

A List of Species Distribution Modeling Software

One of the great advantages of using open source software, such as R, is the multitude of packages available. This helps you avoid reinventing the wheel when working, even if it is in a relatively obscure area. Species Distribution Modeling (or Ecological Niche Modeling) is a field that has embraced R, and this has resulted in a number of packages. This list aims at providing a comprehensive list of software available, together with tutorials, source code, article references and possibly other materials.

Modeling

SDM

- website: <https://cran.r-project.org/web/packages/sdm/index.html>
- documentation: <https://cran.r-project.org/web/packages/sdm/sdm.pdf>
- paper: <https://doi.org/10.1111/ecog.01881>
- tutorials: https://cran.r-project.org/web/packages/sdm/vignettes/quick_sdm.pdf, <http://rsatial.org/sdm/>

An extensible framework for developing species distribution models using individual and community-based approaches, generate ensembles of models, evaluate the models, and predict species potential distributions in space and time.

Dismo

- website: <https://cran.r-project.org/web/packages/dismo/index.html>
- documentation: <https://cran.r-project.org/web/packages/dismo/dismo.pdf>

This package implements a few species distribution models, including an R link to the ‘maxent’ model, and native implementations of Bioclim and Domain. It also provides a number of functions that can assist in using Boosted Regression Trees.

Biomod2

- website: <https://cran.r-project.org/web/packages/biomod2/index.html>
- documentation: <https://cran.r-project.org/web/packages/biomod2/biomod2.pdf>
- tutorials: https://rdr.io/cran/biomod2/man/BIOMOD_Modeling.html

BIOMOD is a computer platform for ensemble forecasting of species distributions, enabling the treatment of a range of methodological uncertainties in models and the examination of species-environment relationships. BIOMOD includes the ability to model species distributions with several techniques, test models with a wide range of approaches, project species distributions into different environmental conditions (e.g. climate or land use change scenarios) and dispersal functions. It allows assessing species temporal turnover, plot species response curves, and test the strength of species interactions with predictor variables. BIOMOD is implemented in R and is a freeware, open source, package.

MaxEnt

- website: https://biodiversityinformatics.amnh.org/open_source/maxent/
- source: <https://github.com/mrmaxent/Maxent>
- paper: <https://www.cs.princeton.edu/~schapire/papers/ecolmod.pdf>

- tutorials: https://biodiversityinformatics.amnh.org/open_source/maxent/Maxent_tutorial2017.pdf, https://web.stanford.edu/~hastie/Papers/maxent_explained.pdf, <https://www.cs.cmu.edu/afs/cs/user/aberger/www/html/tutorial/tutorial.html>

The Maxent software is based on the maximum-entropy approach for modeling species niches and distributions. From a set of environmental (e.g., climatic) grids and georeferenced occurrence localities (e.g. mediated by GBIF), the model expresses a probability distribution where each grid cell has a predicted suitability of conditions for the species. Under particular assumptions about the input data and biological sampling efforts that led to occurrence records, the output can be interpreted as predicted probability of presence (cloglog transform), or as predicted local abundance (raw exponential output).

DOMAIN

- paper: <http://link.springer.com/article/10.1007/BF00051966>
- tutorials: <http://spatialecology.weebly.com/r-code--data/category/sdm-domain>

The DOMAIN procedure computes potential distributions using a point-to-point similarity metric (Gower metric) to assign a classification value to a candidate site based on the proximity in environmental space of the most similar record site. If you believe the impact of serious outliers is an issue with you data, DOMAIN can be directed to use the mean of a number of largest similarity values rather than the maximum similarity value. The DOMAIN model is particularly well suited to applications where available site location records or environmental data are limited.

ENMEval

- source: <https://github.com/bobmuscarella/ENMEval>
- documentation: <https://cran.r-project.org/web/packages/ENMEval/ENMEval.pdf>
- paper: <https://doi.org/10.1111/2041-210X.12261>
- tutorials: <https://cran.r-project.org/web/packages/ENMEval/vignettes/ENMEval-vignette.html>

R package for automated runs and evaluations of ecological niche models.

sdmplay

- website: <https://cran.r-project.org/web/packages/SDMPlay/index.html>
- documentation: <https://cran.r-project.org/web/packages/SDMPlay/SDMPlay.pdf>
- tutorials: <https://cran.r-project.org/web/packages/SDMPlay/vignettes/my-vignette.html>

Functions provided by this pedagogic package allow to compute models with two popular machine learning approaches, BRT (Boosted Regression Trees) and MaxEnt (Maximum Entropy) applied on sets of marine biological and environmental data. They include the possibility of managing the main parameters for the construction of the models. Classic tools to evaluate model performance are provided (Area Under the Curve, omission rate and confusion matrix, map standard deviation) and are completed with tools to perform null models. The biological dataset includes original occurrences of two species of the class Echinoidea (sea urchins) present on the Kerguelen Plateau and that show contrasted ecological niches. The environmental dataset includes the corresponding statistics for 15 abiotic and biotic descriptors summarized for the Kerguelen Plateau and for different periods in a raster format. The package can be used for practicals to teach and learn the basics of species distribution modelling. Maps of potential distribution can be produced based on the example data included in the package, which brings prior observations of the influence of spatial and temporal heterogeneities on modelling performances. The user can also provide his own datasets to use the modelling functions.

virtualspecies

- website: <https://cran.r-project.org/web/packages/virtualspecies/index.html>
- documentation: <ftp://cran.r-project.org/pub/R/web/packages/virtualspecies/virtualspecies.pdf>
- paper: <https://doi.org/10.1111/ecog.01388>
- tutorials: <http://borisleroy.com/files/virtualspeciestutorial.html>

Provides a framework for generating virtual species distributions, a procedure increasingly used in ecology to improve species distribution models. This package integrates the existing methodological approaches with the objective of generating virtual species distributions with increased ecological realism.

MigClim

- website: <https://cran.r-project.org/web/packages/MigClim/index.html>
- documentation: <https://cran.r-project.org/web/packages/MigClim/MigClim.pdf>

Functions for implementing species dispersal into projections of species distribution models (e.g. under climate change scenarios).

Interactive

Zoon

- website: <https://zoonproject.wordpress.com/>
- source: <https://github.com/zoonproject/zoon>
- documentation: <https://cran.r-project.org/web/packages/zoon/zoon.pdf>
- paper: <https://doi.org/10.1111/2041-210X.12858>
- tutorials: <https://zoonproject.wordpress.com/tutorials/>

The zoon r package has been developed specifically to improve reproducibility and comparability of SDMs in r by allowing users to encode entire SDM analyses as repeatable and extensible workflows consisting of independently executable, community-contributed modules. The module-workflow structure enables scientists to more easily create and share components of their analysis, and then, access, modify, reuse and combine the components of others. While zoon's modular nature is similar to other sdm r packages such as BioMod2 and sdm, zoon pulls each module from an open repository that any SDM user can contribute to and makes it easy for non-developers to contribute modules. zoon's focus on reproducible and modifiable workflows is inspired by repositories such as the Cardiac Web Lab and workflow systems such as Taverna and the biovel system, but embeds an SDM-specific workflow system within r, making it much more accessible to the SDM community.

Wallace

- website: <https://wallaceecomod.github.io/>
- source: <https://github.com/wallaceEcoMod/wallace>
- documentation: https://cmerow.github.io/RDataScience/3_4_wallace.html
- paper: <https://doi.org/10.1111/2041-210X.12945>

Wallace is a modular, R-based platform for reproducible modeling of species niches and distributions. The application guides users through a complete analysis, from the acquisition of data to visualizing model predictions on an interactive map, thus bundling complex workflows into a single, streamlined interface.

Visualisation

BDVis

- website: <https://cran.r-project.org/web/packages/bdvis/index.html>
- source: <https://github.com/vijaybarve/bdvis>
- paper: <https://doi.org/10.1093/bioinformatics/btw333>

Provides a set of functions to create basic visualizations to quickly preview different aspects of biodiversity information such as inventory completeness, extent of coverage (taxonomic, temporal and geographic), gaps and biases.

Data

rgbif

- website: <https://cran.r-project.org/web/packages/rgbif/index.html>
- source: <https://github.com/ropensci/rgbif>
- documentation: <https://cran.r-project.org/web/packages/rgbif/rgbif.pdf>
- tutorials: https://cran.r-project.org/web/packages/rgbif/vignettes/rgbif_vignette.html

A programmatic interface to the Web Service methods provided by the Global Biodiversity Information Facility ('GBIF'; <https://www.gbif.org/developer/summary>). 'GBIF' is a database of species occurrence records from sources all over the globe. 'rgbif' includes functions for searching for taxonomic names, retrieving information on data providers, getting species occurrence records, and getting counts of occurrence records.

Note: for suggestions feel free to contact me and I will add them.