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

Jamie M. Kass, Bruno Vilela, Matthew E. Aiello-Lammens, Robert Muscarella, Cory Merow & Robert P. Anderson …

… and Jenna Rios, Sara Varela, Hannah L., Owens, Brian S. Maintner, Olivier Broennimann, Mary E. Blair, Ned Horning, Jorge Velásquez-Tibatá, Gonzalo E. Pinilla-Buitrago, Beth E. Gerstner, Peter J. Galante, Sarah I. Meenan, Valentina Grisales-Betancur, Cecina Babich-Morrow, Erica E. Johnson …

... and many collaborating institutions!

Wallace v1 overview presentation, CC BY SA license, 10 December 2018
Range estimates needed
Applied biodiversity informatics

Anderson (2012)

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Issue: Blavatnik Awards for Young Scientists

Harnessing the world’s biodiversity data: promise and peril in ecological niche modeling of species distributions

Robert P. Anderson¹,²,³,⁴

Agenda: Making data and modeling ready to address critical environmental issues of the 21st century
1. High-quality data, ready to be accessed when the particular problem presents itself

2. Software that achieves an appropriate balance between automation and supervision

3. Scientists capable of building, applying, and appraising high-quality models
Correlative modeling of species niches and ranges

Occurrence records
Environmental data
Algorithm

Model of niche

Suitable areas
Correlative modeling of species niches and ranges

**Algorithm**

- Occurrence records
- Environmental data

**Model of niche**

- Suitable areas
- another time
- another place

Conservation, invasive species, zoonotic diseases, climate change
EL MODELADO DE NICHOS Y DISTRIBUCIONES: NO ES SIMPLEMENTE “CLIC, CLIC, CLIC”

MODELING NICHE AND DISTRIBUTIONS: IT’S NOT JUST “CLICK, CLICK, CLICK”

LA MODÉLISATION DE NICHE ET DE DISTRIBUTIONS: CE N’EST PAS JUSTE “CLIC, CLIC, CLIC”
Software that achieves an appropriate balance between automation and supervision

- *automates* repetitive aspects
- *forces* user to make critical biological and conceptual decisions
- *general* with respect to algorithms used
Problems with existing analyses

Existing GUIs

- (too) easy
- default settings
- inflexible
- infrequently updated
- strings of applications
Problems with existing analyses

Existing GUIs

- (too) easy
- default settings
- inflexible
- infrequently updated
- strings of applications

Code

- error-prone
- not easily generalizable
- slow to be taken up
From R to a web app

Innovative use of GBIF occurrence data
Combine code-based methods with GUI

Foreground:
Flexible, interactive, expandable GUI

Background:
R code from new packages: reproducible
WALLACE: A flexible platform for reproducible modeling of species niches and distributions built for community expansion

Jamie M. Kass\textsuperscript{1,2} | Bruno Vilela\textsuperscript{3} | Matthew E. Aiello-Lammens\textsuperscript{4} | Robert Muscarella\textsuperscript{5} | Cory Merow\textsuperscript{6} | Robert P. Anderson\textsuperscript{1,2,7}
Wallace: currently under expansion …

Jamie M. Kass, Bruno Vilela, Matthew E. Aiello-Lammens, Robert Muscarella, Cory Merow & Robert P. Anderson

… and Jenna Rios, Sara Varela, Hannah L., Owens, Brian S. Maintner, Olivier Broennimann, Mary E. Blair, Ned Horning, Jorge Velásquez-Tibatá, Gonzalo E. Pinilla-Buitrago, Beth E. Gerstner, Peter J. Galante, Sarah I. Meenan, Valentina Grisales-Betancur, Cecina Babich-Morrow, Erica E. Johnson …
What is Wallace?

Point-and-click (GUI) application that includes most steps of a niche/distribution modeling workflow

- Harnesses R packages and gives credit to their authors
R packages on the inside
… and now many other contributors
Wallace
Qualities of Wallace

Wallace’s code is free and **OPEN**

(& users can download data from online databases.)
Qualities of Wallace

Wallace provides **GUIDANCE** that addresses conceptual and methodological issues.
Wallace is **FLEXIBLE**

by providing multiple options, and allowing user inputs and downloads for most components.
Qualities of Wallace

Wallace features INTERACTIVE maps, tables, and graphs to explore data and model predictions.
Qualities of Wallace

Wallace is **REPRODUCIBLE** by providing executable code for documenting and rerunning the analysis.
Wallace is **EXPANDABLE!**

by featuring different methodological options (modules) that researchers can contribute to advance functionality.
Wallace: software for niche/distribution modeling

**Qualities:**
- open
- instructive
- flexible
- interactive
- reproducible
- expandable

**Broadening human diversity**

Female and minority-student coder-biologists

External partners, undergraduates, and interns/fellows:

Sara Varela, Hannah L. Owens, Cecina Babich-Morrow, Jenna Rios, Sarah I. Meenan, Valentina Grisales-Betancur
Installing *Wallace*

- Install R and (optionally) Rstudio
- For Maxent java version, install maxent.jar (https://biodiversityinformatics.amnh.org/open_source/maxent/) and appropriate Java version for your system, then place maxent.jar in `dismo` folder
- Install the `wallace` package from CRAN

```
# install the package
install.packages('wallace')
# load the package
library(wallace)
# run the app
run_wallace()
```

- Troubleshooting on Github page (https://github.com/wallaceEcoMod/wallace)
Wallace today

1. Full Release 1.0.6 (Kass et al. 2018, Methods in Ecology & Evolution software note; Google group; Wallace e-mail)

2. Ongoing NSF funding to work with external partners to add new modules

3. Ongoing NASA funding, led by Mary Blair (AMNH), to develop new R packages and add them to Wallace and interface with BioModelos

Mary E. Blair

NSF DEB-1119915 & DBI-1650241

NSF DBI-1661510

NASA 80NSSC18K0406
Acknowledgments

https://wallaceecomod.github.io/

http://www.andersonlab.ccny.cuny.edu