






APPLICATION

WALLACE: A flexible platform for reproducible modeling of species niches and distributions built for community expansion

Jamie M. Kass^{1,2}  | Bruno Vilela³  | Matthew E. Aiello-Lammens⁴  |
Robert Muscarella⁵  | Cory Merow⁶ | Robert P. Anderson^{1,2,7} 

¹Department of Biology, City College of New York, City University of New York, New York, NY, USA

²Program in Biology, Graduate Center, City University of New York, New York, NY, USA

³Department of Biology, Washington University in St. Louis, St. Louis, MO, USA

⁴Environmental Studies and Science, Pace University - Pleasantville-Briarcliff Campus, Pleasantville, NY, USA

⁵Department of Bioscience, Section of Ecoinformatics and Biodiversity, Aarhus University, Aarhus, Denmark

⁶Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

⁷Division of Vertebrate Zoology (Mammalogy), American Museum of Natural History, New York, NY, USA

Correspondence

Jamie M. Kass
Email: jamie.m.kass@gmail.com

Funding information

Global Biodiversity Information Facility; CUNY Graduate Center Provost's Digital Innovation Grant; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); U.S. National Science Foundation, Grant/Award Number: DBI-1650241, DBI-1661510, DEB-1119915, DEB-1046328 and DBI-1401312

Handling Editor: Robert B. O'Hara

Abstract

1. Scientific research increasingly calls for open-source software that is flexible, interactive, and expandable, while providing methodological guidance and reproducibility. Currently, many analyses in ecology are implemented with “black box” graphical user interfaces (GUIs) that lack flexibility or command-line interfaces that are infrequently used by non-specialists.
2. To help remedy this situation in the context of species distribution modeling, we created WALLACE, an open and modular application with a richly documented GUI with underlying R scripts that is flexible and highly interactive.
3. WALLACE guides users from acquiring and processing data to building models and examining predictions. Additionally, it is designed to grow via community contributions of new modules to expand functionality. All results are downloadable, along with code to reproduce the analysis.
4. WALLACE provides an example of an innovative platform to increase access to cutting-edge methods and encourage plurality in science and collaboration in software development.

KEYWORDS

biogeography, range, reproducibility, software, spatial analysis, species distribution model

1 | INTRODUCTION

Ecological and evolutionary studies have shifted over the past 20 years toward increasingly complex analyses (Bolker, 2008). This has been enabled, in part, by a rise in computing power and the increasing openness of data and software (Gimenez et al., 2014). As a result, most current methods are accessed as either: (1) programming-language scripts run in command-line interfaces (CLIs; e.g. R and PYTHON), or (2) software with graphical user interfaces (GUIs). On one hand, programming

scripts provide flexibility, but custom code is often poorly documented and tailored to specific analyses (Mislán, Heer, & White, 2016). Graphical user interfaces, on the other hand, are easy to navigate and extend accessibility of analyses to more users, but are less flexible than custom code and often necessitate using multiple software packages to complete a study. This exacerbates a problem with GUIs: lack of reproducibility (Hampton et al., 2015). Additionally, GUI implementations of methods often lag behind the cutting-edge analyses enabled by the frequent release of scripts with new publications. Hence, tools

that combine the positive aspects of CLI and GUI methods can help advance ecological research.

We developed WALLACE to address these issues specifically for user communities in ecology and the environmental sciences. WALLACE is an open-source GUI application that offers user-friendly access to R-scripted modern workflows. It is available as the R package `wallace` on CRAN, with a development version on Github (see Data Accessibility for links). WALLACE currently focuses on a workflow for modeling species niches and geographic distributions (Figure 1; Guisan & Thuiller, 2005; Peterson et al., 2011), but we anticipate that future versions will expand the analyses offered for biogeographical and macroecological modeling. WALLACE is written for R (R Core Team, 2017) using `shiny` (a package for developing interactive applications; Chang, Cheng, Allaire, Xie, & McPherson, 2017), and can thus leverage the rapidly expanding suite of R packages authored by the scientific community. Six main qualities of WALLACE distinguish it as a model for providing access and guidance for advanced methodologies (Table 1).

Below, we present several important issues in niche/distribution modeling and explain how we address them with WALLACE, first conceived as a response to the Global Biodiversity Information Facility's 2015 Ebbe Nielsen Challenge (<https://devpost.com/software/wallace-round-2>). We then provide a walkthrough of the application and conclude by discussing the general utility of WALLACE's framework for disseminating scientific methods and encouraging community-wide innovation.

TABLE 1 Advantages of the WALLACE framework

OPEN	Code is free and open-source (GNU GPL 3.0) Users can access data from online databases
EXPANDABLE	Modules (discrete methodological options) can be contributed by the community
FLEXIBLE	Multiple options exist for user data uploads and downloads of results
INTERACTIVE	Sessions are participatory and encourage experimentation A variety of visualizations are provided (maps, tables, figures)
INSTRUCTIVE	Guidance text (theoretical and methodological) is included for all components and modules
REPRODUCIBLE	An annotated and executable R Markdown file is produced for rerunning analyses, sharing results, providing supplemental information/educational resources

2 | CURRENT ISSUES IN NICHE/ DISTRIBUTION MODELING

WALLACE currently implements analyses for species niche/distribution modeling (hereafter "distribution modeling"). These correlative models estimate the response of a species to the environment and with clear assumptions can be used to infer (or hypothesize) geographic ranges, environmental suitability across a landscape, or niche requirements

The screenshot displays the WALLACE web interface. At the top is a navigation bar (1) with tabs: Wallace, Intro, 1 Occ Data, 2 Process Occs, 3 Env Data, 4 Process Envs, 5 Partition Occs, 6 Model, 7 Visualize, 8 Project, and Session Code. On the left is a toolbar (2) with 'Obtain Occurrence Data' and 'Modules Available:' options: 'Query Database' (selected) and 'User-specified'. Below this is the 'Module: Query Database' (2a) section, which includes 'Choose Database' (GBIF, VertNet, BISON), 'Enter species scientific name' (Tremarctos ornatus) (2b), and a 'Set maximum number of occurrences' slider (set to 75). A 'Query Database' button is present. Below the toolbar is a 'Download database occurrence localities (.csv)' section with a 'Download' button. At the bottom left, it lists 'Module Developers: Jamie M. Kass, Bruno Vilela, Robert P. Anderson' and 'spocc references: Package Developers: Scott Chamberlain, Karthik Ram, Ted Hart' with a link to 'CRAN | documentation'. On the right is a log window (3a) showing a message: '> Total gbif records for Tremarctos ornatus returned [75] out of [207] total (limit 75). Records without coordinates removed [39]. Duplicated records removed [3]. Remaining records [33].'. To the right of the log window is a 'Change Base Map' dropdown (3) set to 'ESRI Topo'. Below the log window is a map (3b) of South America showing occurrence points for Tremarctos ornatus. The map includes labels for countries like PANAMA, COLOMBIA, ECUADOR, PERU, VENEZUELA, GUYANA, SURINAME, FRENCH GUIANA, and BRAZIL, as well as major cities and geographical features like the AMAZON BASIN. A navigation bar at the bottom of the map area includes 'Map', 'Occs Tbl', 'Results', 'Component Guidance', and 'Module Guidance'.

FIGURE 1 The WALLACE interface with key features highlighted: (1) Navigation bar with component tabs, (2) toolbar with component name and module selection, (2a) selected module name and featured R package/s, (2b) control panel for selected module, (3) visualization space, (3a) log window, 3b) interactive map, results, and guidance text

(Franklin, 2010a; Peterson et al., 2011). Distribution modeling is used in many disciplines, such as phylogeography (Alvarado-Serrano & Knowles, 2014), community ecology (Guisan & Rahbek, 2011), evolutionary biology (McCormack, Zellmer, & Knowles, 2010), and conservation (Franklin, 2010b). At a minimum, it requires georeferenced occurrence records of the study species (e.g. from field surveys, museum collections, citizen science) and environmental predictors (e.g. climate, land cover, topography). Occurrence data generally represent the primary ecological information available for the vast majority of species. “Presence-only” distribution models use environmental values at occurrences, typically contrasting them with those available in the study region (“background” or pseudoabsence samples; Elith et al., 2011). Since reliable absence data are unavailable for most species, research focusing on presence-only models has grown tremendously over the past two decades. The current implementation of WALLACE concentrates on these models, highlighting two algorithms with differing complexity: BIOCLIM (Booth, Nix, Busby, & Hutchinson, 2014) and Maxent (Phillips, Anderson, & Schapire, 2006). Many approaches exist for making such models, and comparing them conveys to users that the utility of a model does not necessarily improve with complexity (Jiménez-Valverde, Lobo, & Hortal, 2008).

Confusion abounds regarding how best to choose and implement presence-only distribution modeling methods and interpret their outputs (Joppa et al., 2013). There have been numerous calls to address a range of complicating issues (Elith et al., 2011; Merow,

Smith, & Silander, 2013), among them sampling bias (Bean, Stafford, & Brashares, 2012), selection of study extent (VanDerWal, Shoo, Graham, & Williams, 2009), model evaluation (Radosavljevic & Anderson, 2014), model selection (Warren & Seifert, 2011), and considering key assumptions (Yackulic et al., 2013). WALLACE provides extensive guidance text and enables user experimentation with a variety of modules (see Section 4), directly addressing some of these issues and encouraging the use of a diversity of methods (Figure 2).

3 | DIFFICULTIES FOR PROGRAMMERS AND NON-PROGRAMMERS ALIKE

WALLACE combines the strengths of GUI and CLI approaches to enable research in distribution modeling for a broad audience. A number of GUI-based applications have been widely used for distribution modeling analyses (e.g. maxent.jar—Phillips et al., 2006; DesktopGARP—Scachetti-Pereira, 2002; openModeller—de Souza Muñoz et al., 2011), but an ongoing problem is that many researchers treat them like “black boxes,” even though documentation exists in the literature (Joppa et al., 2013). In addition to the shortfalls mentioned above, these GUIs lack adequate guidance within the software. Further, relying on CLIs for distribution modeling can be challenging even for specialists because it involves a combination of map inspection, spatial analysis, and statistical modeling.

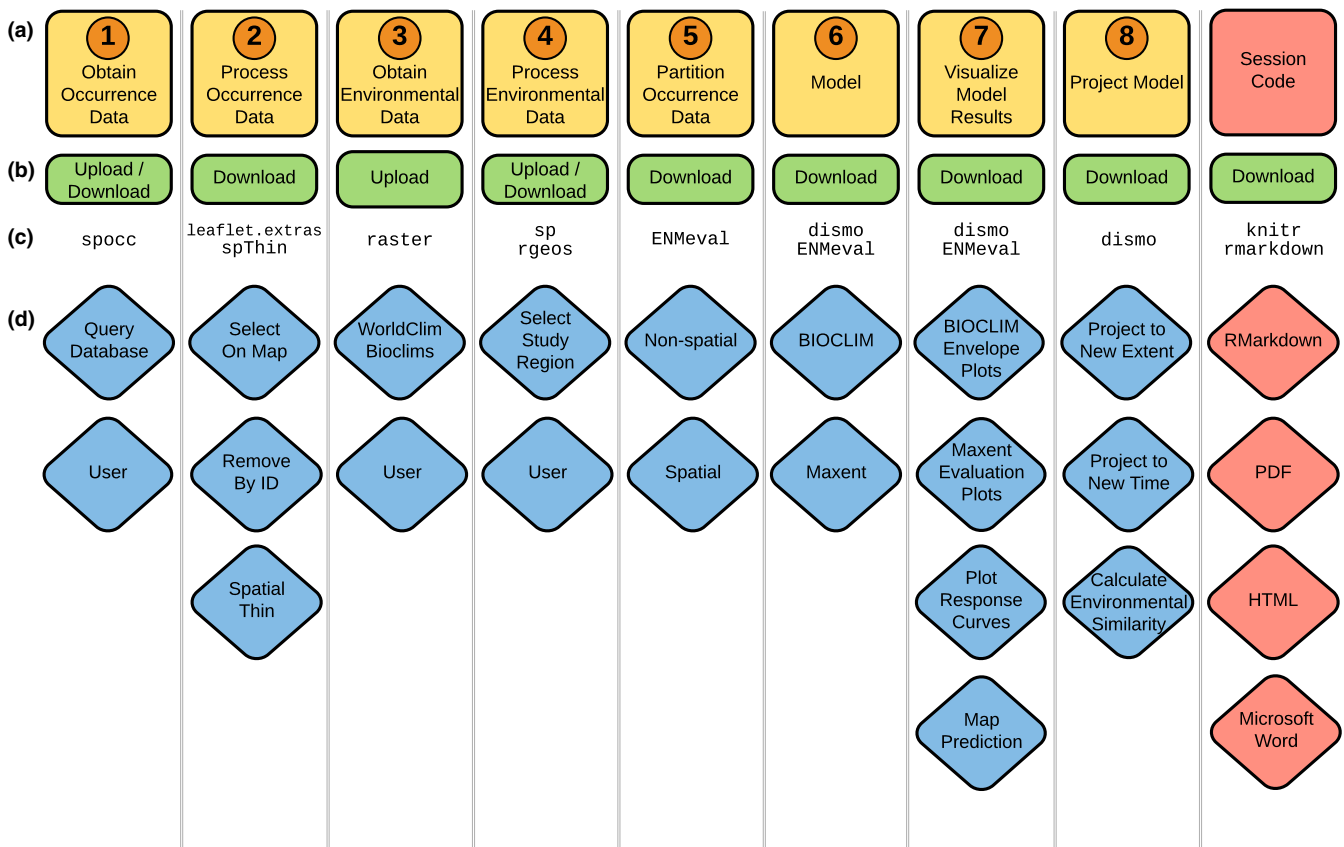


FIGURE 2 Flowchart of WALLACE features: (a) sequential components, (b) upload and download capabilities (c) R packages used, (d) modules available. The “Session Code” tab, although not a component with modules, includes multiple download options

Very recently, a number of new distribution modeling applications present exciting developments in reproducible science and indicate the demand for software that advances accessibility and collaboration (e.g. De Giovanni et al., 2016; Golding et al., 2017; Hallgren et al., 2016; Hardisty et al., 2016; Naimi & Araújo, 2016). However, those that highlight customizability and modularity require programming skills and currently lack integrated guidance on methods, while others that feature user-friendly interfaces and extensive educational resources are less flexible and have fewer opportunities for user contributions. WALLACE aims to provide a wide variety of advantages by having an easily navigable interface featuring advanced and expandable modeling tools, guidance on theory and methods, and access to the underlying code. Further development of these innovative applications and cross-collaboration among them—including WALLACE—would benefit the field greatly. Clearly no single laboratory or research group can address all the needs of biogeography and related fields, and WALLACE is designed to expand in an agile fashion as the field advances and new demands arise.

4 | WALKTHROUGH

We present a brief walkthrough of WALLACE v1.0.0, which is divided into a series of components that feature one or more modules. Module authors and featured R packages are documented in each module (Figure 2). All major modules have associated unit tests, and these will be standard with module submission going forward.

1. *Obtain Occurrence Data*: Species occurrence records can be obtained from online databases or supplied by the user. WALLACE currently accesses GBIF, VertNet, and BISON, removes duplicate coordinates, plots localities on a map, and populates a data table.
2. *Process Occurrence Data*: The user chooses which localities to include in the analysis and can address sampling bias by selecting localities on a map, removing localities by ID, or using a spatial-thinning algorithm (Aiello-Lammens, Boria, Radosavljevic, Vilela, & Anderson, 2015).
3. *Obtain Environmental Data*: For gridded predictor variables to characterize the species' response to the environment, WALLACE currently offers WorldClim bioclimatic rasters (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) or allows user-input rasters.
4. *Process Environmental Data*: The user delineates a study extent to crop the predictor grids and draw background samples, as required by most presence-only models. WALLACE offers four alternatives, with optional buffering: bounding box, minimum convex polygon, buffers around occurrence points, and user input.
5. *Partition Occurrence Data*: To evaluate models, the user chooses among (spatial and non-spatial) methods to partition occurrence localities into groups for k -fold cross-validation.
6. *Build and Evaluate Niche Models*: To examine model complexity, users can fit multiple models and use evaluation statistics to identify optimal settings (e.g. regularization multipliers, feature classes; Hijmans, Phillips, Leathwick, & Elith, 2017; Muscarella et al., 2014).

7. *Visualize Model Results*: The user can pan around the map to explore suitability predictions for the study extent, examine response curves for predictor variables, and view evaluation plots.
8. *Project Model*: The user can project models to other areas or time periods. WALLACE currently allows future projections based on the estimates of different global circulation models (Hijmans et al., 2005). Critically, users can view the magnitude of environmental novelty between the study extent and the projected area/time, which can highlight areas to exercise caution in interpretation (Elith, Kearney, & Phillips, 2010).
9. *Session Code*: The user can download an R Markdown script that reproduces the analysis undertaken during the WALLACE session.

5 | TARGET AUDIENCES

We developed WALLACE with a wide range of audiences in mind. Graduate students interested in distribution modeling and coding but who are not yet advanced programmers should benefit from learning interactively using WALLACE. Conservation practitioners and natural resource managers may want to assess data availability and quality for a study species, learn about methods, run analyses, and share results with colleagues. Experienced programmers can run models, download the session code, and customize it to modify or extend the analysis. Those developing new methods may also want to disseminate their products by contributing new modules to WALLACE. Lastly, educators can use WALLACE to teach interactive lessons about ecology, programming, and scientific best practices.

6 | CONCLUSIONS AND FUTURE DIRECTIONS

WALLACE demonstrates an innovative, open platform for rapid dissemination of scientific methods to a broad audience—specifically encouraging plurality in methodology and ongoing community development. Over the next 3 years under funding from the U.S. National Science Foundation, we plan to work closely with a cadre of international research groups to integrate new modules, both expanding available options within the existing scope of WALLACE and broadening the breadth of its capabilities. Some plans for the future include providing more environmental datasets and modeling algorithms, measuring prediction uncertainty, integrating analyses that use distribution models as inputs (e.g. measuring biodiversity, conservation planning), and model comparison tools. Above all, our vision for an expandable software like WALLACE is that users decide what needs to be added and become contributors themselves.

Although WALLACE is currently focused on distribution modeling, other fields may benefit from adopting a similar framework for software development. Like distribution models, many complex analyses can often be broken down into components and assembled into teachable workflows for disseminating methods to a broad audience. Furthermore, science advances most quickly when researchers share

advancements and build tools together, which WALLACE enables. Finally, WALLACE's interactive nature demonstrates an alternative to static manuals, tutorials, or vignettes for presenting new methods. The next generation of scientific software will benefit from these ideas, which could lead to more individuals learning, contributing to, and engaged in a dynamic process of creative collaboration.

ACKNOWLEDGEMENTS

This work was made possible via an award from the Global Biodiversity Information Facility (Finalist, 2015 Ebbe Nielsen Challenge, led by J.M.K.), a CUNY Graduate Center Provost's Digital Innovation Grant to J.M.K., and grants from the U.S. National Science Foundation (DBI-1650241, DBI-1661510, and DEB-1119915 to R.P.A.). B.V. was supported by a CAPES doctoral grant (Brazil), M.E.A.-L. by NSF DEB-1046328, and R.M. by NSF DBI-1401312. We appreciate helpful comments and suggestions from the Anderson laboratory and other members of the CUNY Ecology, Evolution, and Behavior group; Mary E. Blair; and Morgan W. Tingley.

AUTHORS' CONTRIBUTIONS

J.M.K., R.P.A., M.E.A.-L., B.V., and R.M. conceived of the original genesis of WALLACE. J.M.K. led code development, and B.V., M.E.A.-L., R.M., and C.M. were co-developers. J.M.K. drafted the manuscript, with major input from R.P.A., and all other coauthors provided revisions. R.P.A. led overall project design and drafted guidance text, on which J.M.K. provided revisions.

DATA ACCESSIBILITY

No data were included in this article. The R package is freely available under the GNU GPL-3 license at Github (development version: <https://github.com/wallaceEcoMod/wallace>) and CRAN (stable version: <https://cran.r-project.org/package=wallace>). There is also a project webpage that will be updated with ongoing development (<https://wallacecomod.github.io>).

ORCID

Jamie M. Kass  <http://orcid.org/0000-0002-9432-895X>

Bruno Vilela  <http://orcid.org/0000-0003-4072-0558>

Matthew E. Aiello-Lammens  <http://orcid.org/0000-0002-6180-5959>

Robert Muscarella  <http://orcid.org/0000-0003-3039-1076>

Robert P. Anderson  <http://orcid.org/0000-0002-7706-4649>

REFERENCES

Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, *38*, 541–545. <https://doi.org/10.1111/ecog.01132>

Alvarado-Serrano, D. F., & Knowles, L. L. (2014). Ecological niche models in phylogeographic studies: Applications, advances and precautions. *Molecular Ecology Resources*, *14*, 233–248. <https://doi.org/10.1111/1755-0998.12184>

Bean, W. T., Stafford, R., & Brashares, J. S. (2012). The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. *Ecography*, *35*, 250–258. <https://doi.org/10.1111/j.1600-0587.2011.06545.x>

Bolker, B. M. (2008). *Ecological models and data in R*. Princeton, NJ: Princeton University Press.

Booth, T. H., Nix, H. A., Busby, J. R., & Hutchinson, M. F. (2014). BIOCLIM: The first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. *Diversity and Distributions*, *20*, 1–9. <https://doi.org/10.1111/ddi.12144>

Chang, W., Cheng, J., Allaire, J. J., Xie, Y., & McPherson, J. (2017). shiny: Web application framework for R. R package version 1.0.3. Retrieved from <http://cran.r-project.org/package=shiny>

De Giovanni, R., Williams, A. R., Ernst, V. H., Kulawik, R., Fernandez, F. Q., & Hardisty, A. R. (2016). ENM components: A new set of web service-based workflow components for ecological niche modelling. *Ecography*, *39*, 376–383. <https://doi.org/10.1111/ecog.01552>

de Souza Muñoz, M. E., De Giovanni, R., de Siqueira, M. F., Sutton, T., Brewer, P., Pereira, R. S., ... Canhos, V. P. (2011). openModeller: A generic approach to species' potential distribution modelling. *Geoinformatica*, *15*, 111–135. <https://doi.org/10.1007/s10707-009-0090-7>

Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, *4*, 330–342. <https://doi.org/10.1111/j.2041-210X.2010.00036.x>

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, *17*, 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>

Franklin, J. (2010a). *Mapping species distributions: Spatial inference and prediction*. New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511810602>

Franklin, J. (2010b). Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, *16*, 321–330. <https://doi.org/10.1111/j.1472-4642.2010.00641.x>

Gimenez, O., Buckland, S. T., Morgan, B. J., Bez, N., Bertrand, S., Choquet, R., ... Mérigot, B. (2014). Statistical ecology comes of age. *Biology Letters*, *10*, 20140698. <https://doi.org/10.1098/rsbl.2014.0698>

Golding, N., August, T. A., Lucas, T. C. D., Gavaghan, D. J., van Loon, E. E., & McInerney, G. (2017). The zoon R package for reproducible and shareable species distribution modelling. *Methods in Ecology and Evolution*. <https://doi.org/10.1111/2041-210X.12858>

Guisan, A., & Rahbek, C. (2011). SESAM—a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. *Journal of Biogeography*, *38*, 1433–1444. <https://doi.org/10.1111/j.1365-2699.2011.02550.x>

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, *8*, 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>

Hallgren, W., Beaumont, L., Bowness, A., Chambers, L., Graham, E., Holewa, H., ... Vanderwal, J. (2016). The biodiversity and climate change virtual laboratory: Where ecology meets big data. *Environmental Modelling & Software*, *76*, 182–186. <https://doi.org/10.1016/j.envsoft.2015.10.025>

Hampton, S. E., Anderson, S. S., Bagby, S. C., Gries, C., Han, X., Hart, E. M., ... Mudge, J. (2015). The Tao of open science for ecology. *Ecosphere*, *6*, 120. <https://doi.org/10.1890/ES14-00402.1>

Hardisty, A. R., Bacall, F., Beard, N., Balcázar-Vargas, M. P., Balech, B., Barcza, Z., ... Dobor, L. (2016). BioVeL: A virtual laboratory for data analysis and modelling in biodiversity science and ecology. *BMC Ecology*, *16*, 49. <https://doi.org/10.1186/s12898-016-0103-y>

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land

- areas. *International Journal of Climatology*, 25, 1965–1978. [https://doi.org/10.1002/\(ISSN\)1097-0088](https://doi.org/10.1002/(ISSN)1097-0088)
- Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2017). *dismo: Species distribution modeling*. R package version 1.1-4. Retrieved from <http://cran.r-project.org/package=dismo>
- Jiménez-Valverde, A., Lobo, J. M., & Hortal, J. (2008). Not as good as they seem: The importance of concepts in species distribution modelling. *Diversity and Distributions*, 14, 885–890. <https://doi.org/10.1111/j.1472-4642.2008.00496.x>
- Joppa, L. N., McInerney, G., Harper, R., Salido, L., Takeda, K., O'Hara, K., ... Emmott, S. (2013). Troubling trends in scientific software use. *Science*, 340, 814–815. <https://doi.org/10.1126/science.1231535>
- McCormack, J. E., Zellmer, A. J., & Knowles, L. L. (2010). Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation?: Insights from tests with niche models. *Evolution*, 64, 1231–1244. <https://doi.org/10.1111/j.1558-5646.2009.00900.x>
- Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, 36, 1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>
- Mislan, K. A. S., Heer, J. M., & White, E. P. (2016). Elevating the status of code in ecology. *Trends in Ecology & Evolution*, 31, 4–7. <https://doi.org/10.1016/j.tree.2015.11.006>
- Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution*, 5, 1198–1205. <https://doi.org/10.1111/2041-210X.12261>
- Naimi, B., & Araújo, M. B. (2016). sdm: A reproducible and extensible R platform for species distribution modelling. *Ecography*, 39, 368–375. <https://doi.org/10.1111/ecog.01881>
- Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martinez-Meyer, E., Nakamura, M., & Araújo, M. B. (2011). *Ecological niches and geographic distributions*. Princeton, NJ: Princeton University Press.
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>
- Radosavljevic, A., & Anderson, R. P. (2014). Making better Maxent models of species distributions: Complexity, overfitting and evaluation. *Journal of Biogeography*, 41, 629–643. <https://doi.org/10.1111/jbi.12227>
- Scachetti-Pereira, R. (2002). *DesktopGarp: A software package for biodiversity and ecologic research*. Lawrence, KS: The University of Kansas Biodiversity Research Center. Retrieved from <http://www.nhm.ku.edu/desktopgarp/>
- VanDerWal, J., Shoo, L. P., Graham, C., & Williams, S. E. (2009). Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling*, 220, 589–594. <https://doi.org/10.1016/j.ecolmodel.2008.11.010>
- Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21, 335–342. <https://doi.org/10.1890/10-1171.1>
- Yackulic, C. B., Chandler, R., Zipkin, E. F., Royle, J. A., Nichols, J. D., Campbell Grant, E. H., & Veran, S. (2013). Presence-only modelling using MAXENT: When can we trust the inferences? *Methods in Ecology and Evolution*, 4, 236–243. <https://doi.org/10.1111/2041-210x.12004>

How to cite this article: Kass JM, Vilela B, Aiello-Lammens ME, Muscarella R, Merow C, Anderson RP. WALLACE: A flexible platform for reproducible modeling of species niches and distributions built for community expansion. *Methods Ecol Evol.* 2018;9:1151–1156. <https://doi.org/10.1111/2041-210X.12945>