ECOGRAPHY

News and Views

The ODMAP protocol: a new tool for standardized reporting that could revolutionize species distribution modeling

Matthew C. Fitzpatrick, Susanne Lachmuth and Natalie T. Haydt

M. C. Fitzpatrick ✉ (https://orcid.org/0000-0003-1911-8407) (mfitzpatrick@umces.edu), S. Lachmuth (https://orcid.org/0000-0002-4027-7632) and N. T. Haydt, Appalachian Lab, Univ. of Maryland Center for Environmental Science, Frostburg, MD, USA.

Ecography
44: 1067–1070, 2021
doi: 10.1111/ecog.05700

Subject Editor: Robert P. Anderson
Editor-in-Chief: Miguel Araújo
Accepted 5 April 2021

Several recent papers have highlighted best practices, defined metadata standards and proposed checklists for reproducibility of species distribution models (SDMs). A new paper by Zurell et al. (2020) formalizes, extends and implements many of these ideas in a standardized protocol for reporting SDMs called ODMAP (Overview, Data, Model, Assessment, Prediction). ODMAP has the potential to revolutionize species distribution modeling, not through the development of new methods, but rather by providing a long overdue system to document, communicate and reproduce methods. ODMAP could lead to improved basic and applied science outcomes by increasing comprehension and transparency, and by guiding implementation of best practices. However, the impact that ODMAP has on SDMs depends in large part on the extent to which journals, reviewers and funding agencies encourage its use and whether mentors introduce it to the next generation of biodiversity scientists. Important next steps include formalizing mechanisms for the broader research community to participate in future development of ODMAP and related standards for best practices in biodiversity modeling.

Keywords: best practice, biodiversity modeling, ecological niche modeling, metadata, range, reproducibility, standardization

Species distribution models (SDMs) have become one of the most common quantitative methods in ecology, evolution and conservation, with over 1000 papers published each year on the subject (Peterson and Soberón 2012). The enormous popularity of SDMs means two things. First, as an ecologist, biodiversity scientist or biogeographer, it is likely you routinely encounter SDMs in your own work – whether you are a modeler yourself, use model outputs to inform management or policy, or simply are keeping up with the literature. Second, and more importantly, SDMs are having substantial impact outside academia as demand has grown for spatial biodiversity predictions to inform real world problems like climate change, invasive species, disease and habitat loss. Given the potential of SDMs to (mis)inform applied science especially (García-Díaz et al. 2019), it is crucial that modeling protocols are clearly documented and communicated to ensure quality and reproducibility. This need for greater transparency and defined standards has been acknowledged by major players in

NORDIC SOCIETY OIKOS

www.ecography.org

biodiversity assessment and conservation such as the IPBES (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services) and the IUCN (International Union for Conservation of Nature) (Ferrier et al. 2016, IUCN Standards and Petitions Subcommittee 2017).

Enter ODMAP, a standardized reporting protocol for SDMs proposed in a new paper by Zurell et al. (2020). What is ODMAP (Overview, Data, Model, Assessment, Prediction)? Simply put, it is a crucial advance that has been missing from the field of SDMs since, well, the beginning really. ODMAP is at once a protocol to document and communicate everything that gets baked into an SDM and also a general recipe or checklist that promotes rigor without strictly prescribing how to achieve it. ODMAP helps practitioners think through key steps and decisions in developing models and – by documenting these decisions – provides the details necessary for others to evaluate whether the resulting model aligns with the objectives of the study and its intended applications. A major plus is that ODMAP has been implemented in a user-friendly web-based application (<https://odmap.wsl.ch/>), and also can be run locally to allow for greater customization and flexibility. Either way, ODMAP automates production of SDM metadata (i.e. data documentation formatted in a hierarchical structure that catalogs various aspects of the distribution modeling process) that is ready for journal supplements. Lastly, although it focuses on SDMs (including ecological niche models, habitat suitability models, range models, etc.) ODMAP is applicable to any empirical model that couples spatially referenced biodiversity records (genetic-, species- or community-level) with environmental predictor variables.

It is important to acknowledge some history as ODMAP represents the coalescence of multiple recent efforts to characterize best practices (Araújo et al. 2019, Sofaer et al. 2019), define metadata standards and frameworks (Merow et al. 2019) and propose minimum checklists for reproducibility of SDMs (Feng et al. 2019). Along with Zurell et al. (2020), these studies should be high on the reading list of anyone working with SDMs and intending to use ODMAP's functionality. Araújo et al. (2019) provides a comprehensive discussion of best practices for making SDMs (see also Sofaer et al. 2019). Feng et al. (2019) builds aspects of best practices into a checklist and template (and in this sense can be viewed as a precursor of ODMAP), while providing background details and descriptions that can be useful for understanding the rationale for the metadata elements requested within ODMAP. Read Merow et al. (2019) for details regarding the metadata dictionary integrated into the hierarchical structure underlying ODMAP itself. In Figure 1, we provide a graphical representation of the relationships between these different best-practice tools, ODMAP and the model producer and consumer communities along the workflow from model implementation to applied use of SDMs.

ODMAP is structured into five basic steps that follow a typical SDM workflow (Franklin 2009): Overview, Data, Model fitting, Assessment, Predictions. The Overview section documents details that usually would appear in the main text of a journal article (e.g. objectives, data, scale, assumptions, algorithms and software), whereas the remaining sections comprise the more technical details. Important aspects of ODMAP include:

1. ODMAP's structure takes the form of a straightforward checklist that expands on previous standardization protocols (Feng et al. 2019, Grimm et al. 2020). Novice modelers and experts alike can efficiently structure their methodology and use the application and checklist during modeling development to guide them through the process. Those wishing to reproduce a study's methodology can simply follow the decision workflow.
2. The web-based application interfaces with the range model metadata standards (RMMS; Merow et al. 2019) through the *rangeModelMetadata* package in R to autosuggest possible information as users complete ODMAP. By incorporating an expansion of Merow et al.'s (2019) RMMS into ODMAP, the protocol becomes part of a standardized 'language' that is accepted and accessible across expertise levels and disciplines, allowing for clearer understanding and more efficient replication of SDMs.
3. ODMAP is flexible and open source. In addition to providing feedback to the creators of ODMAP, researchers can engage in developing the protocol by contributing to the *rangeModelMetadata* package (Merow et al. 2019) via GitHub. The authors of ODMAP provide an example of community-engaged development, augmenting the previously existing *rangeModelMetadata* package with ODMAP functionality so that multiple models can be documented simultaneously. How exactly community feedback is considered and implemented in ODMAP and RMMS is an important consideration that we return to below.
4. ODMAP is also flexible to individual customization. As Merow et al. (2019) point out, this flexibility has pros (e.g. user convenience and adaptability) and cons (e.g. deviance from rigid standardization). In ODMAP, the metadata elements are constrained, yet the user is able to provide custom text in some of the text boxes and can add their own sections, subsections and elements after generating a protocol. This flexibility allows ODMAP to accommodate different biodiversity modeling methods beyond standard SDMs (e.g. community-level models), while allowing researchers to include extra information necessary to assess or reproduce their research.

Importantly, ODMAP is not meant to *define* best practices or 'prescribe *how* modelling should be carried out' (Zurell et al. 2020, emphasis ours). Instead, ODMAP focuses on supporting best practices in *reporting* data and modeling choices. However, better SDMs are likely to be a side effect of ODMAP's rigorous documentation standards, for two reasons. First, ODMAP will make the work of reviewers, editors and readers much easier given that it defines standards regarding the scope and level of methodological detail expected for peer review of model quality. By reporting what was done in a standardized way, we will be in a better position to determine whether 1) a model meets minimum standards (Araújo et al. 2019, Sofaer et al. 2019) and 2) is appropriate

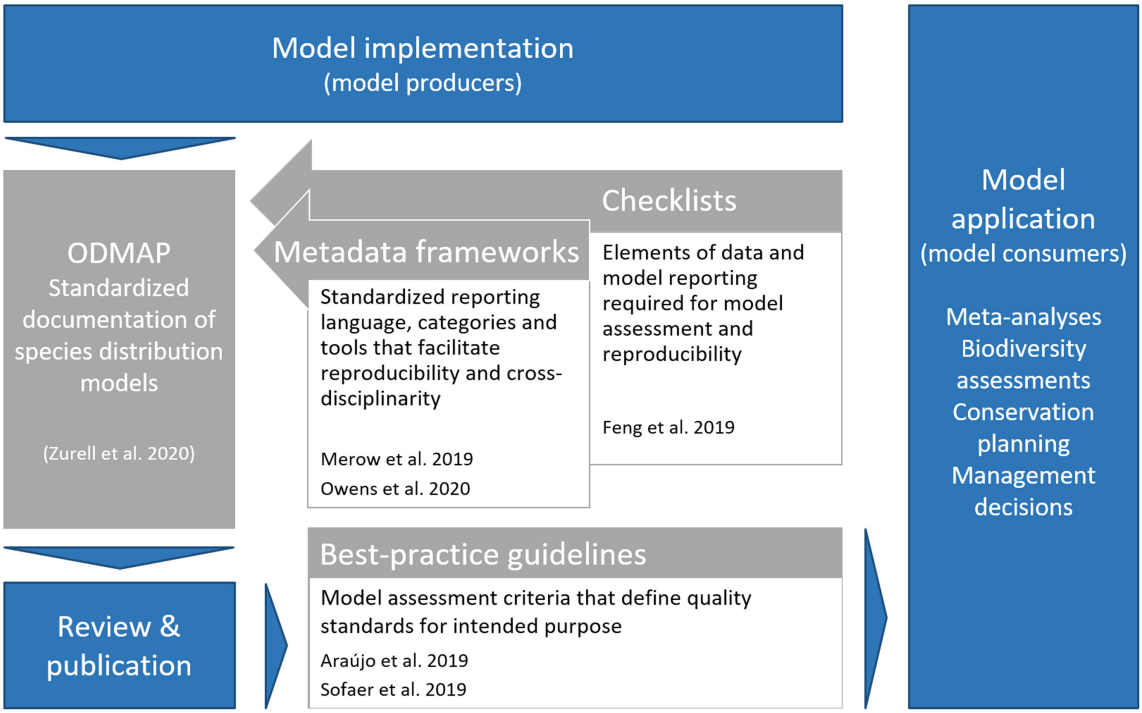

Figure 1. Schematic showing the relationship of ODMAP (Zurell et al. 2020) with existing community developed frameworks and their role in the application of species distribution models. The increasing use of these models for real world applications such as biodiversity assessments and conservation planning has spurred the development and implementation of best-practice standards and tools (light grey elements). ODMAP builds on existing checklists and metadata frameworks and serves all actors involved in an SDM workflow from implementation to application (blue elements). ODMAP guides researchers (model producers) in matching community standards and provides a rigorous control tool for reviewers and users of model outputs (model consumers). Importantly, reviewers should focus on assessing scientific quality and accurate reporting, whereas model consumers – post publication – can benefit from the standardized protocol in assessing model appropriateness for their intended purpose based on best-practice guidelines. Progress in both methodological developments and the application of model outputs necessitates constant updates to the best-practice standards and tools (all light grey elements) by the scientific and stakeholder communities.

for its intended purpose (Guillera-Arroita et al. 2015). In response to this increased transparency, modelers will be more cognizant of their choices and how these compare with best practices, even though these standards are not defined within ODMAP itself. In our own experiences working through the elements of ODMAP, we were struck by how well it enforces rigor. At the same time, one is also reminded of the inherent complexity of SDMs, the multitude of choices that must be made in fitting them, and the sensitivity of models to these choices – things that can be all too easy to forget when mass producing SDMs using available software packages.

By completing the ODMAP form, along with reading the element descriptions and examples provided therein, researchers can develop papers and supplementary materials with all information needed to allow for reproducibility of the *methods*. To reproduce the *results* of a study requires more than is currently covered by ODMAP, namely access to the actual data and scripts used to develop the model. In terms of data documentation, tools like the *occCite* package (Owens et al. 2020) in R and the development of collective stable unique identifiers for both data queries/downloads and the final processed datasets used in a biodiversity modeling analysis (Anderson et al. 2021) could help advance results reproducibility. In this sense, perhaps the most obvious area for expanding ODMAP would be formalized linking of data and code to methods documentation. Stable repositories like GitHub have become the norm for code sharing. GitHub can also host data along with code, as long as data file sizes do not become prohibitive. Data repositories like Dryad (<https://datadryad.org/stash>) and DataONE (<https://www.dataone.org/>) are better suited to host large datasets. Tierney and Ram (2020) provide a practical guide for sharing data alongside code to maximize reproducibility.

Ultimately, we expect ODMAP will have a similar impact on SDMs as the ODD protocol (Grimm et al. 2006) had on individual/agent-based modeling, upon which ODMAP in many ways is based. Reviews of ODD (Grimm et al. 2010, 2020) have documented its extensive usage and revealed that it both increased transparency and resulted in more rigorous formulation of models. We will all be better off if the same holds true for ODMAP and SDMs. In this sense, ODMAP will stand as the advance most likely to influence the direction of the field since the advent of user-friendly software for fitting SDMs – but only if all parties involved (Fig. 1) engage in using and developing ODMAP further.

That brings us to the question of: what's next? Now that ODMAP provides a standardized reporting protocol for SDMs, it is up to the community to determine its fate. We feel the future of ODMAP – and its potential to benefit SDMs – rests on two key aspects of community engagement: 1) the continued updating and improvement of the protocol and 2) ensuring that it enjoys wide usage. In terms of ongoing improvement, ODMAP is the brainchild of a critical mass of leading biodiversity scientists who played major roles in developing and refining SDM methods. As such, ODMAP represents a community effort from the start. Accordingly, the authors promise regular revisions based on community feedback to further improve ODMAP and to keep pace with methodological developments in the field – an informal strategy implemented in ODD (Grimm et al. 2020). Making ODMAP a community-driven enterprise is a commendable goal and one which the authors of ODMAP make clear they support. However, the mechanisms by which community feedback will be considered and implemented remain unclear. The same can be said for the development of agreed-upon best practices (Araújo et al. 2019, Sofaer et al. 2019): who gets to decide what constitutes a best practice and how are these deciders selected? These are important questions that lack easy answers, but which the community will have to grapple with moving forward. Our recommendation is that a working group with rotating membership be established and managed by a scientific society with global presence, such as the International Biogeography Society. We envision this working group would be responsible both for defining best practices for SDMs and for incorporating community feedback into ODMAP.

Regarding usage, ODMAP will realize its greatest impact if mentors and instructors introduce it to students, and if funding agencies, journals and reviewers begin to request ODMAP as part studies that include SDMs. We stress that the goal should be to support peer review through transparency and reporting, not to invite rejection of proposals and manuscripts simply because some aspect(s) of the study do not meet standards. Editors and reviewers must recognize that not all studies will be able to achieve the highest standard for all aspects of model development, nor should they. Instead, it should be a question of whether the model is of sufficient quality for its intended use, whether that use be in support of basic or applied science, and whether any obvious improvements could be made given data limitations and other constraints. Indeed, the best way for authors to navigate instances when they are unable to meet best practices is to acknowledge these constraints and describe what if any impact they could have on the study's objectives.

SDMs are popular in part because they make good use of data that are readily available for understanding and predicting spatial patterns of biodiversity, and thereby hold vast potential to improve our understanding of nature. Such understanding can help us address current unprecedented threats to natural systems and the services they provide humans. At the same time, SDMs, like any model, are a simplification of reality and are sensitive to data quality, statistical assumptions and numerous sources of uncertainty. Efforts like ODMAP to promote best scientific practice should make SDMs better, and in doing so, substantiate the reliability and credibility of research in this field – and that should be everyone's goal.

Acknowledgements – We thank Robert Anderson for inviting us to write this article and his thoughtful feedback on drafts. This article also was improved by comments from an anonymous reviewer.

Transparent Peer Review

The peer review history for this article is available at <https://publons.com/publon/10.1111/ecog.05700>.

References

Anderson, R. P. et al. 2021. Optimizing biodiversity informatics to improve information flow, data quality and utility for science and society. – Front. Biogeogr. 12: e47839.

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

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.

Franklin, J. 2009. Mapping species distributions: spatial inference and prediction. – Cambridge Univ. Press.

García-Díaz, P. et al. 2019. A concise guide to developing and using quantitative models in conservation management. – Conserv. Sci. Pract. 1: e11.

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. 2020. The ODD protocol for describing agent-based and other simulation models: a second update to improve clarity, replication and structural realism. – J. Artif. Soc. Soc. Sim. 23: 7.

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.

IUCN Standards and Petitions Subcommittee. 2017. Guidelines for using the IUCN red list categories and criteria. – .

Merow, C. et al. 2019. Species range model metadata standards: RMMS. – Global Ecol. Biogeogr. 28: 1912–1924.

Owens, H. L. et al. 2020. occCite: querying and managing large biodiversity occurrence datasets. – <https://cran.r-project.org/web/packages/occCite/index.html>.

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.

Sofaer, H. R. et al. 2019. Development and delivery of species distribution models to inform decision-making. – Bioscience 69: 544–557.

Tierney, N. J. and Ram, K. 2020. A realistic guide to making data available alongside code to improve reproducibility. – arXiv:2002.11626.

Zurell, D. et al. 2020. A standard protocol for reporting species distribution models. – Ecography 43: 1261–1277.