigh the growing pains in the long run. Previous experience with the adoption of ODD (Grimm et al. 2006, 2010) suggests important potential benefits of such a standard protocol including more rigorous model formulation, simplified peer review, better comparability between models, easier communication between disciplines, and stronger emphasis on theoretical foundations.

Discussion

Our hope is that the ODMAP protocol can enhance transparency, reproducibility, evaluation and reuse in SDM research to facilitate peer review, meta-analyses and more robust and transparent biodiversity assessments. As the first iteration of a reporting protocol, there are likely to be improvements, refinements and disagreements. However, by developing a 'checklist' of standard operating procedures, we hope to make it easier for authors to report, and for readers to understand, SDM data and methods, as ODD has done for ABM/IBM (Grimm et al. 2006, 2010). Notably, ODMAP is not meant to prescribe how modelling should be carried out but to provide a structured format for how models should be reported. Indeed, comparability and transparency are necessary steps towards developing and applying best-practice standards for the field (Araújo et al. 2019).

Many of the authors of this protocol have played major roles in developing and refining different SDM methods, and represent a critical mass of SDM developers, users and reviewers. Based on this collective experience, we have designed ODMAP to be general enough to accommodate SDM reporting in the very broadest sense. In other words, it applies to any study using a statistical framework to explain, predict and/or project biodiversity distributions. While the specifics of the source data and methods may change for response variables other than the widely-used species occurrence data ('presence'), the requirements for reporting the conceptual underpinnings as well as the Data, Model, Assessment and Prediction sections described in Table 1 remain relevant and applicable.

ODMAP is best suited for empirical-based biodiversity models that are fitted using rule-based, statistical and machine-learning methods. Of course, also other more process-explicit distribution models exist that are used for predicting range dynamics (Zurell et al. 2016, Briscoe et al. 2019) or for testing hypotheses about deep time processes (Rangel et al. 2018). Many ODMAP elements, such as variable selection and approaches to deal with multicollinearity, will not necessarily apply to these models. Despite this, the main sections

of ODMAP – overview, data, model, assessment, prediction – could also provide a useful skeleton for describing more complex, process-explicit models, at least if the general modelling framework is published and known (Lurgi et al. 2015). By contrast, if authors are designing process-explicit models from scratch (Rangel et al. 2018), then we encourage them to use protocols such as ODD (Grimm et al. 2006, 2010), which put more emphasis on specific design decisions.

We have strived to make ODMAP as readily accessible and as easy to use as possible. The protocol explicitly includes a checklist of reporting items and is thus easy to follow and apply in practice (Fig. 1). In particular, the ODMAP table (currently, ver. 1.0) and web application provide a step-by-step guide through modelling and reporting, and integrate with current metadata standards (Merow et al. 2019). Moreover, we have designed ODMAP to apply for a broad range of modelling objectives, and our example applications provide additional guidance on how different study objectives may be reported using this same protocol. As an extra benefit, the ODMAP checklist also provides a roadmap for planning all relevant modelling steps in SDM studies. We anticipate that ODMAP will prompt researchers to consider methodological issues that tend to be more easily overlooked (e.g. uncertainty reporting) and to appropriately address key issues in the modelling process such as model validation. Identifying and addressing these issues at an early stage will ensure robust scientific results and may reduce disagreements among authors, reviewers and editors. Along these lines, we hope that ODMAP will also be positively perceived and implemented by journal editors and reviewers, who stand to benefit from an easier evaluation of the methodological aspects of SDM studies.

Standard protocols are effective tools to support decisions because they establish expectations among readers on what information should be included and where it should be found (Schmolke et al. 2010), thus ensuring that relevant information is delivered in a transparent and efficient way (Grimm et al. 2014). In this context, ODMAP is likely to help overcome two important barriers to the more frequent uptake of SDM outputs in environmental decision making: the perception of biodiversity models – including many frequently used SDMs – as 'black boxes', and the effective communication of model uncertainty (Rapacciuolo 2019).

In summary, ODMAP will help answer the clarion calls for reproducible computational science (Peng 2011), and for improved recording and reporting of methods and data (Mesirov 2010, Munafò et al. 2017), within the field of species distribution modelling, a crucial tool for science and conservation.

Funding – DZ, CK and GF acknowledge support from the German Science Foundation (DFG, grant no. ZU 361-1/1). BS acknowledges support by the German Science Foundation (grant no. SCHR1000/6-2) and by the Volkswagenstiftung and Niedersächsisches Ministerium für Wissenschaft und Kultur (METAPOLIS, ZN3121). BS and PJL acknowledge support by the joint BiodivERsA-project GreenFutureForest (grant no. 01LC1610A). WT was supported by the GAMBAS project funded by the French National Research Agency (ANR-18-CE02-0025). PB was partly funded by OPNAV N45 and the SURTASS LFA Settlement Agreement, being managed by the U.S. Navy's Living Marine Resources program under Contract no. N39430-17-C-1982. CM acknowledges funding from National Science Foundation grant DBI 1565046 and DBI 1661510. JF acknowledges the support of the National Science Foundation (USA) grant no 1853697. GGA and JE acknowledge support of the Australian Research Council via grant DP180101852.

References

Allouche, O. et al. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). – J. Appl. Ecol. 43: 1223–1232.

Anderson, R. P. et al. 2016. Are species occurrence data in global online repositories fit for modeling species distributions? The case of the global biodiversity information facility (GBIF). – Final report of the task group on GBIF data fitness for use in distribution modelling.

Araújo, M. B. and New, M. 2007. Ensemble forecasting of species distributions. – Trends Ecol. Evol. 22: 42–47.

Araújo, M. B. et al. 2005. Validation of species–climate impact models under climate change. – Global Change Biol. 11: 1504–1513.

Araújo, M. B. et al. 2019. Standards for distribution models in biodiversity assessments. – Sci. Adv. 5: eaat4858.

Austin, M. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. – Ecol. Model. 200: 1–19.

Austin, M. P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. – Ecol. Model. 157: 101–118.

Barbet-Massin, M. et al. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? – Methods Ecol. Evol. 3: 327–338.

Barbosa, F. G. and Schneck, F. 2015. Characteristics of the top-cited papers in species distribution predictive models. – Ecol. Model. 313: 77–83.

Beale, C. M. and Lennon, J. J. 2012. Incorporating uncertainty in predictive species distribution modelling. – Phil. Trans. R. Soc. B 367: 247–258.

Bouchet, P. J. and Meeuwig, J. J. 2015. Drifting baited stereo-videography: a novel sampling tool for surveying pelagic wildlife in offshore marine reserves. – Ecosphere 6: art137.

Boyce, M. S. et al. 2002. Evaluating resource selection functions. – Ecol. Model. 157: 281–300.

Breiner, F. T. et al. 2015. Overcoming limitations of modelling rare species by using ensembles of small models. – Methods Ecol. Evol. 6: 1210–1218.

Breiner, F. T. et al. 2018. Optimizing ensembles of small models for predicting the distribution of species with few occurrences. – Methods Ecol. Evol. 9: 802–808.

Briscoe, N. J. et al. 2019. Forecasting species range dynamics with process-explicit models: matching methods to applications. – Ecol. Lett. 22: 1940–1956.

Brown, J. L. 2014. Sdmtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. – Methods Ecol. Evol. 5: 694–700.

Cobos, M. E. et al. 2019. KUENM: an R package for detailed development of ecological niche models using maxent. – PeerJ 7: e6281.
Cohen, J. 1960. A coefficient of agreement for nominal scales. – Educ. Psychol. Measure. 20: 37–46.
Cord, A. F. et al. 2013. Modelling species distributions with remote sensing data: bridging disciplinary perspectives. – J. Biogeogr. 40: 2226–2227.
D'Amen, M. et al. 2017. Spatial predictions at the community level: from current approaches to future frameworks. – Biol. Rev. 92: 169–187.
Daru, B. H. et al. 2018. Widespread sampling biases in herbaria revealed from large-scale digitization. – New Phytol. 217: 939–955.
Di Marco, M. et al. 2018. Changes in human footprint drive changes in species extinction risk. – Nat. Commun. 9: 4621.
Diniz-Filho, J. A. F. et al. 2003. Spatial autocorrelation and red herrings in geographical ecology. – Global Ecol. Biogeogr. 12: 53–64.
Dormann, C. F. et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. – Ecography 30: 609–628.
Dormann, C. F. et al. 2008. Components of uncertainty in species distribution analysis: a case study of the great grey shrike. – Ecology 89: 3371–3386.
Dormann, C. F. et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. – Ecography 36: 27–46.
Dormann, C. F. et al. 2018. Model averaging in ecology: a review of bayesian, information–theoretic and tactical approaches for predictive inference. – Ecol. Monogr. 88: 485–504.
Ehrlén, J. and Morris, W. F. 2015. Predicting changes in the distribution and abundance of species under environmental change. – Ecol. Lett. 18: 303–314.
Elith, J. 2017. Predicting distributions of invasive species. – In: Robinson, A. P. et al. (eds), Invasive species. Cambridge Univ. Press, pp. 93–129.
Elith, J. and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. – Annu. Rev. Ecol. Evol. Syst. 40: 677–697.
Elith, J. et al. 2005. The evaluation strip: a new and robust method for plotting predicted responses from species distribution models. – Ecol. Model. 186: 280–289.
Elith, J. et al. 2008. A working guide to boosted regression trees. – J. Anim. Ecol. 77: 802–813.
Elith, J. et al. 2010. The art of modelling range-shifting species. – Methods Ecol. Evol. 1: 330–342.
Elith, J. et al. 2011. A statistical explanation of maxent for ecologists. – Divers. Distrib. 17: 43–57.
Fandos, G. and Tellería, J. L. 2017. Range compression of migratory passerines in wintering grounds of the western mediterranean: conservation prospects. – Bird Conserv. Int. 28: 462–474.
Feng, X. et al. 2019. A checklist for maximizing reproducibility of ecological niche models. – Nat. Ecol. Evol. 3: 1382–1395.
Ferrier, S. et al. 2016. IPBES – the methodological assessment report on scenarios and models of biodiversity and ecosystem services. – Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn, Germany.
Fielding, A. H. and Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. – Environ. Conserv. 24: 38–49.
Fitzpatrick, M. C. and Hargrove, W. W. 2009. The projection of species distribution models and the problem of non-analog climate. – Biodivers. Conserv. 18: 2255–2261.
Fortin, D. et al. 2005. Wolves influence elk movements: behavior shapes a trophic cascade in yellowstone national park. – Ecology 86: 1320–1330.
Franklin, J. 1995. Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. – Progr. Phys. Geogr. 19: 474–499.
Franklin, J. 1998. Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. – J. Veg. Sci. 9: 733–748.
Franklin, J. 2010. Mapping species distributions: spatial inference and prediction. – Cambridge Univ. Press.
Franklin, J. et al. 2017. Big data for forecasting the impacts of global change on plant communities. – Global Ecol. Biogeogr. 26: 6–17.
Fritz, S. et al. 2017. A global dataset of crowdsourced land cover and land use reference data. – Sci. Data 4: 170075.
García-Díaz, P. et al. 2019. A concise guide to developing and using quantitative models in conservation management. – Conserv. Sci. Pract. 1: e11.
Golding, N. et al. 2018. The zoon R package for reproducible and shareable species distribution modelling. – Methods Ecol. Evol. 9: 260–268.
Goodman, S. N. et al. 2016. What does research reproducibility mean? – Sci. Trans. Med. 8: 341ps312.
Grimm, V. et al. 2006. A standard protocol for describing individual-based and agent-based models. – Ecol. Model. 198: 115–126.
Grimm, V. et al. 2010. The ODD protocol: a review and first update. – Ecol. Model. 221: 2760–2768.
Grimm, V. et al. 2014. Towards better modelling and decision support: documenting model development, testing and analysis using trace. – Ecol. Model. 280: 129–139.
Guillera-Arroita, G. 2017. Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. – Ecography 40: 281–295.
Guillera-Arroita, G. et al. 2015. Is my species distribution model fit for purpose? Matching data and models to applications. – Global Ecol. Biogeogr. 24: 276–292.
Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. – Ecol. Model. 135: 147–186.
Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – Ecol. Lett. 8: 993–1009.
Guisan, A. et al. 2013. Predicting species distributions for conservation decisions. – Ecol. Lett. 16: 1424–1435.
Guisan, A. et al. 2017. Habitat suitability and distribution models with applications in R. – Cambridge Univ. Press.
Hao, T. et al. 2019. A review of evidence about use and performance of species distribution modelling ensembles like biomod. – Divers. Distrib. 25: 839–852.
Harrell, J. F. E. 2006. Regression modeling strategies. – Springer.
Hastie, T. et al. 2009. The elements of statistical learning. – Springer.
Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – Int. J. Climatol. 25: 1965–1978.
Hirzel, A. H. et al. 2006. Evaluating the ability of habitat suitability models to predict species presences. – Ecol. Model. 199: 142–152.
Hosmer, D. W. and Lemeshow, S. 2013. Applied logistic regression. – Wiley.

IPCC 2013. Climate change 2013: the physical science basis. Contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change. – Cambridge Univ. Press.
IUCN Standards and Petitions Subcommittee 2017. Guidelines for using the IUCN red list categories and criteria. – Ver. 13.
Jetz, W. et al. 2012. Integrating biodiversity distribution knowledge: toward a global map of life. – Trends Ecol. Evol. 27: 151–159.
Karger, D. N. et al. 2017. Climatologies at high resolution for the earth's land surface areas. – Sci. Data 4: 170122.
Kass, J. M. et al. 2018. Wallace: a flexible platform for reproducible modeling of species niches and distributions built for community expansion. – Methods Ecol. Evol. 9: 1151–1156.
Kennedy, R. E. et al. 2014. Bringing an ecological view of change to landsat-based remote sensing. – Front. Ecol. Environ. 12: 339–346.
Kéry, M. 2011. Towards the modelling of true species distributions. – J. Biogeogr. 38: 617–618.
Kéry, M. et al. 2013. Analysing and mapping species range dynamics using occupancy models. – J. Biogeogr. 40: 1463–1474.
Kramer-Schadt, S. et al. 2013. The importance of correcting for sampling bias in maxent species distribution models. – Divers. Distrib. 19: 1366–1379.
Lahoz-Monfort, J. J. et al. 2014. Imperfect detection impacts the performance of species distribution models. – Global Ecol. Biogeogr. 23: 504–515.
Lawson, C. R. et al. 2014. Prevalence, thresholds and the performance of presence–absence models. – Methods Ecol. Evol. 5: 54–64.
Leitão, P. J. and Santos, M. J. 2019. Improving models of species ecological niches: a remote sensing overview. – Front. Ecol. Evol. 7: 9.
Leitão, P. J. et al. 2010. Breeding habitat selection by steppe birds in Castro Verde: a remote sensing and advanced statistics approach. – Ardeola 57: 93–116.
Liu, C. et al. 2005. Selecting thresholds of occurrence in the prediction of species distributions. – Ecography 28: 385–393.
Liu, C. et al. 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. – J. Biogeogr. 40: 778–789.
Lurgi, M. et al. 2015. Modelling range dynamics under global change: which framework and why? – Methods Ecol. Evol. 6: 247–256.
Merow, C. et al. 2013. A practical guide to maxent for modeling species' distributions: what it does, and why inputs and settings matter. – Ecography 36: 1058–1069.
Merow, C. et al. 2014. What do we gain from simplicity versus complexity in species distribution models? – Ecography 37: 1267–1281.
Merow, C. et al. 2019. Species range model metadata standards: RMMS. – Global Ecol. Biogeogr. 28: 1912–1924.
Mesgaran, M. B. et al. 2014. Here be dragons: a tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. – Divers. Distrib. 20: 1147–1159.
Mesirov, J. P. 2010. Accessible reproducible research. – Science 327: 415–416.
Meyer, C. et al. 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. – Ecol. Lett. 19: 992–1006.
Michener, W. K. et al. 1997. Nongeospatial metadata for the ecological sciences. – Ecol. Appl. 7: 330–342.
Mod, H. K. et al. 2016. What we use is not what we know: environmental predictors in plant distribution models. – J. Veg. Sci. 27: 1308–1322.
Morrissey, M. B. and Ruxton, G. D. 2018. Multiple regression is not multiple regressions: the meaning of multiple regression and the non-problem of collinearity. – Phil. Theory Pract. Biol. 10: 3.
Morueta-Holme, N. et al. 2018. Best practices for reporting climate data in ecology. – Nat. Clim. Change 8: 92–94.
Munafò, M. R. et al. 2017. A manifesto for reproducible science. – Nat. Hum. Behav. 1: 0021.
Muscarella, R. et al. 2014. ENMEVAL: an R package for conducting spatially independent evaluations and estimating optimal model complexity formaxentecological niche models. – Methods Ecol. Evol. 5: 1198–1205.
Nagelkerke, N. J. D. 1991. A note on a general definition of the coefficient of determination. – Biometrika 78: 691–692.
Naimi, B. and Araújo, M. B. 2016. Sdm: a reproducible and extensible R platform for species distribution modelling. – Ecography 39: 368–375.
Nash, J. E. and Sutcliffe, J. V. 1970. River flow forecasting through conceptual models part I – a discussion of principles. – J. Hydrol. 10: 282–290.
Norberg, A. et al. 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. – Ecol. Monogr. 89: e01370.
Park, D. S. and Davis, C. C. 2017. Implications and alternatives of assigning climate data to geographical centroids. – J. Biogeogr. 44: 2188–2198.
Peng, R. D. 2011. Reproducible research in computational science. – Science 334: 1226–1227.
Peterson, A. T. and Soberón, J. 2012. Integrating fundamental concepts of ecology, biogeography and sampling into effective ecological niche modeling and species distribution modeling. – Plant Biosyst. 146: 789–796.
Peterson, A. T. and Samy, A. M. 2016. Geographic potential of disease caused by Ebola and Marburg viruses in Africa. – Acta Trop. 162: 114–124.
Peterson, A. T. et al. 2011. Ecological niches and geographic distributions. – Princeton Univ. Press.
Phillips, S. J. and Elith, J. 2010. POC plots: calibrating species distribution models with presence-only data. – Ecology 91: 2476–2484.
Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – Ecol. Model. 190: 231–259.
Phillips, S. J. et al. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. – Ecol. Appl. 19: 181–197.
Plesser, H. E. 2018. Reproducibility vs. replicability: a brief history of a confused terminology. – Front. Neuroinform. 11: 76.
Qiao, H. et al. 2018. An evaluation of transferability of ecological niche models. – Ecography 42: 521–534.
Rangel, T. F. et al. 2018. Modeling the ecology and evolution of biodiversity: biogeographical cradles, museums and graves. – Science 361: eaar5452.
Rapacciuolo, G. 2019. Strengthening the contribution of macroecological models to conservation practice. – Global Ecol. Biogeogr. 28: 54–60.
Rapacciuolo, G. et al. 2012. Climatic associations of british species distributions show good transferability in time but low predictive accuracy for range change. – PLoS One 7: e40212.

Renner, I. W. et al. 2015. Point process models for presence-only analysis. – Methods Ecol. Evol. 6: 366–379.
Roberts, D. R. et al. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical or phylogenetic structure. – Ecography 40: 913–929.
Rocchini, D. et al. 2011. Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. – Progr. Phys. Geogr. 35: 211–226.
Rodríguez-Castañeda, G. et al. 2012. Predicting the fate of biodiversity using species' distribution models: enhancing model comparability and repeatability. – PLoS One 7: e44402.
Scherrer, D. et al. 2018. How to best threshold and validate stacked species assemblages? Community optimisation might hold the answer. – Methods Ecol. Evol. 9: 2155–2166.
Schmolke, A. et al. 2010. Ecological models supporting environmental decision making: a strategy for the future. – Trends Ecol. Evol. 25: 479–486.
Schröder, B. et al. 2009. Predictive species distribution modelling in butterflies. – In: Settele, J. et al. (eds), Ecology of butterflies in europe. Cambridge Univ. Press, pp. 62–77.
Schwieder, M. et al. 2018. Landsat phenological metrics and their relation to aboveground carbon in the brazilian savanna. – Carbon Balance Manage. 13: 7.
Sequeira, A. M. M. et al. 2018. Transferring biodiversity models for conservation: opportunities and challenges. – Methods Ecol. Evol. 9: 1250–1264.
Serra-Diaz, J. M. et al. 2017. Big data of tree species distributions: how big and how good? – For. Ecosyst. 4: 30.
Strobl, C. et al. 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. – BMC Bioinform. 8: 25–25.
Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. – Science 240: 1285–1293.
Thuiller, W. et al. 2004. Effects of restricting environmental range of data to project current and future species distributions. – Ecography 27: 165–172.
Thuiller, W. et al. 2009. Biomod – a platform for ensemble forecasting of species distributions. – Ecography 32: 369–373.
Thuiller, W. et al. 2019. Uncertainty in ensembles of global biodiversity scenarios. – Nat. Commun. 10: 1446.
Thurfjell, H. et al. 2014. Applications of step-selection functions in ecology and conservation. – Movement Ecol. 2: 4.
Valavi, R. et al. 2018. BLOCK CV: an R package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. – Methods Ecol. Evol. 10: 225–232.
van Vuuren, D. P. and Carter, T. R. 2013. Climate and socio-economic scenarios for climate change research and assessment: reconciling the new with the old. – Clim. Change 122: 415–429.
Venter, O. et al. 2016. Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. – Nat. Commun. 7: 12558.
Wieczorek, J. et al. 2012. Darwin core: an evolving community-developed biodiversity data standard. – PLoS One 7: e29715.
Wüest, R. O. et al. 2018. Integrating correlation between traits improves spatial predictions of plant functional composition. – Oikos 127: 472–481.
Wüest, R. O. et al. 2020. Macroecology in the age of big data – where to go from here? – J. Biogeogr. 47: 1–12.
Yates, K. L. et al. 2018. Outstanding challenges in the transferability of ecological models. – Trends Ecol. Evol. 33: 790–802.
Zurell, D. et al. 2012. Predicting to new environments: tools for visualising model behaviour and impacts on mapped distributions. – Divers. Distrib. 18: 628–634.
Zurell, D. et al. 2016. Benchmarking novel approaches for modelling species range dynamics. – Global Change Biol. 22: 2651–2664.
Zurell, D. et al. 2018. Long-distance migratory birds threatened by multiple independent risks from global change. – Nat. Clim. Change 8: 992–996.
Zurell, D. et al. 2020. Testing species assemblage predictions from stacked and joint species distribution models. – J. Biogeogr. 47: 101–113.
Zuur, A. F. et al. 2009. Mixed effects models and extensions in ecology with R. – Springer.

Supplementary material (available online as Appendix ecog-04960 at <www.ecography.org/appendix/ecog-04960>). Appendix 1–9 + Table A1.