ECOGRAPHY

Software notes

GARPTools: R software for data preparation and model evaluation of GARP models

Catherine G. Haase, Anni Yang, Kristina M. McNyset and Jason K. Blackburn

C. G. Haase, Dept of Biology, Austin Peay State Univ., Clarksville, TN, USA. – CGH, A. Yang and J. K. Blackburn (https://orcid.org/0000-0003-0928-4831) ✉ (jkblackburn@ufl.edu), Spatial Epidemiology and Ecology Research Laboratory, Dept of Geography, Univ. of Florida, Gainesville, FL, USA. AY and JKB also at: Emerging Pathogens Inst., Univ. of Florida, Gainesville, FL, USA. – K. M. McNyset, Silverton, OR, USA.

Ecography
44: 1790–1796, 2021
doi: 10.1111/ecog.05642

Subject Editor: Thiago F. Rangel
Editor-in-Chief: Miguel Araújo
Accepted 13 August 2021

The aim of the GARPTools package is to provide tools to prepare data for input into the desktop version of the genetic algorithm for rule-set prediction (GARP), for the evaluation of the accuracy of output models, and for summary/examination of environmental coverages used in GARP rule sets for best models in an experiment. GARP is a software package for biodiversity and ecological research that allows the user to predict and analyze wild species' geographic distributions. GARP is a presence-background genetic algorithm that models species' potential geographic distributions through an iterative process of training and testing that occurs through resampling and replacement of input data. GARP develops rule-sets of if/then logic statements that assign presence or absence, which are then mapped on the landscape. Toward this, GARPTools provides preparation functions including splitting species presence locations into training and testing data sets and resampling environmental layers to the same spatial resolution (raster cell size) and extent. Model evaluation functions are relevant to current procedures applied to species distribution modeling, including the receiver operating characteristic curves and omission and commission indices. There are also functions to estimate the contribution of environmental coverages (covariates) based on the GARP outputs by examining the individual rules. GARPTools intends to provide a means to systematically prepare data, evaluate models, and summarize environmental coverages among multiple systems and species without the need for complex software programming.

Keywords: AUC, data preparation, ecological niche modeling, GARP, model evaluation, ROC, R software

Background

Ecological niche modeling approaches identify non-random relationships between point data describing known locations reporting species' occurrences and gridded (raster) covariates describing the environment (Peterson 2001, Blackburn 2010). These approaches result in presence/absence maps that describe the likelihood that a pixel presents suitable environmental conditions for the species on the landscape. There is a diversity of modeling approaches, such as logistic regression (Kessler et al. 2018),

boosted regression trees (Elith et al. 2008), and random forest models (Evans et al. 2011) – each requiring input of presence and absence locations. Using these approaches, the user may have confirmed absence locations or generate pseudoabsence data with geographic information system (GIS) tools (Kracalik et al. 2017). Presence-background modeling approaches, such as the genetic algorithm for rule-set prediction (GARP) (Stockwell and Peters 1999) and MaxEnt (Phillips and Dudík 2008), derive pseudoabsence iteratively within the algorithm, requiring the user to only provide input data on species' occurrences (presence) and covariates. In all the modeling approaches, users are required to perform extensive pre-processing to prepare input and covariate data ahead of model development. Additionally, many tools, such as GARP, require the user to post-process data to perform accuracy assessments to evaluate experiment quality and predictive success. Similarly, Cobos et al. (2019) identified this need for that MaxEnt system and introduced the kuenm R package.

Here we present a set of pre- and post-processing tools for the desktop version of the GARP algorithm. The GARP approach employs a superset of modeling approaches, whereby the type of model (e.g. logistic regression or range rules) are selected and applied iteratively using the rules of genetics to add, delete, revise, and develop extensive rulesets relating presence or absence to ecological conditions. The algorithm has been described in detail elsewhere (Stockwell and Peters 1999) and used extensively to predict species distributions for diverse taxa and pathogens across the landscape since the late 1990s. In several papers, authors have examined the individual rules output from the desktop version of GARP, DesktopGARP ver. 1.1.3 (McNyset 2005, Blackburn et al. 2007) and described variable contribution using early python-based tools (Joyner 2010, Mullins et al. 2011, Barro et al. 2016). However, there has yet to be a comprehensive way to prepare and evaluate model outputs in a standard format. Here we focus on an R software package to assist users in preparing data for and standardizing the evaluation of model outputs from GARP experiments.

Aim of GARPTools

The GARPTools package provides tools to prepare data for input into the desktop version of GARP and for the evaluation of the accuracy of output models. DesktopGARP is a software package for biodiversity and ecological research that allows the user to predict and analyze wild species' distributions. GARP is a presence-background genetic algorithm that models species' potential geographic distributions through an iterative process of training and testing that occurs through resampling and replacement of input data (Pereira 2002). In particular, the package GARPTools was designed to support:

1) Data preparation for raster layers and occurrence data
   - Resampling of environmental layers to the same spatial scale and extent
   - Cropping of environmental layers to study area
   - Removal of duplicate presence points within a single raster cell
   - Splitting presence data into training (for input into GARP) and testing (for model evaluation in R) data points
2) Accuracy metrics evaluation
   - Summation of the best models generated by the best subset procedure in GARP into a single best subset raster with model agreements
   - Calculation of the confusion matrix, including model sensitivity and specificity
   - Assessment of model accuracy with AUC, the AUC ratio, commission, and omission metrics
   - Plotting of the receiver operating characteristic curve
3) Summary of environmental coverages and variable contribution
   - Extraction of dominant presence rules for the best models in the best subset generated by GARP
   - Plotting the rescaled median range of environmental variables in the dominant presence rules
   - Estimation of variable contributions in GARP models

The package is particularly useful when running multiple GARP experiments for multiple species; the data prep and accuracy tools were developed to allow for use during an iterative process. It is also useful when using environmental layers with multiple spatial scales and extents while focusing on a specific study area.

Overview of R functions and available datasets

The main functions available in the GARPTools package supporting data input, model evaluation, and ruleset extractions are described in Table 1. Figure 1 illustrates the work flow for implementing GARPTools steps in R before and after running GARP experiments in the DesktopGARP software.

Species location dataset and environmental layers

The package provides both species location data and environmental variable layers that can be used to demonstrate data preparation for DesktopGARP (ver. 1.1.03), model evaluation, visualization of environmental variables in the best model, and estimation of variable contributions. GARPTools includes location data for white-tailed deer *Odocoileus virginianus* available from GBIF.org, a free and open access facility for global biodiversity data. This dataset is available in both data frame (*wtdeer_df*) and shapefile formats (*wtdeer_locations*) and includes 169 GPS locations from North Carolina, USA. Environmental layers available include gridded raster layers for altitude (*alt*), annual mean temperature (*bio_1*), and annual precipitation (*bio_12*) with a spatial resolution of 1 km^2 and are available for download from WorldClim (<www.worldclim.org/>), a database for global climate data (as described in Hijmans et al. 2005).

Table 1. Main functions and their descriptions of the GARPTools package.

Function	Description
aucGARP	Calculates the area under the receiver operating characteristic curve
centroid	Returns the centroid coordinates of raster cells containing presence locations
checkData	Checks that presence locations overlap environmental layers
commissionGARP	Calculates the commission or the proportion of the landscape predicted as present when the species has not been reported as present, of the best models of GARP
confuseMatrix	Returns the confusion matrix containing the sensitivity and specificity of the best model subset of GARP
extractRules	Extracts the dominant presence (negated) range rules and logit rules for the best model subset
findModRule	Finds the dominant rules (both presence and absence rules) in the best models produced by GARP
getMedMinMax	Returns the minimum and maximum median values for variables in dominant presence rules from a given rule-set
omissionGARP	Calculates the omission or the true occurrence points that are not predicted as present, of the best models output by GARP
plotRange	Plots the rescaled median ranges for environment variables used across dominant presence rules
plotROC	Plots the receiver operating characteristic curve of the best model subset, relating 1-model specificity to model sensitivity
prevalence	Calculate the prevalence of each environmental covariate in the dominant presence rules of best models
ptsSum.df	Summarizes the number of original sampling points, the number of remaining points in the spatially unique selection process, and the number of points reduced down from the centroid function
ptsSum.sp	Get a list of points that remain in the spatially unique selection process and the points that are removed from the centroid function
rasterPrep	Resamples input rasters to the same cell size and extent of a sample raster
rescale	Rescales the minimum and maximum median values for each environment variable used in dominant presence rules to be between 0 and 1
seAUC	Returns the standard error of the AUC of the best model subset output
splitData	Uses a user-defined percentage to split presence locations into training and testing datasets
sumRasters	Sums the best models into a single raster that represents the number of models that agree on a predicted presence location
totPresRules	Get the total number of dominant presence rules in the ruleset file output of the best models
unimportIdx	Calculate the unimportance index of each environmental covariate for the evaluation of the variable contribution
wGARP	Returns the Wilcoxon test statistic of the best model subset output
zAUC	Returns the z-score associated with the AUC of the best model subset

Data preparation for DesktopGARP

One of the key functionalities of GARPTools is to prepare both species location and environmental covariate data files to the formats required for DesktopGARP. More specifically, GARPTools provide functions to resample raster layers to the same spatial extent and resolution (function *resample*), finding and filtering multiple presence locations within a single raster cell (functions *ptsSum.df* and *centroid*, respectively) and split presence locations into training and testing datasets (function *splitData*).

Like most ecological niche modeling applications, GARP requires all environmental covariates to have the same spatial resolution and extent. Models with higher resolution datasets usually perform better over lower resolution data (Ross et al. 2015); therefore using layers with multiple resolutions may skew the relative influence of variables on predicting presence (Scott et al. 2002). The *rasterPrep* function resamples the raster layers of environmental covariates to the same spatial resolution (i.e. cell size) and extent of a user-defined raster using the bilinear method due to the continuous values.

GARP requires only a single presence location to occur within a pixel of an environmental layer since it does not consider abundance when assessing presence. Multiple points within a single pixel may inflate the importance of the environmental variable that occurred at that location (Scott et al. 2002) and will result in an upload error in DesktopGARP. Therefore, the *centroid* function returns the centroid location of all raster cells that contain a presence location, which circumvents multiple locations. The function uses the cell size and extent of user-defined raster to generate centroid coordinates of all raster cells that contain presence locations that can then be used in GARP. The function is useful when input rasters' cells are large (and thus contain multiple sampling locations) or when sampling was unevenly clustered.

A GARP experiment is by design a random walk (Stockwell and Peterson 2002) because there is no single solution due to the nature of random walks; GARP is evaluated by the total number of models in an experiment (a set of models) that predict a pixel as present. The *splitData* function randomly splits presence locations into training and testing subsets outside of GARP based on a user-defined percentage and allows for multiple iterations of the subsampling process to use in multiple GARP models. The training data are then used as input occurrence data in the GARP experiment, and the testing data are reserved externally for model evaluation.

Model evaluation: receiver operating characteristic, omission, and commission indices

Model variation associated with random walks can be sifted through a 'best subsets' procedure in the GARP model (Anderson et al. 2003). The best models are selected by estimating the omission and commission rates of each model

based on the internal training and testing sets split inside GARP and the user-defined thresholds (Anderson et al. 2003). The discrimination capacity of the best models is assessed by the area under the curve (AUC) of the receiver operative characteristic (ROC) using the external testing set which we hold out from the model building (Fielding and Bell 1997). The AUC evaluates the discrimination capacity of the best models by relating the model sensitivity (true positive rate) to 1-specificity (true negative rate). AUC determines if the model set is discriminating better than a randomly model (AUC = 0.5) and can be described as the probability that a randomly selected presence has a higher model output value than a randomly selected absence (DeLong et al. 1988).

The *aucGARP* function calculates the AUC for predicting presence locations; as presence is a discrete variable, individual cutpoints of the best subset models act as thresholds of the ROC curve. By calculating the AUC for a finite number of potential thresholds (here defined by the *n* argument), it is equivalent to calculating the non-parametric Wilcoxon's statistic of model evaluation (Phillips et al. 2006).

We also provide a means to calculate a modified version of the AUC and compare it to a modified reference value, noted as the AUC ratio. AUC can often undervalue certain niche models, such as those provided by GARP, and lead to incorrect comparisons among model variations (Lobo et al. 2008). More specifically, AUC is highly influenced by the spatial extent of the models and summarizes the model performance over ROC ranges that the model would rarely operate. Additionally, AUC weighs both types of error (omission and commission) equally, which is not adequate in most applications of distribution modeling. Thus, the following methods outlined by Peterson et al (2008), we also provide a means to calculate partial-area ROC and AUC ratios that provide a means to better compare these models with the *aucRatio* function.

GARPTools also provides for calculations of internal accuracy metrics of omission (false negatives; real occurrence points not predicted; *omissionGARP* function) and commission (the proportion of the landscape where the species has not been reported as present but is predicted as present; *commissionGARP* function) to select the best performing models in the experiment (Scott et al. 2002). Total and average commission and omission are calculated based on the confusion matrix (extracted by *confuseMatrix* function). Total commission is calculated as the total number of pixels predicted present across all the best models in the best subset generated by GARP, divided by the total number of unoccupied pixels in the study area. The average commission is calculated as the average of the total number of pixels predicted present, divided by the total number of unoccupied pixels within the study area on a model-by-model basis for each of the models of the best subset. The total omission is calculated as the total number of independent test points predicted absent by the summated grid of all the best models. The average omission is calculated as an average omission across each of the best models.

Extraction of environmental coverages and estimation of variable contribution based on prevalence and median range of covariates

It is of primary importance to understand the environmental coverages of the species' ecological niche. The GARPTools package provides a series of functions to extract and visualize the environmental coverages estimated by GARP models. There are four types (range, negated range, logit, and atomic) of if-then genetic rules considering different combinations of environmental variables in GARP (Stockwell and Noble 1992). Each model representing a genetic ruleset includes 50 rules which predict the presence and absence of the species for each pixel. Extraction of rules from the best models helps estimate ecological requirements of species survival and the potential effects of the covariates on the prediction of the species' distribution. Here we only focus on the presence rules since the absence rules predicting the absence of the species normally have erratic coverages (Mullins et al. 2011). The *findModRules* function finds the dominant rules (both presence rules and absence rules) for each model, or those that together predict over 90% of the landscape, and the *extractRules* function extracts the dominant presence rules for each GARP model. Finally, the *getMedMinMax* function summarizes the median range of the environmental variables in the dominant rules of the best models, which can also be rescaled to 0–1 using the *rescale* function. As noted by Yang et al. (2020), the median range is the difference between the median range from the upper and lower bounds of a ruleset. Rescaling allows for the quantitative comparison of the median ranges for different variables at the same level (Mullins et al. 2011).

The estimation of environmental coverages can also be used as a potential criterion for evaluating variable contributions in GARP models. Most species distribution models incorporate the algorithms or tools to estimate the variable contributions, for example, the ANOVA table in regression technique, the relative influences in boosted regression trees (BRTs), and the regularized training gain with jackknifing in MaxEnt (Araujo and Guisan 2006, Phillips et al. 2006, Elith et al. 2008). Although GARP incorporates a jackknife procedure, there is no integrated rubrics/methodologies to evaluate the variable contributions. Here, we provide two metrics as proxies to assess the variable contribution in GARP models: 1) the prevalence of the environmental variable in the dominant presence rules of the best models and 2) its median range (as described above).

We adopt a similar idea of relative influences in BRTs to define the prevalence of the covariates in a GARP model. The relative influences in BRTs are defined as the frequency the variable is used to split the individual trees, weighted by the squared improvement resulting from the final trees (Friedman and Meulman 2003). Here, we define the prevalence of the environmental covariate in the dominant presence rules of the GARP best subset as the frequency the environmental covariate is used to predict the presence of the species. It can be calculated using the *prevalence* function.

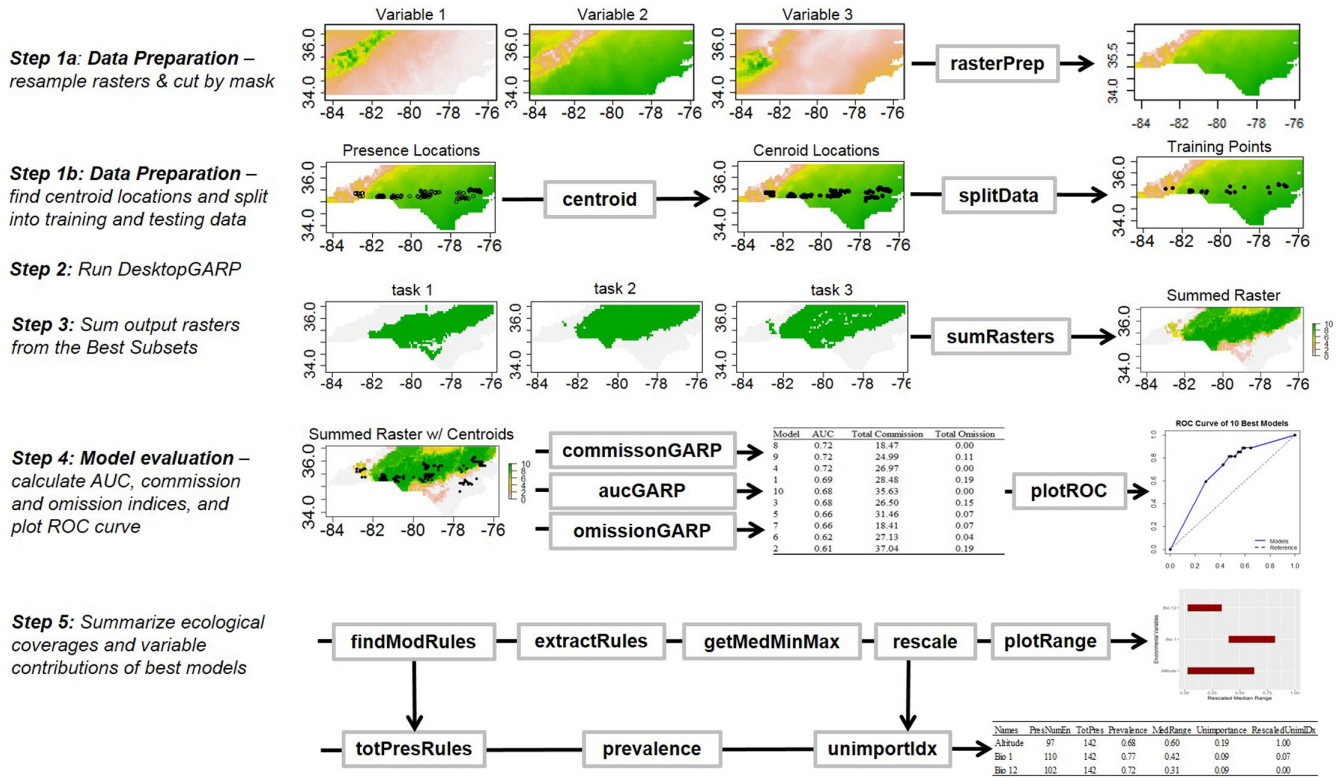

Figure 1. A flow chart describing the functionality of many functions available in the package GARPTools: Step 1a is for raster preparation; Step 1b is for occurrence location preparation.

We use an index to consider both criteria, prevalence and median range, to assess the variable contribution. The variables in the GARP models with the least contribution are those with a wide median range and low prevalence. Thus, the index of unimportance for a covariate in a GARP model is defined as:

$$\text{Unimportance} = (1 - \text{prevalence}) \times \text{rescaled median range}$$

and can be calculated using the *unimportIdx* function. The multiplication of the rescaled median range and the frequency the covariate is not used to predict the presence of the species in dominant presence rules will exclude variables with 1) high prevalence and wide median range, and 2) low prevalence and narrow median range. Variables with high prevalence are frequently used to predict the presence of the species. Variables with a narrow median range suggest that the survival of species is constrained to some specific conditions regarding this covariate. The aforementioned functions and tools to extract rules and estimate variable contribution are specifically designed for DesktopGARP 1.1.03 version which has rulesets recorded as part of the output.

Conclusions

There are multiple advantages to using GARP over other distribution modeling techniques, specifically those models that require absence locations, which can be difficult to obtain accurately for most species. Additionally, GARP is a machine learning technique, which does not require assumptions about the data structure often necessary in linear modeling. Likewise, GARP uses four different types of rulesets to describe relationships between predictor variables and species' presence (referred to as a superset), and GARPTools allow the user to compare the rulesets selected. GARPTools also provides a stepping off point for increased functionality of GARP and highlights the value of GARP as one of many modeling tools available to the ecological niche and/or species distribution modeling community.

Often incorrect comparisons are made between GARP and other species distribution modeling programs, such as Maxent (Phillips et al. 2006) and BIOMOD (Thuiller et al. 2009), specifically when comparing the accuracy metrics noted by AUC values. As noted by Peterson et al. (2008), the ROC curve computed by GARP is not completely characterized, and thus cannot be directly compared to ROC curves output by other modeling programs. ROCs often favor algorithms that are not limited in their range of commission errors and models that define absence differently than GARP. Previous work has noted that GARP often can outperform these other programs when comparisons are made over AUC ratios estimated by partial ROC tests (Encarnación-Luévano et al. 2013). Thus we now include a means to compare the AUC ratios among models with a function to calculate these ratios as described by Peterson et al. (2008).

As in any modeling technique, GARP does have some disadvantages, specifically in the ease of application and appropriateness for all species. The desktop version of GARP can be difficult to configure for non-Windows operating systems, and often requires the use of grids output by ArcView. However, GARP is a free tool with a user-friendly interface, which is often not the case in other software programs. Additionally, GARPTools provides the tools to streamline the data generation process required for GARP. GARP models often have difficulty in modeling distribution of isolated or sessile organisms because of the relevancy of spatial scale of the environmental layers (Stockman et al. 2006). However, overprediction of isolated or rare species can be an issue in many modelling techniques, and care must be taken in including the critical delimiting factors for these species' distributions.

Here we provide extensive data preparation and model evaluation tools for the desktop version of GARP. These tools relieve some of the disadvantages in using GARP compared to other species' distribution modeling techniques and provide a means to prepare data and evaluate output models with ease.

To cite GARPTools or acknowledge its use, cite this Software note as follows, substituting the version of the application that you used for 'version 1.0':

Haase, C. G. et al. 2021. GARPTools: R software for data preparation and model evaluation of GARP models. – Ecography 44: 1790–1796 (ver. 1.0).

Acknowledgements – We thank the editors and two blind reviewers for assistance with the manuscript. C Griffin provided firsthand experience and suggestions from early versions of GARPTools that were critical to developing the vignette.
Funding – This project was sponsored by the US NIH EEID r01GM117617 to JKB.

Author contributions

Catherine Haase: Conceptualization (supporting); Data curation (equal); Formal analysis (equal); Methodology (lead); Software (lead); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Anni Yang**: Formal analysis (equal); Methodology (equal); Software (equal); Writing – original draft (supporting); Writing – review and editing (supporting). **Kristina McNyset**: Conceptualization (equal); Methodology (equal); Software (supporting); Writing – review and editing (supporting). **Jason Blackburn**: Conceptualization (equal); Data curation (equal); Formal analysis (supporting); Funding acquisition (lead); Investigation (lead); Methodology (equal); Project administration (lead); Resources (equal); Software (supporting); Supervision (lead); Validation (equal); Visualization (supporting); Writing – original draft (equal); Writing – review and editing (equal).

Transparent Peer Review

The peer review history for this article is available at <https://publons.com/publon/10.1111/ecog.05642>.

Data availability statement

All of the code presented in GARPTools is stored within an R package housed on GitHub at <https://rdrr.io/github/cghaase/GARPTools/>. All the data for the program are stored within the package and described in a vignette within the package.

References

Anderson, R. P. et al. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. – Ecol. Model. 162: 211–232.

Araujo, M. B. and Guisan, A. 2006. Five (or so) challenges for species distribution modelling. – J. Biogeogr. 33: 1677–1688.

Barro, A. S. et al. 2016. Redefining the Australian anthrax belt: modeling the ecological niche and predicting the geographic distribution of *Bacillus anthracis*. – PLoS Negl. Trop. Dis. 10: e0004689.

Blackburn, J. 2010. Integrating geographic information systems and ecological niche modeling into disease ecology: a case study of *Bacillus anthracis* in the United States and Mexico. – In: O'Connell, K. P. et al. (eds), Emerging and endemic pathogens: advances in surveillance, detection and identification. Springer, pp. 59–88.

Blackburn, J. et al. 2007. Modeling the geographic distribution of *Bacillus anthracis*, the causative agent of anthrax disease, for the contiguous United States using predictive ecologic niche modeling. – Am. J. Trop. Med. Hyg. 77: 1103.

Cobos, M. E. et al. 2019. kuenm: an R package for detailed development of ecological niche models using Maxent. – PeerJ 7: e6281.

DeLong, E. R. et al. 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. – Biometrics 44: 837–845.

Elith, J. et al. 2008. A working guide to boosted regression trees. – J. Anim. Ecol. 77: 802–813.

Encarnación-Luévano, A. et al. 2013. Activity response to climate seasonality in species with fossorial habits: a niche modeling approach using the lowland burrowing treefrog *Smilisca fodiens*. – PLoS One 8: e78290.

Evans, J. S. et al. 2011. Modeling species distribution and change using random forest. – In: Drew, C. et al. (eds.), Predictive species and habitat modeling in landscape ecology. Springer, pp. 139–159.

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.

Friedman, J. H. and Meulman, J. J. 2003. Multiple additive regression trees with application in epidemiology. – Stat. Med. 22: 1365–1381.

Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – Int. J. Climatol. 25: 1965–1978.

Joyner, T. A. 2010.Ecological niche modeling of a zoonosis: a case study using anthrax and climate change in Kazakhstan.MSc thesis, Univ. of Florida, Gainesville, FL, USA. 150 p.

Kessler, W. H. et al. 2018. Estimating the geographic distribution of host-seeking adult Amblyomma americanum (Acari: Ixodidae) in Florida. – J. Med. Entomol. 56: 55–64.

Kracalik, I. T. et al. 2017. Modeling the environmental suitability of anthrax in Ghana and estimating populations at risk: implications for vaccination and control. – PLoS Negl. Trop. Dis. 11: e0005885.

Lobo, J. M. et al. 2008. AUC: a misleading measure of the performance of predictive distribution models. – Global Ecol. Biogeogr. 17: 145–151.
McNyset, K. 2005. Use of ecological niche modelling to predict distributions of freshwater fish species in Kansas. – Ecol. Freshwater Fish 14: 243–255.
Mullins, J. et al. 2011. Ecological niche modelling of the *Bacillus anthracis* A1. A sub-lineage in Kazakhstan. – BMC Ecol. 11: 32.
Pereira, R. S. 2002. Desktop GARP. – Univ. of Kansas Biodiversity Research Center, Lawrence, KS, USA.
Peterson, A. 2001. Predicting species' geographic distributions based on ecological niche modeling. – Condor 103: 599–605.
Peterson, A. T. et al. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. – Ecol. Model. 213: 63–72.
Phillips, S. and Dudík, M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. – Ecography 31: 161–175.
Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – Ecol. Model. 190: 231–259.
Ross, L. K. et al. 2015. The influence of data resolution on predicted distribution and estimates of extent of current protection of three 'listed' deep-sea habitats. – PLoS One 10: e0140061.
Scott, J. M. et al. 2002. Predicting species occurrences: issues of accuracy and scale. – Island Press.
Stockman, A. K. et al. 2006. An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. – Divers. Distrib. 12: 81–89.
Stockwell, D. and Peters, D. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. – Int. J. Geogr. Inf. Sci. 13: 143–158.
Stockwell, D. R. B. and Noble, I. R. 1992. Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. – Math. Comput. Simul. 33: 385–390.
Stockwell, D. R. B. and Peterson, A. T. 2002. Controlling bias in biodiversity data. – In: Scott, J. M. et al. (eds), Predicting species occurrences: issues of scale and accuracy. Island Press, pp. 537–546.
Thuiller, W. et al. 2009. BIOMOD – a platform for ensemble forecasting of species distributions. – Ecography 32: 369–373.
Yang, A. et al. 2020. Exploring environmental coverages of species: a new variable contribution estimation methodology for rulesets from the genetic algorithm for rule-set prediction. – PeerJ 8: e8968.