Volume 18 Issue 10
Originally issued as B1



Global Journal of human-social science: B
Geography, Geo-Sciences, Environmental Science & Disaster Management
Volume 18 Issue 1 Version 1.0 Year 2018
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-460X & Print ISSN: 0975-587X

Assessing Accuracy Methods of Species Distribution Models: AUC, Specificity, Sensitivity and the True Skill Statistic

By Farzin Shabani, Lalit Kumar & Mohsen Ahmadi
University of New England

Abstract- We aimed to assess different methods for evaluating performance accuracy in species distribution models based on the application of five types of bioclimatic models under three threshold selections to predict the distributions of eight different species in Australia, treated as an independent area. Five discriminatory correlative species distribution models (SDMs), were used to predict the species distributions of eight different plants. A global training data set, excluding the Australian locations, was used for model fitting. Four accuracy measurement methods were compared under three threshold selections of *i)* maximum sensitivity + specificity, *ii)* sensitivity = specificity and *iii)* predicted probability of 0.5 (default). Results showed that the choice of modeling methods had an impact on potential distribution predictions for an independent area. Examination of the four accuracy methods underexamined threshold selections demonstrated that TSS is a more realistic and practical method, in comparison with AUC, Sensitivity and Specificity. Accurate projection of the distribution of a species is extremely complex.

Keywords: *AUC, sensitivity, specificity, TSS, bioclimatic model, correlative model.*

GJHSS-B Classification: *FOR Code: 040699*



Strictly as per the compliance and regulations of:



© 2018. Farzin Shabani, Lalit Kumar & Mohsen Ahmadi. This is a research/review paper, distributed under the terms of the Creative Commons Attribution-Noncommercial 3.0 Unported License http://creativecommons.org/licenses/by-nc/3.0/), permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Assessing Accuracy Methods of Species Distribution Models: AUC, Specificity, Sensitivity and the True Skill Statistic

Farzin Shabani [α], Lalit Kumar [σ] & Mohsen Ahmadi [ρ]

Abstract- We aimed to assess different methods for evaluating performance accuracy in species distribution models based on the application of five types of bioclimatic models under three threshold selections to predict the distributions of eight different species in Australia, treated as an independent area. Five discriminatory correlative species distribution models (SDMs), were used to predict the species distributions of eight different plants. A global training data set, excluding the Australian locations, was used for model fitting. Four accuracy measurement methods were compared under three threshold selections of *i)* maximum sensitivity + specificity, *ii)* sensitivity = specificity and *iii)* predicted probability of 0.5 (default). Results showed that the choice of modeling methods had an impact on potential distribution predictions for an independent area. Examination of the four accuracy methods under-examined threshold selections demonstrated that TSS is a more realistic and practical method, in comparison with AUC, Sensitivity and Specificity. Accurate projection of the distribution of a species is extremely complex. As models provided slight variances in projections of the same group of species, it may be more expedient to use TSS as an intuitive method for measuring the performances of the SDMs, in comparison to AUC, Sensitivity, and Specificity.

Keywords: AUC, sensitivity, specificity, TSS, bioclimatic model, correlative model.

I. Introduction

There is evidence of more widespread application of species distribution models (SDMs) to a broader range of practical and hypothetical questions (Guisan and Thuiller, 2005; Jeschke and Strayer, 2008). Also termed habitat or ecological niche models, bioclimatic envelopes and resource selection functions, these are examples of correlative models employing environmental and/or geographical data in order to describe the observed distribution patterns of particular species. This more widespread usage implies that such models are now being used to process alternative data forms, particularly recently having focused on occurrence records of museums and herbaria (Graham et al., 2004). In research into climate change and invasive species, predictions of SDMs may extend beyond the environmental or geographic areas in which the training samples originated (e.g. Araújo et al. (2005)). In the field of epidemiology, for example, SDMs are being used to predict the distributions and occurrences of diseases Peterson et al. (2002). Technological advancement of geographic information systems (Foody, 2008) and progress in data analysis (Breiman, 2001b), has supported the implementation of new modeling methods and applications, which have grown from simple environmental matching techniques, such as in Bioclim(Busby, 1991) and DOMAIN Carpenter et al. (1993), to non-linear relationships of greater complexity between the presence of a species and its environment (e.g., Generalised Additive Models (GAM))Hastie and Tibshirani (1990) and Maximum Entropy Modeling (MaxEnt) (Phillips et al., 2006)). The recent concentration on Bayesian methods and machine learning support the development of further new methods (Latimer et al., 2006; Prasad et al., 2006).

SDM uncertainty can generally be classified into two fundamental categories: model uncertainty and measurement uncertainty (Elith et al., 2002). The former arises from model simplifications, limitations or assumptions in describing processes of extreme complexity,such as future climate projections, or the algorithms of the relationships of species to environment. The latter arises from data imprecision and error, occurring through incorporation of incorrect geographic coordinates of species observations, or climatic datasets created inconsistently from a variety of weather stations, time periods, and interpolated into the mapping process. The origins of uncertainty in SDM predictions have been studied by comparison of the predictions of different types of modeling algorithms, based on a common species, or group thereof, or common environmental predictors (Anderson et al., 2006) or by maintaining a common set of species and algorithms and altering predictor variables (Watling et al., 2012). A few studies have made comparisons combining these multiple factors into a single structure(Buisson et al., 2010; Hanspach et al., 2011). One such example, using four sources of model and measurements of uncertainty regarding the modeling of a single species, ascertained that the algorithm was the main cause of uncertainty, and subsequently occurrence data and co linearity of predictor variables (Dormann et al., 2008).

Author α σ: Ecosystem Management, School of Environmental and Rural Science, University of New England, Armidale, NSW, 2351, Australia. e-mail: fshaban2@une.edu.au
Author ρ: Department of Natural Resources, Isfahan University of Technology, Isfahan, Iran.

© 2018 Global Journals

Assessing predictive accuracy is critical in the development process of distribution models (Barry and Elith, 2006; Guisan and Thuiller, 2005). Quantitative performance assessment for the determination of model suitability to application can be used to uncover aspects requiring improvement (Anderson et al., 2006; Barry and Elith, 2006; Vaughan and Ormerod, 2005), as well as providing the basis for selection of the most appropriate modeling technique for the specific application (Loiselle et al., 2003; Segurado and Araujo, 2004) in that it enables a researcher to investigate the impact of different data and species' properties on the degree of accuracy of the predictive maps generated (Kadmon et al., 2003). In practice, there are two facets in measuring SDM accuracy; discrimination capacity and reliability (i.e. classification accuracy)(Pearce and Ferrier, 2000), with the former generally considered more imposing on outcome than the latter (Ash and Shwartz, 1999). In modeling, discrimination capacity implies the ability to differentiate presence sites (those where the subject species is detected) and absence sites (i.e. pseudo-absence or background sites where it is known or supposed to be absent). Alternatively, reliability implies concord of the predicted occurrence probabilities and proportions of sites observed to be occupied by the species (Pearce and Ferrier, 2000). Reliability is a core facetof quality in probabilistic predictive modeling.

In modeling exercises, the selection of appropriate modeling techniques (e.g., DOMAIN, CLIMEX, MaxEnt, BRT, RF, Bioclim) and methods of measuring accuracy (e.g.,AUC, Sensitivity, Specificity, the True Skill Statistic) are crucial to the outcome. A variety of methods for accuracy measurers are available, each functioning in a slightly different manner. For the layman or novice, the basic decisions at the commencement of the process is which of these is most appropriate to the specific application. Thus, it is necessary to make a comparison of a variety of modeling techniques, associated accuracy measure methods and different species, since techniques perform differently with particular species and the distributions of each.

This study assessed four different methods of measures of accuracy (the area under the ROC curve (AUC), Specificity, Sensitivity and the True Skill Statistic (TSS)) on each of five types of correlative model (General Linear Model (GLM), Max Ent, Bioclim, Random Forest (RF), Boosted Regression Tree (BRT)) under three threshold selections of *i)*maximum sensitivity + specificity, *ii)*sensitivity =specificity and *iii)*probability value of 0.5 (hereafter default) on *Asparagus asparagoides*, *Triticumaestivum* L., *Lantana camara*L., *Opuntiarobusta*,*Triadicasebifera*, *Fusarium oxysporum*f. spp., *Phoenix dactylifera* L. and *Gossypium* (cotton) species distribution records for Australia and the remainder of the world. For this research, we purposefully selected different types of species covering cultivated, fungus, and invasive species and three different thresholds as these give a better basis for validation of the model and thresholds compared to selecting one type of species and threshold. In the primary stage five models were constructed, and thereafter compared using the four measures of accuracy and three different thresholds for each of the five modeling techniques based on projections of suitable climate, derived from observed distribution records of these eight species.

II. Materials and Methods

a) Distribution Records

Distribution data was collected from a variety of sources. Global distribution data was sourced from the Global Biodiversity Information Facility (2015), Atlas of Living Australia (2017), as well as published literature. ENM Tools (Warren et al., 2010) was used in the processing of each grid cell's georeferenced occurrence data to equal 1. Thus, the fact that a single grid cell may display multiple records is of no consequence to the projections or performance evaluation. Distribution records for each of the eight species at Global (GLS) and Australian (AUS) scale numbered as follows: i) *Asparagus asparagoides* GLS: 4924, AUS: 3836, ii) *Phoenix dactylifera* L. GLS: 529, AUS: 51, iii) *Fusarium oxysporum* f. spp GLS: 230, AUS: 30, iv) *Gossypium* GLS: 17322, AUS: 2656, v) *Lantana camara* L. GLS: 17856, AUS: 8324, vi) *Opuntiarobusta* GLS: 299, AUS: 57, vii) *Triadicasebifera* GLS: 1724, AUS: 53 and viii) *Triticumaestivum* L. GLS 50337, AUS: 142. Both native and exotic distribution records were included in the dataset, as it was beyond the parameters of the study scope to distinguish between the inclusion of only native, exotic, or both, in terms of the techniques to project climate suitability and the accuracy methods employed.

b) Species distribution modeling

- Generalized Linear Model (GLM)

The technique of iterative weighted linear regression was employed in GLM to estimate maximum probability of parameters, with a linear expression of the distributions of observations by transformation of the exponential family and systematic effects. For GLM, parametric functions were employed to link the combined linear and quadratic explanatory variables. A standard polynomial approach in combination with an automatic stepwise model selection based on the Akaike Information Criterion (AIC) was used to fit the model. Modeling was done in R v. 3.3.2 (R Development Core Team, 2016).

- MaxEnt

MaxEnt desktop version 3.3.3k (Phillips et al., 2006) was used with modified parameters (Phillips and Dudík, 2008). MaxEnt is dependent on user coordinated

© 2018 Global Journals

geographical background data (Guillera-Arroita et al., 2014) in order to compare the climate factors of the sampled reference set of grid cells with those grid cells where the species is observed to be present. The definition of the background data set significantly affects output (Elith et al., 2011) and the complete range of the species across the searched areas should be included (Elith et al., 2010). Our MaxEnt algorithm compared presence locations and variable interactions to similar interactions of background locations, and established the maximum entropy probability distribution approximating uniformity, subject to the limitations imposed by observed spatial distributions and associated environmental factors. The minimizing of relative entropy between known locations and background point data in such a manner optimizes the maximum entropy probability distribution (Phillips et al., 2006).

- Bioclim

Bioclim (similar to GLM, MaxEnt, BRT and RF) employs the principle that current distribution is the fundamental indicator of the climatic needs of a species, in order to correlate these climate variables with the observed distributions of the species. The model uses the realized niche to describe bioclimatic envelopes, in that non-climatic factors, inclusive of biotic interactions, impose limitations on observed distributions. In contrast, a mechanistic relationship with a more physiological basis is established between the climatic parameters and species response in other types of bioclimatic models (Pearson and Dawson, 2003; Woodward, 1987). Thus, in these models, the fundamental niche is established by modeling the physiological limiting mechanisms in terms of climatic factors. An area of criticism of bioclimatic modeling has been that biotic interactions, species dispersal and evolutionary changes are excluded from the modeling process. These limiting factors and human impacts show that realized niches, as utilized in methodologies of correlative bioclimatic envelopes, are not necessarily the absolute limits of a range and that a future distribution may well be based on alternative factors comprising the realized niche (Pearson and Dawson, 2003). Thus, Bioclim, and its associated environmental envelope models, produce a 'climate profile' of a species, sometimes termed a 'boxcar' descriptor or 'parallelepiped classifier' (Busby, 1991). This basic hyper-box classificatory method thus describes the potential range of a species in terms of a multi-dimensional environmental space whose parameters are the minimum and maximum values for all presences (or 95% of these, or similar variations). In order to extrapolate the prediction within an independent area, we parameterized the model on the outlier-corrected (Skov and Svenning, 2004) observed minimum and maximum values of presence of the species for each variable climatic factor, to provide more conservative results. Bioclimmodel was implemented using the 'Dismo' package (Hijmans and Elith, 2015).

- Random Forest (RF)

The Random Forest is, in performance, one of the most accurate classificatory regression tree-based models. In RF, bootstrap aggregation is used to select many subsamples from the data, generated through a bagging algorithm, a large number of de-correlated regression trees (Breiman, 2001a). RF tree predictors are combined in a manner that each is dependent on the values of independently sampled random vectors, assuming similar distribution for each tree in the forest (Breiman, 2001a). An aggregating (averaging or majority vote) of the predictions of the ensemble forms the basis of the prediction (Svetnik et al., 2003). Out-of-bag observations from each tree are used in predicting model errors and the importance of variables. As in an ensemble approach, decision tree predictions are averaged. We used the 'RandomForest' package (Liaw and Wiener, 2002) to fit the RF models.

- Boosted Regression Tree (BRT)

In our BRT model we used a similar background area to the MaxEnt model, fitting sufficient combinations (decision trees) iteratively, and combining these to produce an optimal model with refined predictive performance. BRT incorporates two multiple regression tree algorithms. Using a binary division into rectangles of the predictor space, it relates the predictor responses to identify areas with the closest responses to predictors and incorporates boosting, an additional procedure, which merges the fitted trees for greater accuracy. For BRT model we employed the 'Dismo' package (Ridgeway, 2006)using an additional setting code recommended by Elith et al. (2008).

c) Bioclim variables, Background data and the methods for providing weights for species records

To remove models' complexity and screening explanatory variables we used the jack-knife analysis method and calculated pairwise Pearson correlation matrix of the variables to select the more important variables with low correlation ($R^2 < 0.5$). For example, the following variables; bio1 (Annual mean temperature (°C)), bio3 (Isothermality), bio8 (Mean temperature of wettest quarter (°C)), bio12 (Annual precipitation (mm)), bio15 (Precipitation seasonality (C of V)), bio17 (Precipitation of driest quarter (mm)), bio20 (Annual mean radiation (W m^{-2})), bio21 (Highest weekly radiation (W m^{-2}), bio24 (Radiation of wettest quarter (W m^{-2})), bio31 (Moisture index seasonality (C of V)), bio34 (Mean moisture index of warmest quarter) and bio35 (Mean moisture index of coldest quarter) were selected for the species *Asparagus asparagoides*. To broaden the background data in terms of the likelihood of fewer record returns from more recent locations of invasion

© 2018 Global Journals

and those poorly sampled, we gave greater importance to records with less geographic proximity. However, it was taken into account that without records on survey effort in terms of time, it is impossible to distinguish between unsuitable and under-sampled areas, and that the above-mentioned adjustments would unavoidably thus confuse these two categories of geographical area. For calculation of the weighting surface, we divided the number of weighted records (using Gaussian kernel method with standard deviations of default values in ArcGIS) in the selected geographical environment for each cell globally, but excluding Australia, by the weighted number of terrestrial cells of the specific area, to eliminate edge effects along coastal regions. Thereafter, the resulting grid was adjusted to maximum 20 and minimum 1, which excluded extreme values. This weighting method, as advocated by Elith et al. (2010), minimizes bias favouring records from densely sampled areas in relation to those from less sampled areas. The kernel density layer of each species and Hawths Tools extension (Beyer, 2004) were used to generate background points for the world, excluding Australia, for training purposes. The same method was used to generate background points for Australia, for comparing model performances. Thus, all SDM performances were evaluated against the same background data for every species.

d) Accuracy Methods

- The area under the ROC curve (AUC)

The receiver operating characteristic (ROC) curve provides an alternative technique for assessment of accuracy of ordinal score models (Fielding and Bell, 1997b). The construction of ROC curves uses all possible thresholds for classifying the scores into confusion matrices, obtaining each matrix' sensitivity and specificity; then comparing sensitivity against the corresponding proportion of false positives (equal to 1 − specificity). Using all thresholds avoids the arbitrary choice of a single threshold (Liu et al., 2005; Manel et al., 2001), and takes into account the trade-off of sensitivity and specificity (Pearce and Ferrier, 2000). The area below the ROC curve (AUC) is also valid as a single threshold-independent measurement of model performance (Brotons et al., 2004; Thuiller et al., 2005). AUC has been demonstrated to be independent of prevalence (McPherson et al., 2004; Somodi et al., 2017) and is seen to be an accurate measure of ordinal score model performance. However, in practice, SDMs used in conservation, such as for selection of representative sites and identification of biodiversity hotspots, frequently needs presence–absence maps of distributions of a species, and requires the selection of a threshold for the transformation of the ordinal scores into presence–absence predictions (Berg et al., 2004). In these circumstances, evaluation accuracy of prediction should be based on the specific threshold selected, as opposed to threshold-independent ROC curves. It is important to note that among the more frequently usedspecies distribution models (e.g. Bioclim, Nix (1986); GARP, Stockwell (1999)) dichotomous presence–absence distribution predictions are generated, to which it is not possible to apply ROC curves.

- Sensitivity and Specificity

Sensitivity represents the proportion of correctly predicted presence records and thus the quantification of omission errors. In calculation, Sensitivity equals $\frac{a}{a+c}$ where *a*denotes the number of correctly predicted presence cells and *c* the number of cells in which the species was found, but absence is predicted by the model. Specificity represents the proportion of correctly predicted absences and thus the quantification of commission errors. In calculation, Specificity equals $\frac{d}{b+d}$ where *b* denotes the number of cells in which the species was not found but presence is predicted by the model, and *d* is the number of cells correctly predicting absence. It is important to note that compared across models, sensitivity and specificity are independent of one another, as well as being independent of prevalence, which represents the proportion of sites where the species was recorded as present.

- True Skill Statistic (TSS)

The TSS is independent of prevalence and equals $\frac{ad-bc}{(a+c)(b+d)}$. Allouche et al. (2006) have shown that TSS is an intuitive method of performance measurement of SDMs in which predictions are expressed as presence–absence maps. It was further shown that TSS gives results showing significant correlation with those of the threshold-independent AUC statistic (Allouche et al., 2006).

e) Thresholds

There are many methods of thresholds selections including taking 0.5 as the threshold (default), which is widely used in ecology (Pearson et al., 2002) or a specific level of sensitivity or specificity (e.g. 95%) is desired or deemed acceptable (Cantor et al., 1999) or thresholds are chosen to maximize the agreement between observed and predicted distributions. A third category of threshold selection identifies a threshold value that maximizes the percent of points correctly classified; maximizes sensitivity plus specificity; or maximizes Kappa, a measure that utilizes both sensitivity and specificity (Guisan et al., 1998). In this study the most commonly used thresholds of *i)*maximum sensitivity + specificity, *ii)*sensitivity = specificity and *iii)*default were examined to evaluate four accuracy methods of the species distribution models.

f) Evaluating accuracy methods

Presence points in this study were divided into two sample categories; training and test points per

© 2018 Global Journals

species. The training dataset comprised presence points of the complete global distribution of the species, excluding the Australian continent, while out-of-sample data (occurrences on the Australian continent) was used as a test of SDM performance. We concentrated on the area below the ROC curve (AUC), Sensitivity, Specificity and True Skill Statistic (TSS) of an independent area under three different thresholds, in order to evaluate accuracy for each species and model separately. Thus, eight species were evaluated using five correlative models. In that there was no data representing true absence of each species in Australia, the proportions of the extent of Australia identified as suitable were calculated, as an index of potential overestimations of the models.

III. Results

Differences in the four methods of accuracy evaluation (AUC, Specificity, Sensitivity and TSS) of Bioclim, BRT, GLM, MaxEnt and RF in the projections of suitable climate under the three different thresholds, based on independent records of all eight species, are shown in Figure 1.

a) AUC

AUC produced similar results in all models. For example, AUC values for all models for *Asparagus asparagoides*, is around 0.94 (Fig 1) even though the output shows a clear difference. Similar comparative results occurred for *Fusarium oxysporum*f. spp.(≈ 0.63), *Gossypium* (≈ 0.70), *Lantana camara* L. ≠ 0.95), *Phoenix dactylifera* L. (≈ 0.55), *Triadicasebifera*(≈ 0.98) and *Triticumaestivum* L. (≈ 0.77) (Fig 1). However, in the case of *Opuntiarobusta*, AUC values of different models had some variation (inconsistency), giving AUC values from Bioclim, BRT, GLM, MaxEnt and RF as 0.51, 0.88, 0.85, 0.90 and 0.50 respectively. Results also show the mean AUC values, using five correlative modeling techniques on eight species, were above 0.77. Consistent with this moderate AUC value, the training dataset model did not predict occurrences of the studied species in certain places where these are known to occur (Fig 1).

b) Specificity

A comparison of specificity in all five models, based on the test data under three different thresholds, shows relatively comparable values for *Asparagus asparagoides*, *Fusarium oxysporum*f. spp., *Gossypium, Lantana camara* L.,*Opuntiarobusta, Phoenix dactylifera* L., *Triadicasebifera* and *Triticumaestivum* L. (Fig 1). For example, specificity values under default threshold for *Triticumaestivum* L. and *Fusarium oxysporum*f. sppfor Bioclim, BRT, GLM, MaxEnt and RF were 1, 0.79, 0.76, 0.87, 0.91 and 1, 0.72, 0.07, 0.00 and 1respectively. Similar comparison on specificity values under "sensitivity = specificity" threshold for *Triticumaestivum* L. and *Fusarium oxysporum*f. sppfor Bioclim, BRT, GLM, MaxEnt and RF were 0.68, 0.68, 0.70, 0.68, 0.74 and 0.67, 0.60, 0.51, 0.59 and 0.98 in turn. Finally, a comparison of specificity values under "maximum sensitivity + specificity" threshold for *Triticumaestivum* L. and *Fusarium oxysporum*f. sppfor Bioclim, BRT, GLM, MaxEnt and RF were 0.63, 0.47, 0.52, 0.73, 0.74 and 0.74, 0.60, 0.88, 0.93 and 0.99 in that order. Results also show that the mean specificity values under different thresholds, using the five modeling techniques on the eight specieswere above 0.78 (Fig. 1).

c) Sensitivity

Sensitivity presented variable results for most models under different examined thresholds. For example, sensitivity values for *Phoenix dactylifera* L. under default threshold were 0.00, 0.38, 0.85, 0.23, and 0.00 for Bioclim, BRT, GLM, MaxEnt and RF, respectively. Sensitivity values for this species under threshold of "sensitivity = specificity" were close to each other while values of sensitivity under threshold of "maximum sensitivity + specificity" were 0.91, 0.17, 0.85, 0.21, and 0.21 for Bioclim, BRT, GLM, MaxEnt and RF, respectively. Similar variations on sensitivity values under default threshold for *Opuntiarobusta* on Bioclim, BRT, GLM, MaxEnt and RF were 0, 0.23, 0.64, 0.19, and 0 respectively. Similar contrast on sensitivity values under "sensitivity = specificity" threshold for this speciesfor Bioclim, BRT, GLM, MaxEnt and RF were 0.02, 0.66, 0.76, 0.80, and 0.00 in turn. Finally, an assessment of sensitivity values under "maximum sensitivity + specificity" threshold for *Opuntiarobusta* for Bioclim, BRT, GLM, MaxEnt and RF were 0.02, 0.66, 0.76, 0.88, 0.00 in that order(Fig. 1).

d) TSS

More realistic value can be seen between the TSS index obtained under different thresholds and/or most of the SDMs output. For example, TSS values for *Triticumaestivum* L.under default threshold were 0.37, 0.36, 0.27, and 0.23 for BRT, GLM, MaxEnt and RFrespectively, which indicates better consistency with areas projected as climatically suitable for the species. TSS values for this speciesunder threshold of "sensitivity = specificity" were 0.37, 0.36, 0.40, 0.25, and 0.28 for Bioclim, BRT, GLM, MaxEnt and RF respectively. Similar consistency for this species were also found under threshold of "maximum sensitivity + specificity" on BRT, GLM, MaxEnt and RF. It should be mentioned that some variation were also seen under different thresholds for this species on Bioclim. Similar consistency was shown for *Fusarium oxysporum*f. spp., *Gossypium, Lantana camara* L.,*Opuntiarobusta, Phoenix dactylifera* L., and *Triadicasebifera* (Fig. 1).

© 2018 Global Journals

IV. Discussion

In this study, the five correlative modeling techniques under three different thresholds were examined through extrapolation (Fig 1). The assessment of SDM correlative and envelope performance, based on AUC, Sensitivity, Specificity and TSS in modeling eight species under threshold selections of *i)* maximum sensitivity + specificity, *ii)* sensitivity = specificity and *iii)* default, indicates that TSS gives varying, but more realisticvalues (Fig 1), in comparison with specificity which represents the probability of correct classification of absence by the model. Caruana and Niculescu-Mizil (2006) note, however, that some researchers have attempted to explain the tests' relative performances and their sensitivity to data characteristics, but movement toward the establishment of a comprehensive assessment toolbox has been hindered by disagreement on the valid applicability of some statistics. SDM evaluation measurements could benefit from the identification of techniques useful in other fields, and from more concentration of research on topics such as the analysis of spatial patterns in errors, dealing with uncertainties, and assessment performance in the context of specific applications, including decision making (Austin, 2007).

We believe that the utilized method to generate absence or background points in the study was appropriate as this method is recommended by Elith et al. (2010) for species which have been presented in different portions of the range for different periods of time. In contrast, the recognized best practice when using museum data is to use what has been termed the 'target group background' approach (Phillips et al., 2009). It should be highlighted that although one of the examined threshold was the default one (0.5)it does not mean that we are suggesting this threshold as the best one.

We believe that use of a combination of distribution modeling techniques such as Bioclim, MaxEnt, BRT, RF and GLM in a complementary method, together with species accuracy estimators, allows us to better represent the geographical distribution of species and the species composition at localities, including a measure of its accuracy. However, it is necessary to assess and evaluate accuracy of species distribution modeling with different techniques as there are biases and limitations in representation of the results purely based on one modeling technique or one accuracy method. Using a combination of methodological approaches as executed in this study facilitates identification of an overall pattern, provided by all of the individual model predictions, that represent the geographical patterns of richness and composition of species, regardless of the degree of accuracy of the predictions by each individual model for each species.

Accurate projection of a dynamic phenomenon such as the richness of the distribution of a species is extremely complex. It has been shown that the results of SDMs are unreliable projections of the range of a species. Rather, they produce a provisional description of ranges, which require continuous updating as new data becomes available or environmental factors alter. Species distributions predicted by the relating of biological data to environmental variables showed a tendency toward overestimation of the actual range extents, due in part to the limitations of using only the environmental conditions as model predictors for the sites where the species has a known presence. Where absences due to historical, dispersal or biotic factors (Pulliam, 2000) are not accounted for, model predictions willinevitably tend toward the potential distribution of species (i.e. sites of environmental suitability in which a species could occur, based on a group of environmental variables; see (Jiménez-Valverde et al., 2008)). Under such circumstance, a set of errors and biases will result when predictive distribution maps are overlaid to create a representation of the richness of a species, producing an unrealistic representation (Hortal et al., 2007). Thus, the creation of a valid representation of species richness demands a deeper analysis of results, in order to detect areas with notable levels of omission, as well as account for presences located in areas where no representation was predicted.

Why not AUC? SDMs are invaluable for addressing questions and issues in biogeography, as well as evolutionary and conservation biology. Understanding performance, assessment of correlative and mechanistic models is essential to their valid application (Guisan and Thuiller, 2005). AUC is a frequently used technique for measurement of model performance (Lobo et al., 2008; Manel et al., 2001; Thuiller et al., 2005), proven to be independent of prevalence, in theoretical (Hanley and McNeil, 1982; Zweig and Campbell, 1993) and empirical applications (McPherson et al., 2004). In performance measurement, AUC is threshold independent and thus suitable for evaluating performance in ordinal score models, like logistic regression with true presence-absence data. However, in practice, absence data is often unavailable and only the presence data is accessible. Under such circumstances, envelope (eg. Bioclim) or distance-based models (e.g. Domain or Mahalanobis) are the SDMs of choice (Farber and Kadmon, 2003). However, in practice, a comparative prediction of presence-absence is often necessary, thus necessitating a threshold application for transforming the probability/ suitability scores into presence–absence data. For most reverse selection algorithms, presence–absence data of composition of species in specific locations is necessary (Tsuji and Tsubaki, 2004). As available data is frequently not complete, SDMs are often used to predict presence or absence in a potential locality for a

© 2018 Global Journals

particular species (Sánchez-Cordero et al., 2005). Biodiversity hotspot estimations are also frequently based on presence–absence predictions (Schmidt et al., 2005). Assessing impacts at community level of global change could be achieved by stacked binary SDM species assemblage prediction (D'Amen et al., 2015; Guisan and Rahbek, 2011). Presence–absence predictions exclude ROC plotting and, thus, AUC is not a technique for evaluating accuracy of the predictive maps used in such applications. The results in Figure 1 indicate that the high values of AUC for each species and model is no guarantee of output accuracy. Further, MESS (Multivariate Environmental Similarity Surface) maps do not specify changes in correlations between variables, and tests for these are also essential because parameters are estimated on the structure of correlations between training data predictors. Generally in SDMs, predictions will be unreliable for areas with substantial variance in correlations of important variables (Harrell, 2001). When available predictors have only indirect relationships to distributions of species, this is particularly problematic (Austin, 2002). While the selected set of variables might reasonably well represent the unmeasured directly influential variable, if inherent correlations change in new areas, there will be compromises in predictions.

Regarding the necessity of producing presence/absence predictions from SDMs, evaluating this binary prediction using confusion matrix and classification accuracy criteria should be taken into account. However, the selection of an optimal threshold is a critical issue, raisinga literary criticism(Liu et al., 2005). How well a binary prediction can classify presence and absence observations, which is called as sensitivity and specificity, respectively, is the cornerstone of the classification accuracy evaluation. Although,these metrics have been solely used for evaluating binary predictions(Ahmadi et al., 2013), they show an inherent inconsistency. For examples models with ahigh value of sensitivity donot necessarily show high specificity. It seems that models capability for extrapolation and/or interpolation compromise the resulting values of sensitivity and specificity(Franklin, 2010; Merow et al., 2014). This can be seen in our case where for almost all species RF results in the lowermost probability of occurrence in the independent area, and accordingly, high values of specificity but low values of sensitivity. Furthermore, the niche shift, the tendency of the species to establish in areas beyond the native niche in out-of-sample areas (e.g. independent area), also affects the prediction performance of the SDMs [34].In this situationTSS (i.e. sensitivity + specificity – 1) through combining the capability of correctly predicting both presence and absence (e.g. background points) observations, and therefore, taking into account both omission and commission errors, provides a reasonable viewpoint of the models performance.

Comparison of the initial distributions of species richness from model predictions with the observed ones and the analysis of errors are the successive phases for adjustment of predicted distributions of a species subset, thereby refining the picture of species richness. Reductions in the errors of omission or commission can be executedby prioritizing either sensitivity or specificity (Fielding and Bell, 1997a). The accuracy of a model must be always interpreted in terms of its intended purpose (Araujo and Guisan, 2006) by differential weighting of false-positives and false-negatives. In our study, the impact of omitting observed species was assumed to be greater, and we thereforeminimized errors of omission. Both commission and omission errors need consideration, however, from the perspective of conservation, ignoring a species where it is present may lead to the underestimation or minimization of the conservation needs of an area, while erroneously including a species in a particular locality might result in unnecessary or wasted conservation efforts and resources(Rondinini et al., 2006). A specific strategy is demanded, based on the need to reduce commission or omission errors.

Choosing a threshold is required when assessing model performance using the indices derived from the confusion matrix, which also facilitates the interpretation of modeling outputs, and in line with this matter we refer to Liu et al.(Liu et al., 2005) who reviewed different threshold determination approaches. Furthermore, refer to Bean et al.(Bean et al., 2012) who investigated the effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. In line with their finding, and based on the results of this study, selecting an arbitrary default threshold (for example predicted probability of 0.5) may underestimate the performance of the model to classify presence/absence areas. In such situations, taking into account the behaviour of the model to characterize presence and absence points, for example where sensitivity of the model equals to specificity or their summation reaches maximum, is more reasonable for selecting thresholds and producing binary presence/absence maps.

In this study attempts were made to answer the question "*in the use of species distribution models, should we rely on the result of a single accuracy method or a single species distribution method?"* through evaluating AUC, Sensitivity, Specificity and TSS performance accuracy methods based on the application of five types of bioclimatic models under three different thresholds to predict the distributions of eight different species in an independent area. As discussed earlier, SDMs are based on different algorithms and thus they perform differently; and for the users, the decisions at the commencement of the process is which of these is most appropriate is complicated; and the situation would become more

© 2018 Global Journals

challenging if the users rely on in appropriate accuracy measure methods. Our findings show that evaluating performance of accuracy gives different results among different techniques and the TSS method is better compared to the other three examined methods. We note that this study adds to one undertaken by Allouche et al. (2006) who assessed the accuracy of species distribution models through prevalence, kappa and TSS.

V. Conclusion

The extensive array of methods, data types and novel research questions imply the need for many modeling decisions. Different modeling techniques (e.g., DOMAIN, CLIMEX, MaxEnt, BRT, RF, Bioclim) and different methods of measuring accuracy (e.g., AUC, Sensitivity, Specificity, the True Skill Statistic)have different requirements. In selecting the most appropriate method of measuring accuracy, knowledge is required in terms of which method is most appropriate for the data available and its intended application. However, the information facilitating an informed choice of method is currently scattered throughout the modeling literature and incomplete, making it problematic for most users to make decisions on the adoption of newer methods, and for newcomers to know where to begin. Knowledge of a particular algorithm gives insight into the features and limitations of its predictions, and why particular patterns occur. As Bioclim, GLM, MaxEnt, BRT and RF provided slight variances in projections of the same group of species, it may be more expedient to use TSS as an intuitive method for measuring the performances of species distribution models, in comparison with the area under the ROC curve (AUC), Sensitivity and Specificity.

References Références Referencias

1. Ahmadi, M., Kaboli, M., Nourani, E., Shabani, A., Ashrafi, S., 2013. A predictive spatial model for gray wolf (Canis lupus) denning sites in a human-dominated landscape in western Iran. Ecological research 28, 513-521.
2. Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). Journal of applied ecology 43, 1223-1232.
3. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J Hijmans, R., Huettmann, F., R Leathwick, J., Lehmann, A., Li, J., G Lohmann, L., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129-151.
4. Araujo, M., Guisan, A., 2006. Five (or so) challenges for species distribution modeling. Journal of biogeography 33, 1677-1688.
5. Araújo, M., Whittaker, R., Ladle, R., Erhard, M., 2005. Reducing uncertainty in projections of extinction risk from climate change. Global Ecology and Biogeography 14, 529-538.
6. Ash, A., Shwartz, M., 1999. R2: a useful measure of model performance when predicting a dichotomous outcome. Statistics in medicine 18, 375-384.
7. Atlas of Living Australia, 2017. Atlas of Living Australia, Available at: :http://www.ala.org.au/,. Accessed: July 2017.
8. Austin, M., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modeling. Ecological modeling 157, 101-118.
9. Austin, M., 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecological modeling 200, 1-19.
10. Barry, S., Elith, J., 2006. Error and uncertainty in habitat models. Journal of Applied Ecology 43, 413-423.
11. Bean, W., Stafford, R., Brashares, J., 2012. The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. Ecography 35, 250-258.
12. Berg, Å., Gärdenfors, U., Von Proschwitz, T., 2004. Logistic regression models for predicting occurrence of terrestrial molluscs in southern Sweden–importance of environmental data quality and model complexity. Ecography 27, 83-93.
13. Beyer, H., 2004. Hawth's analysis tools for ArcGIS.
14. Breiman, L., 2001a. Random forests. Machine learning 45, 5-32.
15. Breiman, L., 2001b. Statistical modeling: The two cultures (with comments and a rejoinder by the author). Statistical Science 16, 199-231.
16. Brotons, L., Thuiller, W., Araújo, M., Hirzel, A., 2004. Presence-absence versus presence-only modeling methods for predicting bird habitat suitability. Ecography 27, 437-448.
17. Buisson, L., Thuiller, W., Casajus, N., Lek, S., Grenouillet, G., 2010. Uncertainty in ensemble forecasting of species distribution. Global Change Biology 16, 1145-1157.
18. Busby, J., 1991. BIOCLIM-a bioclimate analysis and prediction system. Plant Protection Quarterly (Australia).
19. Cantor, S., Sun, C., Tortolero-Luna, G., Richards-Kortum, R., Follen, M., 1999. A comparison of C/B ratios from studies using receiver operating characteristic curve analysis. Journal of clinical epidemiology 52, 885-892.
20. Carpenter, G., Gillison, A., Winter, J., 1993. DOMAIN: a flexible modeling procedure for mapping potential distributions of plants and animals. Biodiversity & Conservation 2, 667-680.
21. Caruana, R., Niculescu-Mizil, A., 2006. An empirical comparison of supervised learning algorithms,

© 2018 Global Journals

Proceedings of the 23rd international conference on Machine learning. ACM, pp. 161-168.
22. D'Amen, M., Dubuis, A., Fernandes, R., Pottier, J., Pellissier, L., Guisan, A., 2015. Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models. Journal of Biogeography.
23. Dormann, C., Purschke, O., Márquez, J., Lautenbach, S., Schröder, B., 2008. Components of uncertainty in species distribution analysis: a case study of the great grey shrike. Ecology 89, 3371-3386.
24. Elith, J., Burgman, M., Regan, H., 2002. Mapping epistemic uncertainties and vague concepts in predictions of species distribution. Ecological modeling 157, 313-329.
25. Elith, J., Kearney, M., Phillips, S., 2010. The art of modeling range-shifting species. Methods in ecology and evolution 1, 330-342.
26. Elith, J., Leathwick, J., Hastie, T., 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77, 802-813.
27. Elith, J., Phillips, S., Hastie, T., Dudík, M., Chee, Y., Yates, C., 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17, 43-57.
28. Farber, O., Kadmon, R., 2003. Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. Ecological Modeling 160, 115-130.
29. Fielding, A., Bell, J., 1997a. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental conservation 24, 38-49.
30. Fielding, A.H., Bell, J., 1997b. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental conservation 24, 38-49.
31. Foody, G., 2008. GIS: biodiversity applications. Progress in Physical Geography 32, 223.
32. Franklin, J., 2010. Mapping species distributions: spatial inference and prediction. Cambridge University Press.
33. Global Biodiversity Information Facility, 2015. Global Biodiversity Information Facility (GBIF), Available at: http://www.gbif.org, Accessed: July 2015.
34. Graham, C., Ferrier, S., Huettman, F., Moritz, C., Peterson, A., 2004. New developments in museum-based informatics and applications in biodiversity analysis. Trends in ecology & evolution 19, 497-503.
35. Guillera-Arroita, G., Lahoz-Monfort, J., Elith, J., 2014. Maxent is not a presence–absence method: a comment on Thibaud et al. Methods in Ecology and Evolution 5, 1192-1197.
36. 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.
37. Guisan, A., Theurillat, J., Kienast, F., 1998. Predicting the potential distribution of plant species in an alpine environment. Journal of Vegetation Science 9, 65-74.
38. Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecology letters 8, 993-1009.
39. Hanley, J., McNeil, B., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143, 29-36.
40. Hanspach, J., Kühn, I., Schweiger, O., Pompe, S., Klotz, S., 2011. Geographical patterns in prediction errors of species distribution models. Global Ecology and Biogeography 20, 779-788.
41. Harrell, F., 2001. Regression modeling strategies: with applications to linear models, logistic regression and survival analysis.
42. Hastie, T., Tibshirani, R., 1990. Generalized additive models. CRC Press.
43. Hijmans, R.J., Elith, J., 2015. Species distribution modeling with R. Citeseer.
44. Hortal, J., Lobo, J., Jiménez-valverde, a., 2007. Limitations of Biodiversity Databases: Case Study on Seed-Plant Diversity in Tenerife, Canary Islands. Conservation Biology 21, 853-863.
45. Jeschke, J., Strayer, D., 2008. Usefulness of bioclimatic models for studying climate change and invasive species. Annals of the New York Academy of Sciences 1134, 1-24.
46. Jiménez-Valverde, A., Lobo, J., Hortal, J., 2008. Not as good as they seem: the importance of concepts in species distribution modeling. Diversity and distributions 14, 885-890.
47. Kadmon, R., Farber, O., Danin, A., 2003. A systematic analysis of factors affecting the performance of climatic envelope models. Ecological Applications 13, 853-867.
48. Latimer, A., Wu, S., Gelfand, A., Silander, J., 2006. Building statistical models to analyze species distributions. Ecological applications 16, 33-50.
49. Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R news 2, 18-22.
50. Liu, C., Berry, P., Dawson, T., Pearson, R., 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28, 385-393.
51. Lobo, J., Jiménez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. Global ecology and Biogeography 17, 145-151.
52. Loiselle, B., Howell, C., Graham, C., Goerck, J., Brooks, T., Smith, K., Williams, P., 2003. Avoiding pitfalls of using species distribution models in

© 2018 Global Journals

conservation planning. Conservation Biology 17, 1591-1600.
53. Manel, S., Williams, H., Ormerod, S., 2001. Evaluating presence–absence models in ecology: the need to account for prevalence. Journal of applied Ecology 38, 921-931.
54. McPherson, J., Jetz, W., Rogers, D., 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? Journal of applied ecology 41, 811-823.
55. Merow, C., Smith, M., Edwards, T., Guisan, A., McMahon, S., Normand, S., Thuiller, W., Wüest, R., Zimmermann, N., Elith, J., 2014. What do we gain from simplicity versus complexity in species distribution models? Ecography 37, 1267-1281.
56. Nix, H., 1986. A biogeographic analysis of Australian elapid snakes. Atlas of elapid snakes of Australia 7, 4-15.
57. Pearce, J., Ferrier, S., 2000. Evaluating the predictive performance of habitat models developed using logistic regression. Ecological modeling 133, 225-245.
58. Pearson, R., Dawson, T., 2003. Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? Global Ecology and Biogeography 12, 361-371.
59. Pearson, R., Dawson, T., Berry, P., Harrison, P., 2002. SPECIES: a spatial evaluation of climate impact on the envelope of species. Ecological modeling 154, 289-300.
60. Peterson, A., Sánchez-Cordero, V., Beard, C., Ramsey, J., 2002. Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. Emerging infectious diseases 8, 662-667.
61. Phillips, S., Anderson, R., Schapire, R., 2006. Maximum entropy modeling of species geographic distributions. Ecological Modeling 190, 231-259.
62. Phillips, S., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31, 161-175.
63. Phillips, S., Dudík, M., Elith, J., Graham, C., Lehmann, A., Leathwick, J., Ferrier, S., 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications 19, 181-197.
64. Prasad, A., Iverson, L., Liaw, A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9, 181-199.
65. Pulliam, H., 2000. On the relationship between niche and distribution. Ecology letters 3, 349-361.
66. R Development Core Team, 2016. R Development Core Team. R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna.
67. Ridgeway, G., 2006. gbm: Generalized boosted regression models. R package version 1.
68. Rondinini, C., Wilson, K., Boitani, L., Grantham, H., Possingham, H., 2006. Tradeoffs of different types of species occurrence data for use in systematic conservation planning. Ecology letters 9, 1136-1145.
69. Sánchez-Cordero, V., Cirelli, V., Munguial, M., Sarkar, S., 2005. Place prioritization for biodiversity content using species ecological niche modeling. Biodiversity Informatics 2.
70. Schmidt, M., Kreft, H., Thiombiano, A., Zizka, G., 2005. Herbarium collections and field data-based plant diversity maps for Burkina Faso. Diversity and Distributions 11, 509-516.
71. Segurado, P., Araujo, M.B., 2004. An evaluation of methods for modeling species distributions. Journal of Biogeography 31, 1555-1568.
72. Skov, F., Svenning, J., 2004. Potential impact of climatic change on the distribution of forest herbs in Europe. Ecography 27, 366-380.
73. Somodi, I., Lepesi, N., Botta-Dukát, Z., 2017. Prevalence dependence in model goodness measures with special emphasis on true skill statistics. Ecology and evolution 7, 863-872.
74. Stockwell, D., 1999. The GARP modeling system: problems and solutions to automated spatial prediction. International journal of geographical information science 13, 143-158.
75. Svetnik, V., Liaw, A., Tong, C., Culberson, J., Sheridan, R., Feuston, B., 2003. Random forest: a classification and regression tool for compound classification and QSAR modeling. Journal of chemical information and computer sciences 43, 1947-1958.
76. Thuiller, W., Lavorel, S., Araújo, M., 2005. Niche properties and geographical extent as predictors of species sensitivity to climate change. Global Ecology and Biogeography 14, 347-357.
77. Tsuji, N., Tsubaki, Y., 2004. Three new algorithms to calculate the irreplaceability index for presence/absence data. Biological Conservation 119, 487-494.
78. Vaughan, I., Ormerod, S., 2005. The continuing challenges of testing species distribution models. Journal of Applied Ecology 42, 720-730.
79. Warren, D., Glor, R., Turelli, M., 2010. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33, 607-611.
80. Watling, J., Romanach, S., Bucklin, D., Speroterra, C., Brandt, L., Pearlstine, L., Mazzotti, F., 2012. Do bioclimate variables improve performance of climate envelope models? Ecological Modeling 246, 79-85.
81. Woodward, F., 1987. Climate and plant distribution. Cambridge University Press.

© 2018 Global Journals

82. Zweig, M., Campbell, G., 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical chemistry 39, 561-577.

Author contribution statement:
Conceived and designed the experiments: FS LK MA.
Performed the experiments: FS, MA.
Analysed the data: FS, MA.
Contributed reagents/materials/analysis tools: FS, MA.
Wrote the paper: FS, LK, MA.

Additional Information:

The authors have declared that no competing financial interests exist.

© 2018 Global Journals

Figure Caption

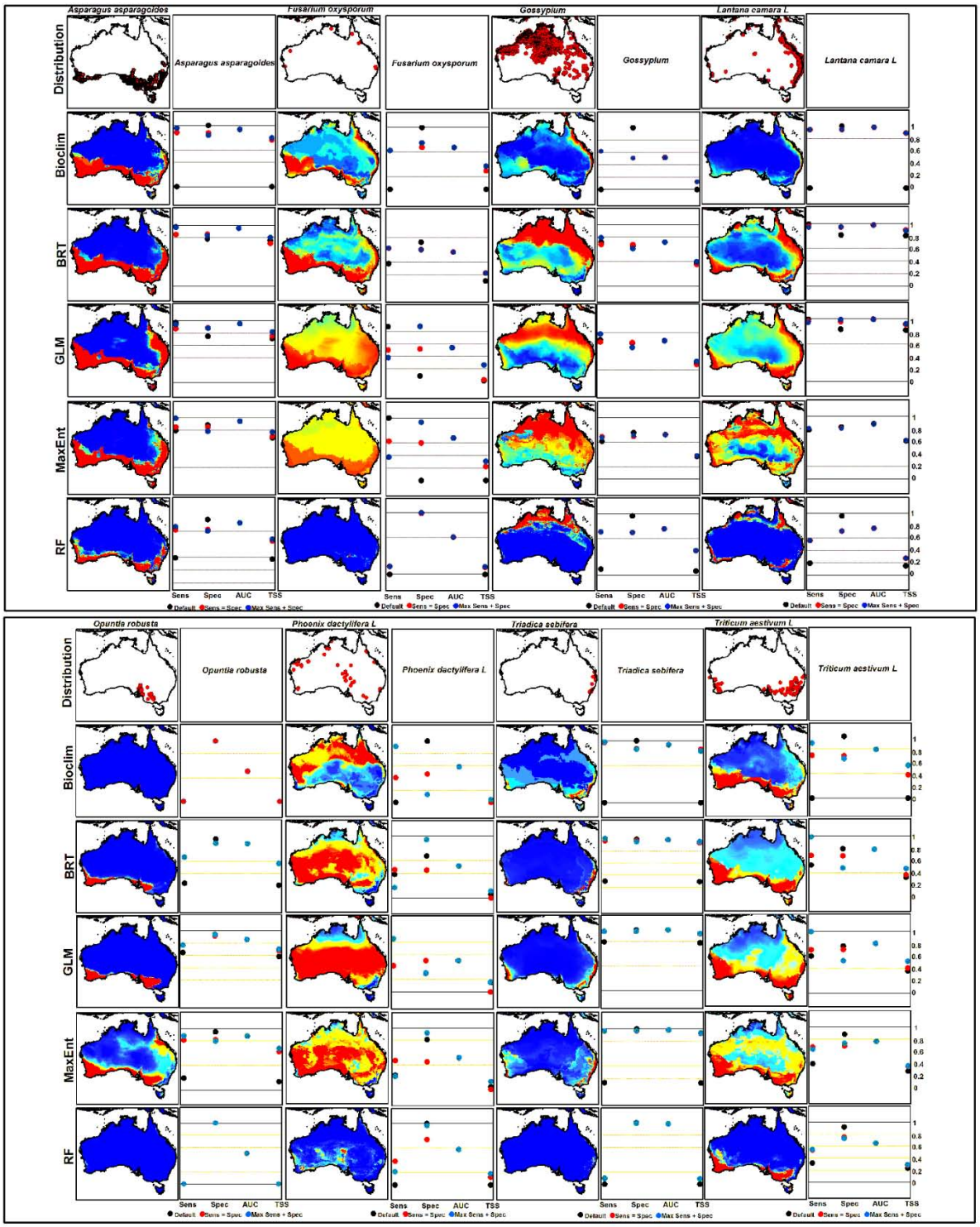


Figure 1: Projections of an independent area for the potential distribution of eight species using five different correlative niche models and the area under the ROC curve (AUC), Sensitivity, Specificity and the True Skill Statistic (TSS) values. Warmer colors show areas with better-predicted conditions. The top row of maps, shows the distribution of species. This map was generated by ArcMap 10.2. Available at http://www.esri.com/arcgis/about-arcgis. The dot graphs show the performance of the different accuracy methods tested by different threshold selection.

© 2018 Global Journals