

Transferability of species distribution models: The case of *Phytophthora cinnamomi* in Southwest Spain and Southwest Australia



J. Duque-Lazo^{a,b,c,*}, H. van Gils^b, T.A. Groen^b, R.M. Navarro-Cerrillo^a

^a Department of Forestry, School of Agriculture and Forestry, University of Córdoba, Laboratory of Dendrochronology, DendrodatLab-ERSAF Edif. Leonardo da Vinci, Campus de Rabanales s/n, 14071 Córdoba, Spain

^b Department of Natural Resources, Faculty of Geo-information Science and Earth Observation (ITC), University of Twente, Hengelosestraat 99, 7500 AE Enschede, The Netherlands

^c Department of Systematic Botany and Functional Biodiversity, Institute of Biology, University of Leipzig, Johannisallee 21–23, 04103 Leipzig, Germany

ARTICLE INFO

Article history:

Received 27 January 2015

Received in revised form

17 September 2015

Accepted 22 September 2015

Keywords:

General Linear Models (GLM)

Mahalanobis distance

Maximum entropy models (MaxEnt)

Phytophthora cinnamomi

Species distribution models

Model transferability

ABSTRACT

Species distribution models (SDMs) predict the probability of species occurrence based on their statistical relationships with environmental variables. Many studies have compared SDM algorithms and accuracies, but few have compared transferability of SDMs across disjunctive geographic areas and the impact of the number of variables on their performance and transferability. We compared the predictive power and transferability of ten SDM algorithms using *Phytophthora cinnamomi* point locations. *P. cinnamomi* Rands is an oomycete associated with Oak Decline in Spain and Jarrah Dieback in Australia. The environmental similarity between the two areas was assessed using Mahalanobis distances. Model performance was evaluated by area under the curve (AUC). The spatial distribution of *P. cinnamomi* was predicted satisfactorily (AUC > 0.7) in Spain by nine algorithms, but only marginally (AUC > 0.55) by all except one in Australia. Further, three machine-learning methods showed relatively high predictive power in the calibration areas combined with low transferability to the other continent, while two of the five regression models achieved reasonable accuracies within the calibration areas combined with a better intercontinental transferability compared to machine-learning methods. The SDM algorithms generated diverse responses to using less than five environmental variables. The most desirable combinations of model performance and transferability were achieved by General Linear Models (GLM) and Generalized Additive Models (GAM). MaxEnt performed at par with GLM and GAM in-situ as well as in transferring for one of the two directions. Moreover, MaxEnt provided the most robust response across the number of environmental variables tested.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Species Distribution Models (SDMs) predict species occurrences in space and time (Elith and Leathwick, 2009; Franklin, 2010) based on statistical relationships between presence-only or presence-absence point records of species and environmental variables. These relations are projected across a specific geographic region using a gridded environmental layer. Therefore, species distributions may be assessed in three ways: interpolation for contiguous/adjacent areas without species occurrence (Wenger and Olden, 2012), extrapolation for a geographic range wider than the

calibration area of the model (Thuiller et al., 2011), or transferring a model calibrated in one region or time period to a disjunctive region or to a different period (Elith and Leathwick, 2009). Transferability of SDMs is helpful in tasks such as assessment of the impacts of climate change on biodiversity (Thuiller et al., 2011), prediction of the distribution of threatened species (Navarro-Cerrillo et al., 2011; van Gils et al., 2012), forecasting of forest dynamics (van Gils et al., 2008), and risk assessment of invasion by alien species (Václavík et al., 2010). The latter includes *Phytophthora* SDMs in California, USA (Václavík and Meentemeyer, 2009), South West Australia (Shearer et al., 2007) and South West Spain (Consejería de Medio Ambiente, 2010). However, transferability performance of SDMs between two disjunctive areas is poorly understood preventing an evidence-based selection of SDM algorithms and environmental variables to obtain the most reliable predictions by transferring a model.

Developing models that consider a sufficient number of environmental variables without becoming over-fitted (Thuiller et al.,

* Corresponding author at: Department of Forestry, School of Agriculture and Forestry, University of Córdoba, Laboratory of Dendrochronology., DendrodatLab-ERSAF Edif. Leonardo da Vinci, Campus de Rabanales s/n, 14071 Córdoba, Spain. Tel.: +34957218786.

E-mail addresses: jduque@uco.es, jduquelazo@gmail.com (J. Duque-Lazo).

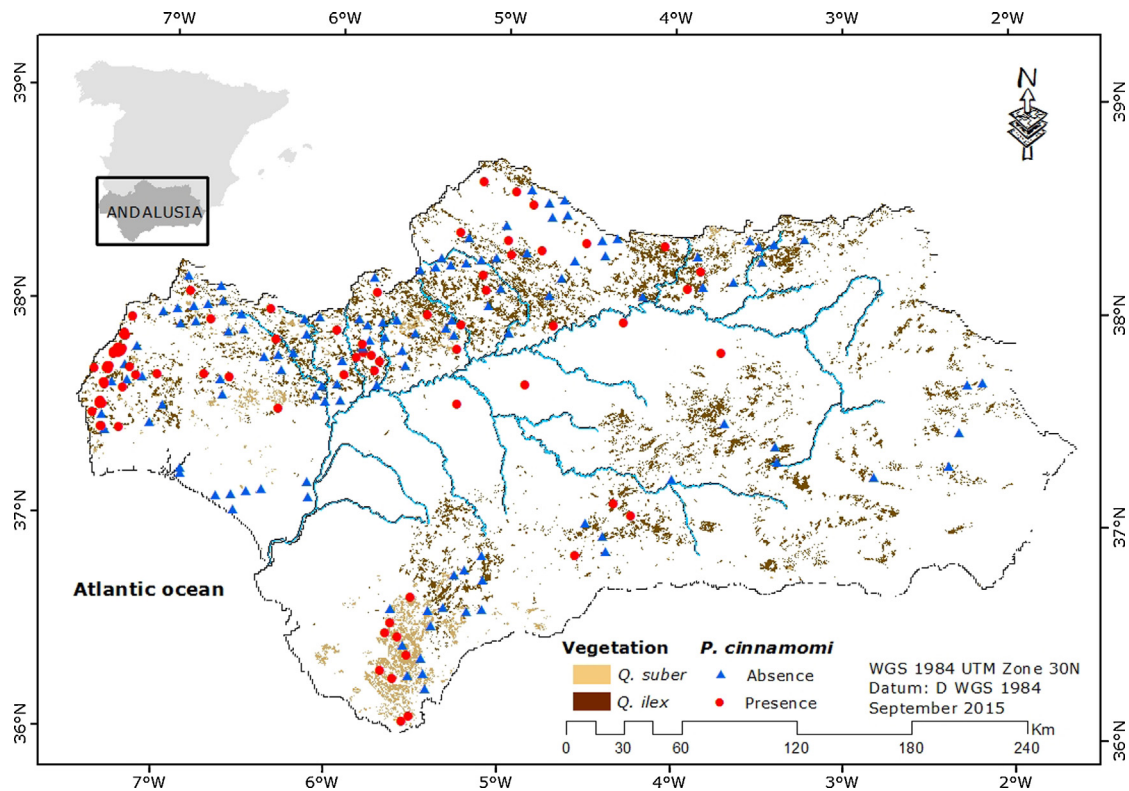


Fig. 1. *P. cinnamomi*, holm (*Quercus ilex*) and cork oak (*Q. suber*) point locations in Spain (Andalusia; oak distribution source: MAGRAMA, 2007).

2008) is a challenge for achieving transferability. Over-fitting may lead to loss of generality of the model, and thus compromise correct inferences in a disjunctive region. A large number of environmental variables (e.g. Elith and Leathwick, 2009; Franklin, 2010; Merow et al., 2014) may generate accurate estimations for the calibration area, but also lead to over-fitting (Anderson and Gonzalez, 2011). However, large (5–10) numbers of environmental variables can be reduced sometimes without substantial loss of SDM accuracy; this applies across regression-based and machine-learning algorithms (Merow et al., 2014; van Gils et al., 2012, 2014).

Comparatively few studies on SDM have tested their transferability. Previous studies disagree about the supremacy of regression over machine-learning algorithms for transferability (Heikkinen et al., 2012; Wenger and Olden, 2012). Also, several studies that test model transferability (Heikkinen et al., 2012; Wenger and Olden, 2012) do not consider the influence of the number or type of environmental variables used, even though it has been demonstrated that the number of variables (Moreno-Amat et al., 2015) and the type of variable (Verbruggen et al., 2013) used can impact on model transferability.

This research explores transferability of *Phytophthora cinnamomi* SDMs between Spain and Australia the questions raised in this study are: (a) Are models that show high accuracies locally also better when transferred to a disjunctive area? (b) Which SDM algorithm achieves the best transferability accuracy? And (c) Do model transferability accuracies depend on the number of variables included in the models?

2. Method

2.1. Study area

South West Spain (Andalusia, Spain, 87.3 km²) is located at the oceanic edge of the Eurasian continent and, as a consequence, its climate presents typical Mediterranean features. In South West

Australia (Australia), we delineated a coastal area of equivalent size within the Mediterranean climatic belt (Peel et al., 2007). Both areas show a high botanical biodiversity (Myers et al., 2000) (Figs. 1 and 2).

The study area within Spain is characterized by a central valley surrounded by mountains with an elevation increasing rapidly from sea level up to 3800 m a.s.l. The study area in Australia is less mountainous with an elevation up to 700 m a.s.l. further inland.

2.2. The species

Phytophthora cinnamomi Rands is a soil-borne oomycete with a worldwide distribution including the USA, Australia, South Africa and Europe (Zentmyer, 1988). *P. cinnamomi* causes root rot, dieback and cankers in >3000 woody plants species, including eucalyptus and oak (Hardham, 2005). Soil moisture and high temperature enhance its activity (Weste and Marks, 1987). *P. cinnamomi* is on the list of the 100 most invasive species worldwide (Lowe et al., 2000) and its effects have been reported in Spain (Sánchez et al., 2003) and Australia (Shearer et al., 2007). This oomycete is considered a major factor for Oak Decline in Spain (Brasier, 1996) and Jarrah Dieback in Australia (Weste and Marks, 1987).

The study areas in both Spain and Australia have a Mediterranean climate and similar environmental conditions (Peel et al., 2007) suggesting that model transferability could be high (Elith and Leathwick, 2009). The similarity in the value range of each environmental variable was assessed between the study areas in Spain and Australia areas using Mahalanobis distances (Quinn and Keough, 2002).

2.3. Presence/absence data

We used *P. cinnamomi* point locations from the study areas in Spain and Australia to test SDM transferability. For the Spain area a total of 215 presence/absence points (presence: $n=95$; absence:

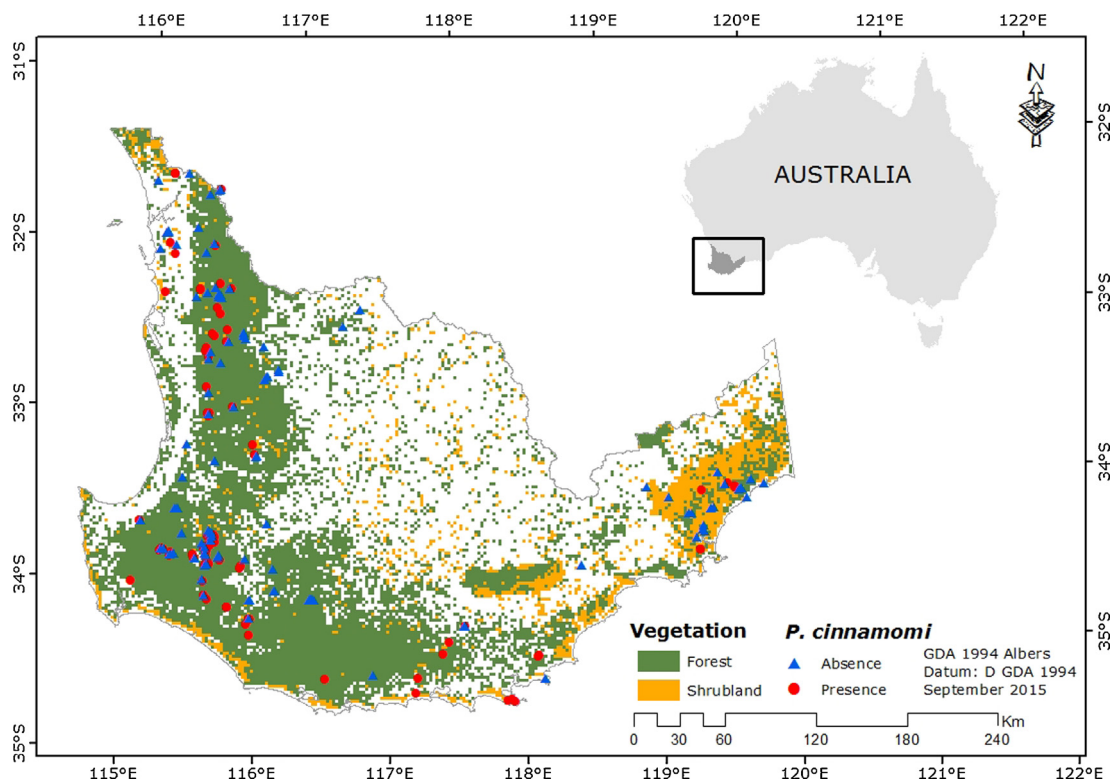


Fig. 2. *P. cinnamomi* point locations and woody vegetation cover of the SW Australia study area; vegetation cover source: ABARES, 2011.

$n = 120$; Fig. 1) of *P. cinnamomi* were evaluated. Most of the Spain points ($n = 167$; presence: $n = 47$; absence: $n = 120$) were obtained from the Andalusia Forest Monitoring Network and the Forest Phytosanitary Alert Network (Consejería de Medio Ambiente, 2010). Soil samples for *P. cinnamomi* presence/absence determination were extracted from forest surveys (Fig. 1). The soil was sampled at the vertices of 8×8 km grids, laid out by the “International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests” (CEE-ICP Forest). The vertices were located on the ground using a hand-held GPS with a spatial accuracy of 5–10 m. The samples were analyzed at the Agroforestry Pathology Department (University of Córdoba). We added 48 presence points from which 42 points were obtained from field surveys and 6 points from the literature (Gomez-Aparicio et al., 2012).

The Australian *P. cinnamomi* dataset included 1,625 points (presence: $n = 552$; absence: $n = 1073$) recorded by the Vegetation Health Service laboratory between the 1st of July 2011 and the 30th of June 2012 (Stukely et al., 2012). Although, reducing the number of occurrence points might influence model accuracy (Hanberry et al., 2012), the dataset was randomly sub-sampled in R (R Core Development Team, 2014) to a comparable number of presences and absences as contained in the Spain dataset (Fig. 2).

2.4. Spatial autocorrelation

Spatial autocorrelation is a common issue in environmental datasets (Legendre and Fortin, 1989) and occurs when the values of variables sampled at close proximity are not independent from each other (Tobler, 1970). Spatial autocorrelation in occurrence data may cause over-predictions of SDMs. Consequently, models that do not account for spatial autocorrelation tend to present higher accuracy estimates (e.g. AUC) than those that consider spatial autocorrelation (Veloz, 2009). To assess whether this could apply to our dataset, we calculated the range of the semi-variogram per environmental variable. This range was compared with the

average distance between the point occurrences of *P. cinnamomi* in each study site and the pixel resolution (Yu et al., 2015).

2.5. Environmental data

For the study areas in Spain and Australia 29 meteorological variables were available as gridded layers, at 1 and 2 km² resolution, respectively (Supplementary material Appendix 1, Tables A1 and A2). The set of meteorological variables was complemented with four additional meteorological (10 km²), NDVI (8 km²), topographic (DEM and DEM-derived), soil, substrate and land cover (Supplementary material Appendix 1, Tables A1 and A2). The data layers were spatially trimmed to the study extent and re-sampled by nearest neighbourhood assignment to 2 km² grids for both study areas; the meteorological variables were standardized over the period from 1976 to 2006 (Supplementary material Appendix 1, Table A1).

The meteorological variables were selected based on known seasonal requirements of *P. cinnamomi*, which include seasons with high temperatures and soil moisture availability (Hardham, 2005; Weste and Marks, 1987). Therefore, seasonal (winter: Dec, Jan, Feb; spring: Mar, Apr, May; summer: Jun, Jul, Aug, autumn: Sep, Oct, Nov; for Spain, and winter: Jun, Jul, Aug; spring: Sep, Oct, Nov; summer: Dec, Jan, Feb; autumn: Mar, Apr, May; for Australia) and annual averages (Supplementary material Appendix 1, Table A2) were included. The final dataset contained 93 environmental variables for each site.

2.6. Variable selection

We reduced the number of environmental variables by the combination of two approaches. First, we performed a collinearity analysis using the Variance Inflation Factor (VIF) followed by variable importance analysis until we found the 15 non-collinear ($VIF < 10$), most important variables shared by both study areas. Fifteen variables were considered a maximum number to avoid

Table 1

Importance ranking of the five non-collinear, similar and most important environmental variables shared between the two studies areas.

Australia		Spain	
Variable	Rank	Variable	Rank
Mean summer temperature	1	Elevation	1
Distance to river	2	Distance to river	2
SD NDVI in summer	3	Mean summer temperature	3
Elevation	4	Slope steepness	4
Slope steepness	5	SD NDVI in summer	5

SD: standard deviation; NDVI: normalized difference vegetation index of satellite-sensed spectral data.

over-fitting and more than enough for accurate SDMs. Second, we calculated dissimilarity by the Squared Mahalanobis Distance (SMD) between the environmental spaces of both study areas. Variables with a SMD lower than the p -value for the chi-squared distribution ($SMD < P(\chi_p^2)$) were considered dissimilar. Finally, we combined the two set of variables resulting from the first and second procedures to a set of shared variables which are non-collinear, important and similar (Supplementary material Appendix 1, Fig. A1).

Collinear variables were deleted for each study area separately (Franklin, 2010; Guisan and Thuiller, 2005), considering a Variance Inflation Factor (VIF) >10 as critical threshold (Quinn and Keough, 2002).

The least important of the non-collinear variables were eliminated stepwise (Elith and Leathwick, 2009) until 15 variables remained for each study area. The stepwise approach consisted of determining the importance of all variables each time a variable was excluded from the analysis.

Variable importance was determined by calculating the correlation (r) between model predictions including all remaining variables (a “full model”) and predictions excluding the tested variable (a “reduced model”). If a variable contributed little to a model, both outputs would be similar and $(1 - r)$ low, while the opposite would be the case for important variables (Thuiller et al., 2009). We deleted those variables that showed the lowest importance scores for the most accurate models. From the remaining 15 variables for each study area, we selected the variables that were similar between the two areas (Table 1).

The similarity of the environmental space was compared by pairwise calculation of the SMD between the 93 variables of each study area. The SMD measures the squared distance in a multivariate space between a value per variable and the means of all values (Quinn and Keough, 2002). It considers; (1) the correlation within the data, (2) the variance in all directions and (3) the covariance between variables (Farber and Kadmon, 2003; Filzmoser, 2004).

Variables with high SMD values can be considered as outliers (Farber and Kadmon, 2003; Filzmoser, 2004). The outlying (dissimilar) variables were identified and deleted in a step-by-step procedure by comparing the SMD with the Chi-squared distribution (χ_p^2); until the requirements were fulfilled, based on the assumption that the SMD can be considered to follow a Chi-squared distribution (Farber and Kadmon, 2003). In each step, the SMD of each variable was computed and compared with the value of the $P(\chi_p^2)$; If the highest value of SMD from one variable was higher than the $P(\chi_p^2)$ that variable was deleted; and the process was started again until the SMD of all remaining variables was $<P(\chi_p^2)$. A confidence level of 95% was used to exclude variables that were dissimilar from the population (Quinn and Keough, 2002). The final set of environmental variables for modelling contained the uncorrelated, most important and most similar variables (Table 1).

Table 2

The SDM algorithms in biomod2 used in this research and their required type of dependent variables (Thuiller et al., 2009).

SDM	Variable		Reference
ANN	Artificial neural networks	P/A	Lek and Guegan (1999)
BRT	Boosted regression trees	P/A	Elith et al. (2008)
CART	Classification and regression trees	P/A	Vayssières et al. (2000)
FDA	Flexible discriminant analysis	P/A	Trevor et al. (1994)
GAM	Generalized additive models	P/A	Guisan et al. (2002)
GLM	Generalized linear models	P/A	Guisan et al. (2002)
MaxEnt	Maximum entropy	P/B	Phillips et al. (2006)
MARS	Multivariate adaptive regression splines	P/A	Friedman (1991)
RF	Random forest	P/A	Breiman (2001)
SRE	Surface range envelope	P/B	Busby (1991)

P: Presence; A: Absence; B: Background. MaxEnt and SRE generate a large number (default: ten thousand) of random background absence point locations (aka pseudo-absences).

2.7. Models (biomod2)

The biomod2 package in R (R Core Development Team, 2014) predicts a species' distribution using ten different algorithms and evaluates models using a wide range of approaches as: Cohen's Kappa (K ; Cohen, 1960), Area Under the Curve (AUC) of the Receiver Operating Characteristic plot (ROC; Franklin, 2010), True Skills Statistics (TSS; Allouche et al., 2006), sensitivity (true positive rate) and specificity (true negative rate). We selected AUC as evaluation metric despite its drawbacks, because we compared two SDMs of one species in two areas with the same extent (Lobo et al., 2008, 2010).

AUC represents the model's discriminative capacity by plotting the commission error ($1 - \text{specificity}$; false positives) on the horizontal axis, vs. omission error (sensitivity; correctly identified positives) at the vertical axis. AUC ranges between 0.5–1, where 1 represents a perfect discrimination between presence and absence and 0.5 represents a random fit. AUC, as evaluation metric, is threshold independent (Franklin, 2010).

The ten algorithms included in biomod2 may be classified in four categories: regression, machine-learning, classification and enveloping. Regression-based SDMs build linear or non-linear relationships between species occurrence and their environmental space; that is Generalized Linear Models (GLMs) and Generalized Additive Models (GAMs), respectively. The machine-learning methods: Artificial Neural Networks (ANN), Boosted Regression Trees, (BRT), Multivariate adaptive regression splines (MARS), Maximum Entropy (MaxEnt) and Random Forest (RF), extract the environmental space of the species occurrence directly from the training data. Classification methods, Classification and Regression Trees (CART) and Flexible discriminate Analysis (FDA), are based on successive data partitions into homogeneous groups of responses. The enveloping method, Surface Range Envelope (SRE), defines the climatic conditions of the species occurrence and extrapolates the results to similar areas (Elith and Leathwick, 2009; Franklin, 2010; Guisan and Thuiller, 2005; Merow et al., 2014) (Table 2). The default settings of the biomod2 R-package version 3.1.25, (Phillips et al., 2006; Thuiller et al., 2009) were used.

2.8. Model calibration and evaluation

From the *P. cinnamomi* presence records a thousand random data partitions were created. Seventy percent was used for model calibration and the remaining thirty percent for evaluation. This yielded a thousand different fits per model, each with its own accuracy indicator. Mean and quantiles of AUC were calculated (Heikkinen et al., 2012; van Gils et al., 2014). Models with higher mean accuracy (AUC) values and smaller variations were considered more precise.

The impact of the number of variables on model performance was tested by a stepwise variable reduction. We started from the set of five environmental variables and stepwise deleted the least important variable until one variable remained (Elith and Leathwick, 2009; van Gils et al., 2008, 2014). The importance value per variable was calculated as described under *Variable selection*. This importance does not allow comparison of importance between algorithms. Therefore, we used a ranking system (Syphard and Franklin, 2009) to compare variable selection among algorithms. This procedure was carried out separately for both study areas.

2.9. Model transferability

Models trained and calibrated in Spain were transferred to Australia and vice-versa. Model transferability was quantified as the loss in accuracy relative to the accuracy when applied in-situ as follows (“Accuracy” refers to mean values of AUC) (Heikkinen et al., 2012):

$$\text{Accuracy}_{\text{INDEX}} = \frac{\text{Accuracy}_{\text{Transferred}}}{\text{Accuracy}_{\text{in-situ}}} \quad (1)$$

Accuracy index values equal to unity (=1) indicate that models performed equally in the training and transfer areas. Values <1 show that models performed better in the training area, and values >1 mean that models performed better in the transfer area.

3. Results

3.1. Spatial autocorrelation

The environmental variables used for both study sites showed positive spatial autocorrelation up to a distance of 10 km across resolutions in Spain and 20 km in Australia (Supplementary material Appendix 1, Fig. A2). However, the average distance between the point occurrences of *P. cinnamomi* in each study site (Australia = 3.1 km; Spain = 7.4 km) are larger the pixel resolution (2 km) used (Supplementary material Appendix 1, Fig. A2). Consequently, none of the variables has been removed from the set because of autocorrelation.

3.2. Comparison of variables between study areas

The total number of variables ($n = 93$) was reduced after collinearity ($n = 46$) and variable importance ($n = 15$) analysis; and twenty nine variables were found that were dissimilar between the two sites (Supplementary material Appendix 1, Fig. A1).

The Mahalanobis analysis revealed that the two study areas differ significantly (95% confidence level of the χ_p^2 distribution) for about a third ($n = 29$) of the environmental variables tested ($n = 93$). After removal of the dissimilar variables, the remaining ones ($n = 64$) showed similar environmental conditions in the two areas (Fig. 3).

3.3. Final variables

The most important, uncorrelated, similar variables selected to test transferability ($n = 5$) were ranked by their importance (1 = most important and 5 is least important) to model the distribution of *P. cinnamomi*. Distance to rivers is ranking high (2nd) in both countries and slope steepness low (4th or 5th). The three other variables are direct (mean summer temperature) or indirect (elevation; NDVI in summer) climate variables (Table 1 & Supplementary material Appendix 1, Table A3).

3.4. Model performance

GAM, MaxEnt, GLM, BRT and RF produced, in this order, the overall higher accuracy values when validated within the training area (Supplementary material Appendix 1, Fig. A3). GAM presented the highest AUC when five variables are used, but this AUC decreased at lower numbers of variables. BRT and RF showed reasonable AUCs, but accuracies also decreased at reduced numbers of variables (Supplementary material Appendix 1, Fig. A3).

Models trained and evaluated in Spain resulted in higher AUCs (>0.7) than those trained and evaluated in Australia (>0.55). The AUC values changed with the number of variables (5–1) in both situations (Supplementary material Appendix 1, Figs. A3 and A4). Models trained in Spain, except SRE, generated at each number of variables reasonable predictions (AUC > 0.70) at each number of variables. However, the AUC decreased with the number of variables for BRT, FDA, GAM, GLM, RF and SRE; increased with the number of variables tested for ANN and CART while the AUC values for MaxEnt and MARS remained akin (Supplementary material Appendix 1, Fig. A3).

Models trained in Australia, at each number of variables (1–5) presented accuracies hardly better than random (AUC > 0.55). GAM, BRT, RF, MaxEnt and GLM generated, in this order, better predictions. BRT, GAM, RF and SRE decreased their AUC with the number of variables, while the AUCs of ANN, FDA and GLM increased and those of CART, MARS and MaxEnt did not vary (Supplementary material Appendix 1, Fig. A4).

3.5. Model transferability

Model transferability results are presented in Fig. 4 and Supplementary material Table A4. GLMs showed the highest transferability for both directions and MaxEnt showed the most stable transferability across the number of environmental variables. Models transferred to Spain showed higher transferability values than vice-versa (transfer to Australia) (Supplementary material Appendix 1, Figs. A5 and A6). Model transfer to Spain generally presented a transferability index >1, i.e. AUC values of transferred models were higher than those in the training area (Fig. 4 and Supplementary material Table A4). Transferability indices for models transferred to Australia were <1, i.e. AUC values within the training area were higher than in the transferred area (Fig. 4 and Supplementary material Table A4). GLM, MaxEnt and, to some extent, GAM were the model algorithms that showed relatively high transferability indices. SRE presented reasonable transferability indices but its overall accuracy was hardly better than random. MaxEnt showed similar transferability across a number of variables (5–2) but not with one variable; (Fig. 4). However, MaxEnt performed better from Australia to Spain than vice-versa. The response of GAM to the number of variables depended on the direction of the transfer (Fig. 4). GLM and MaxEnt generated the highest AUCs combined with the highest transferability (Fig. 4 and Supplementary material Table A4).

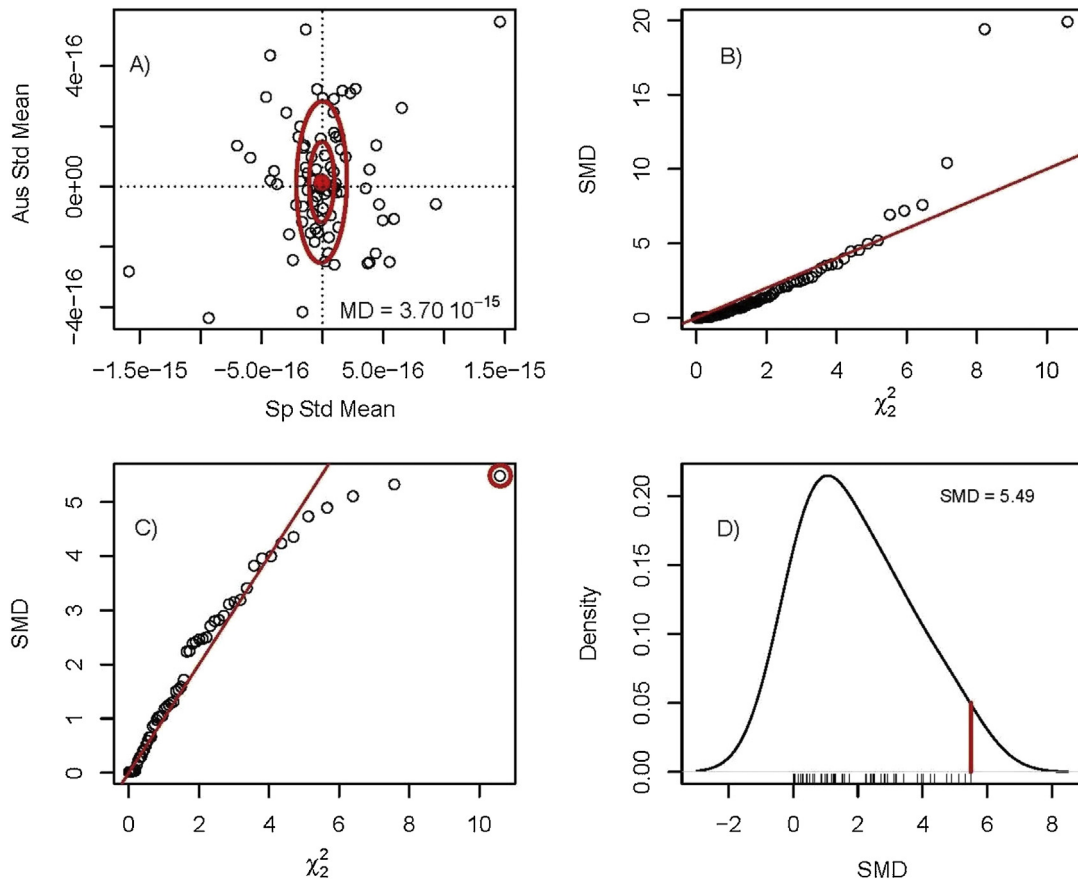


Fig. 3. Mahalanobis Distance (MD) analysis of the 93 environmental variables. (AUS: Australia; Sp: Spain; Std: Standardize; SMD: squared Mahalanobis distance; 5.49 is the SMD value at the edge of the distribution. (A) The Mahalanobis centroid position (red point) and two ellipses with a minor axis equal to 1 and 2 MD units away from the centroid.; (C) The point surrounded by the red circle is the remaining variable with the highest SMD. (D) The red line indicates the position of the remaining variable with the highest SMD. (A), displays the standardized mean of variable values in the study areas. Points close to the centroid and inside the ellipses represent similar variables. (B) shows a quantile–quantile plot of SMD values vs. quantiles of the χ^2_p -distribution using all 93 variables. Variables that differ much between the two study areas are points far away from the 1:1 line. (C), shows a Q–Q plot of SMD values vs. quantiles of the χ^2_p -distribution after deleting the 29 outliers, showing that some points representing the remaining variables are situated far away from the 1:1 line, although not significantly different according to the χ^2_p test. (D) shows the SMD density distribution after removal of the outliers.

An example of model projections within the training and transfer area is presented in Fig. 5.

4. Discussion

The SDM algorithms responded differently to the number of environmental variables used and to the direction of model transfer. Accurate predictive performance within the calibration area did not guarantee an equally good performance after intercontinental transfer. The GLM and MaxEnt algorithms showed a steadier AUC across both areas and number of variables compared to GAM. **Two** regression SDMs (GAM & GLM) were best transferred across the continents although GAM dropped in transferability accuracy when fewer than three variables were used. MaxEnt confirms its robustness in predictive modelling (Navarro-Cerrillo et al., 2011; van Gils et al., 2014) independently of the number of selected variables in addition to performing acceptable transferability results from Australia to Spain.

The difference in model transferability from Spain to Australia and vice-versa might be due to differences in the value ranges of environmental variables (Table 3).

Our Mahalanobis dissimilarity analysis compared mean values of the environmental variables. However, the ranges in values of four out of the five environmental variables used in modelling differed substantially between study areas. The range in elevation,

Table 3

Value ranges of the non-collinear, similar and important environmental variables used to test model transferability.

Variable	Value range	
	Australia	Spain
Distance to river	0–166 km	0–23 km
Elevation	0–782 m	0–3261 m
SD NDVI in summer	1–30	1–25
Mean summer temperature	17–24 °C	13–28 °C
Slope steepness	0–22%	0–100%

summer temperature, and slope steepness was much wider in Spain, while the distance-to-river showed a wider range in Australia (Table 3). These range differences might have reduced model transferability due to extrapolation outside the environmental ranges in the calibration areas of the model (Elith and Leathwick, 2009; Merow et al., 2014). To reduce uncertainty in model transferability we recommended that not only the (dis)similarity in the mean value of variables is standardized, but also their range (Merow et al., 2014). Particularly, the difference of ranges in elevation and distance-to-rivers (Table 3), together with the importance of both variables in Spain and the distance-to-rivers in Australia (Table 1) might have caused the lower model transferability from Spain to Australia. The high probability of the *P. cinnamomi* occurrence in the central river valley of the study area in Spain when

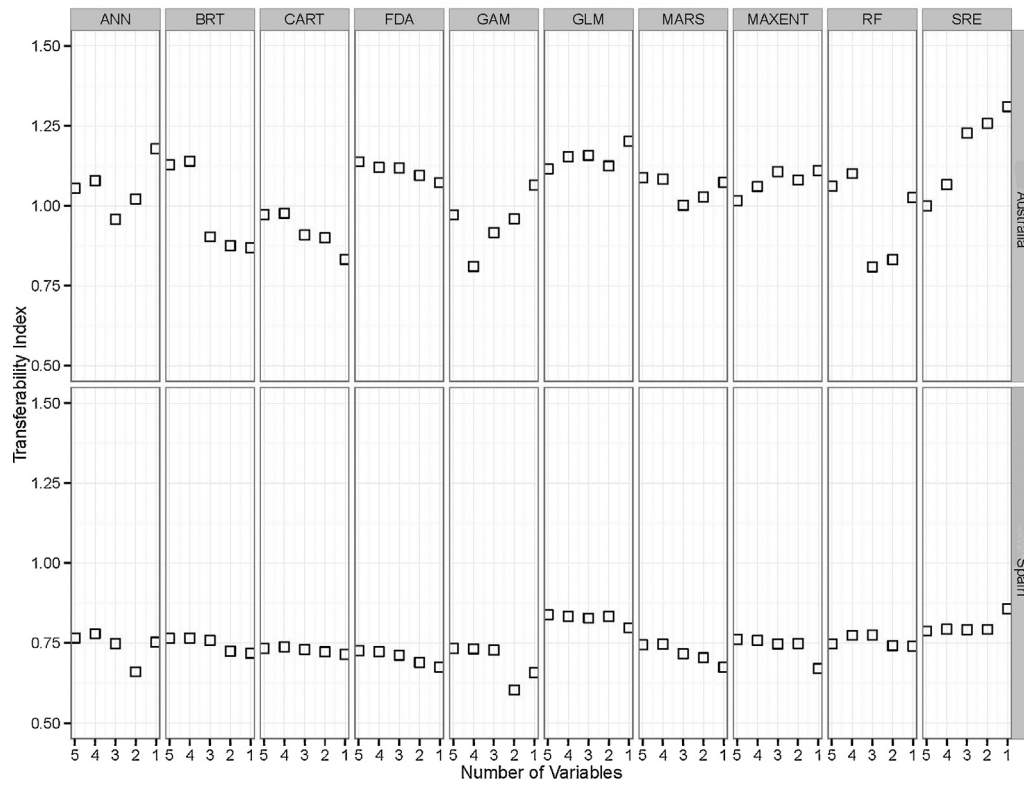


Fig. 4. Mean transferability index of the 10 models. The mean transferability index of models trained and evaluated within Australia and their subsequent evaluation in Spain (upper half) and vice-versa (lower half) using Eq. (1).

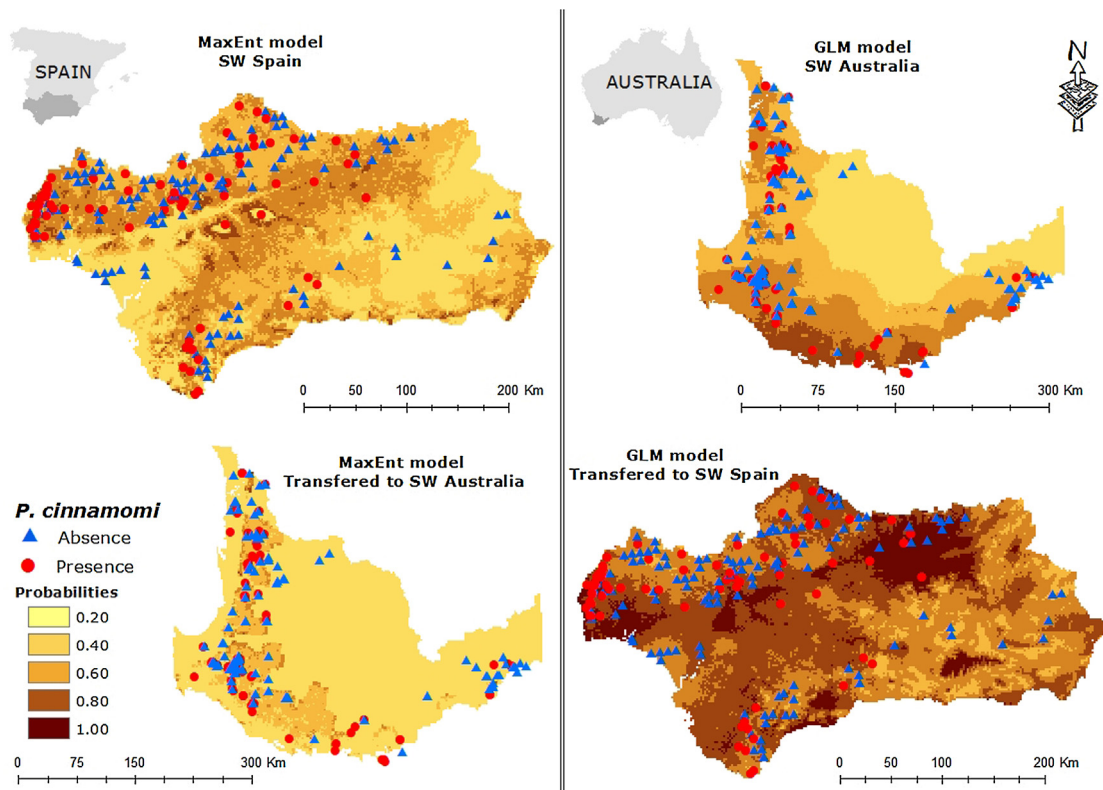


Fig. 5. Examples of predictive maps of models trained and evaluated within the same area and predictive maps of the transferred model.

predicted from models transferred from Australia (Fig. 5), might be due to the presence of *Phytophthora* (Casas and Díaz, 1985) but was not accounted for within the sampling design because the lack of Oak forest areas (Fig. 1). However, the superiority in transferability of GAM, GLM and MaxEnt over the other algorithms, even when considering dissimilar variables (Supplementary material Appendix 1, Figs. A3–A6), might be due to their capacity to build simple relationships between explanatory and response variables (Merow et al., 2014).

A further explanation for the differences in accuracies between models transferred to Australia and vice-versa could be the contrasting species sampling strategies in the study areas. The point locations of *P. cinnamomi* in Spain were obtained by systematic stratified sampling in Oak forest while in Australia the points resulted from a purposive, clustered sampling (Figs. 1 and 2) within areas where decline symptoms in trees and scrublands were identified. Edwards et al. (2006) found a positive impact of systematic stratified sampling on model accuracy and Franklin (2010) a negative one of clustered sampling. The importance of sampling design derives from covering the entire environmental range of species occurrence (Edwards et al., 2006) and avoiding spatial autocorrelation (Yu et al., 2015). However, spatial autocorrelation might be less relevant in our comparative case where both study areas showed a similar level of potential spatial autocorrelation.

Another issue that might have influenced model accuracy negatively is our maximum of five environmental variables when other studies have found accurate responses on model transferability with six (Wenger and Olden, 2012) and eight (Heikkinen et al., 2012) variables. However, Moreno-Amat et al. (2015) found that model complexity, i.e. a high ($n > 15$) number of variables impacts negatively on model transferability. Our results also showed that accuracy values varied according to the number of environmental variables included and the algorithm used. This finding underlines that optimal model complexity and number of explanatory variables remains an important issue in SDM. Franklin (2010) presented an illustrative table with the number and type of predictive variables used in SDMs studies. In most of the presented cases, the number and type of variables used was selected by the user based on what was available. The final set of ($n = 5$) environmental variables used in this study were climatic variables or proxies thereof and DEM derived similar to other studies of model transferability (Moreno-Amat et al., 2015; Wenger and Olden, 2012) on the other hand, Heikkinen et al. (2012) also considered land cover variables which were scorned by the analysis of collinearity. However, our findings showed that the optimal number of variables may depend (Fig. 4) on the SDM algorithm and the direction of transfer. For the better algorithms (GLM and MaxEnt) in our case three environmental variables seems to be sufficient both in-situ and for transfer.

The transferability test highlighted the poor transferability of most machine-learning algorithms (ANN, BRT, MARS and RF). Our results diverge from findings of Heikkinen et al. (2012) who found a better transferability performance by ANN, BRT and MaxEnt compared with regression algorithms (GLM and GAM), although they used more (eight) predictive variables. However, both Heikkinen's and this study are in agreement on GLM's good transferability. Wenger and Olden (2012) also identified GAM and GLM as suitable algorithms for model transfer purposes using six predictive variables. Further, Wenger et al. (2013) found accurate temporal transferability results also at lower numbers of variables (2–6) with GLM. The transferability strengths of GLM versus machine-learning algorithms might be related to avoidance of over-fitting when few variables are selected (Wiens et al., 2009). Over-fitting may result from fitting multivariate complex relationships of species occurrences and environmental variables, as is possible in many machine-learning techniques (Elith et al., 2006; Merow et al., 2014). Over-fitting means that more parameters are estimated than

needed. The number of parameters may be higher than the number of variables. For example, hidden nodes in ANN, data partitioning in BRT or number of trees in RF (Franklin, 2010) may also lead to over-fitting. Machine learning methods extract as much information as possible from the data available; therefore MaxEnt could have chosen simpler features when reducing the number of variables and increase complexity at higher numbers of variables. Heikkinen et al. (2012) and Wenger and Olden (2012) found poor transferability for RF in comparison with other methods using eight and six predictive variables, respectively. A low transferability for MARS has been previously reported both in space (Heikkinen et al., 2012) and over time (Prasad et al., 2006). Likewise, the tendency of CART to over-fit predictions has been previously suggested to explain its poor transferability over time with four predictive variables (Thuiller, 2003).

5. Conclusion

Most SDM algorithms predicted the *P. cinnamomi* distribution satisfactorily in Spain (AUCs > 0.70) and hardly better than random in Australia (AUCs > 0.55). The predictive powers of GAM, GLM and to some extent MaxEnt were superior to the other models tested. The lower model performance in Australia is attributed to the purposive sample strategy design (Edwards et al., 2006).

Our results also showed that high predictive performance within the training area is not a guarantee of high transferability, such as for example the machine-learning methods ANN, BRT, MARS and RF that had high AUC values in the training area but low transferability indices for AUC. Regression methods, especially GLM, showed less high, but acceptable AUC values and high transferability indices. CART and FDA had similar AUC values as GLM, but had lower transferability. MaxEnt presented accurate transferability performance from Australia to Spain but not in the reverse direction. SRE performed the least well both in terms of accuracy and transferability. In conclusion, GLMs combine good in-situ performance and transferability in both directions, while MaxEnt performed well in-situ but transferred only well in one direction, and provided the most robust response across the number of variables tested.

Acknowledgements

We thank Mike Stukely, Vegetation Health Service, Science Division, Department of Environment and Conservation, Western Australia for providing the *Phytophthora cinnamomi* data. We would also like to thank Carmen de Vos, editorial assistant at University of Wolverhampton, for English-language proofreading of an early draft. We acknowledge the project DIVERBOS (CGL2011-30285-C02-02). The Erasmus Mundus program (EU) is acknowledged for funding and Twente University and Cordoba University for hosting the first author during the research.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ecolmodel.2015.09.019>.

References

- ABARES, 2011. The 2011 Australian national map layers. Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES), Geoscience Australia, CSIRO Land and Water, the Department of Sustainability, Environment, Water, Population and Communities (Environmental Resources Information Network), the Australian Bureau of Statistics, and the National Centre for Social Applications of Geographic Information Systems ed. Australian Bureau of Agricultural

- and Resource Economics and Sciences (ABARES), Retrieved 15 October, 2012, from (www.daff.gov.au/abares/mcass).
- Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* 43, 1223–1232.
- Anderson, R.P., Gonzalez Jr., I., 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with Maxent. *Ecol. Modell.* 222, 2796–2811.
- Brasier, C.M., 1996. *Phytophthora cinnamomi* and oak decline in southern Europe. Environmental constraints including climate change. *Ann. Sci. Forest.* 53, 347–358.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
- Busby, J.R., 1991. BIOCLIM—a bioclimatic analysis and prediction system. In: Margules, C., Austin, M. (Eds.), *Nature Conservation: Cost Effective Biological Surveys and Data Analysis* (1991), vol. 1. CSIRO Publishing, Australia, pp. 64–68.
- Casas, A.T., Díaz, R.M.J., 1985. Etiología, importancia y distribución de la seca del garbanzo en el Valle del Guadalquivir. *Boletín de Sanidad Vegetal—Plagas* 11, 69–145.
- Cohen, J., 1960. A Coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20, 37–46.
- Consejería de Medio Ambiente, 2010. Estudio de procesos de mortalidad de pinares (sierra de Filabres y Baza - Almería) y especies del género *Quercus* en Andalucía. Consejería de Medio Ambiente, Córdoba.
- Edwards, T.C., Cutler, D.R., Zimmermann, N.E., Geiser, L., Moisen, G.G., 2006. Effects of sample survey design on the accuracy of classification tree models in species distribution models. *Ecol. Modell.* 199, 132–141.
- Elith, J.H., Graham, C.P., Anderson, R., Dudík, M., Ferrier, S., Guisan, A.J., Hijmans, R., Huettmann, F.R., Leathwick, J., Lehmann, A., Li, J.G., Lohmann, L.A., Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC.M., Overton, J., Townsend Peterson, A.J., Phillips, S., Richardson, K., Scachetti-Pereira, R.E., Schapire, R., Soberón, J., Williams, S.S., Wisz, M.E., Zimmermann, N., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129–151.
- Elith, J., Leathwick, J.R., 2009. Species distribution models: ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Syst. Evol. Syst.* 40, 677–697.
- Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. *J. Anim. Ecol.* 77, 802–813.
- Farber, O., Kadmon, R., 2003. Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecol. Modell.* 160, 115–130.
- Filzmoser, P., 2004. A multivariate outlier detection method. In: *Proceedings of the Seventh International Conference on Computer Data Analysis and Modeling*. Belarusian State University, Minsk, pp. 18–22.
- Franklin, J., 2010. *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge University Press, NY, U.S.A., pp. 340.
- Friedman, J.H., 1991. Multivariate adaptive regression splines. *Ann. Stat.* 19, 1–67.
- Gomez-Aparicio, L., Ibanez, B., Serrano, M.S., De Vita, P., Avila, J.M., Perez-Ramos, I.M., Garcia, L.V., Sanchez, M.E., Maranon, T., 2012. Spatial patterns of soil pathogens in declining Mediterranean forests: implications for tree species regeneration. *New Phytol.* 194, 1014–1024.
- Guisan, A., Edwards Jr., T.C., Hastie, T., 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol. Modell.* 157, 89–100.
- Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* 8, 993–1009.
- Hanberry, B.B., He, H.S., Dey, D.C., 2012. Sample sizes and model comparison metrics for species distribution models. *Ecol. Modell.* 227, 29–33.
- Hardham, A.R., 2005. *Phytophthora cinnamomi*. *Mol. Plant Pathol.* 6, 589–604.
- Heikkinen, R.K., Marmion, M., Luoto, M., 2012. Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography* 35, 276–288.
- Legendre, P., Fortin, M.J., 1989. Spatial pattern and ecological analysis. *Vegetatio* 80, 107–138.
- Lek, S., Guegan, J.F., 1999. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecol. Modell.* 120, 65–73.
- Lobo, J.M., Jiménez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecol. Biogeogr.* 17, 145–151.
- Lobo, J.M., Jiménez-Valverde, A., Hortal, J., 2010. The uncertain nature of absences and their importance in species distribution modelling. *Ecography* 33, 103–114.
- Lowe, S., Browne, M., Boudjelas, S., De Poorter, M., 2000. 100 of the World's Worst Invasive Alien Species: A Selection from the Global Invasive Species Database. Invasive Species Specialist Group Auckland, New Zealand.
- MAGRAMA, 2007. Tercer Inventario Forestal Nacional (IFN3). Ministerio de Medio Ambiente, Gobierno de España, Madrid.
- Merow, C., Smith, M.J., Edwards, T.C., Guisan, A., McMahon, S.M., Normand, S., Thuiller, W., Wüest, R.O., Zimmermann, N.E., Elith, J., 2014. What do we gain from simplicity versus complexity in species distribution models? *Ecography* 37, 1267–1281.
- Moreno-Amat, E., Mateo, R.G., Nieto-Lugilde, D., Morueta-Holme, N., Svenning, J.-C., García-Amorena, I., 2015. Impact of model complexity on cross-temporal transferability in Maxent species distribution models: An assessment using paleobotanical data. *Ecol. Modell.* 312, 308–317.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A., Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858.
- Navarro-Cerrillo, R.M., Hernandez-Bermejo, J.E., Hernandez-Clemente, R., 2011. Evaluating models to assess the distribution of *Buxus balearica* in southern Spain. *Appl. Veg. Sci.* 14, 256–267.
- Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci. Discuss.* 4, 439–473.
- Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecol. Modell.* 190, 231–259.
- Prasad, A., Iverson, L., Liaw, A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* 9, 181–199.
- Quinn, G.G.P., Keough, M.J., 2002. *Experimental Design and Data Analysis for Biologists*. Cambridge University Press, NY, U.S.A., pp. 557.
- R Core Development Team, 2014. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Sánchez, M.E., Sánchez, J.E., Navarro Cerrillo, R.M., Fernández, P., Trapero, A., 2003. Incidencia de la podredumbre radical causada por *Phytophthora cinnamomi* en masas de Quercus en Andalucía. *Boletín de Sanidad Vegetal—Plagas* 29, 87–108.
- Shearer, B.L., Crane, C.E., Barrett, S., Cochrane, A., 2007. *Phytophthora cinnamomi* invasion, a major threatening process to conservation of flora diversity in the South-west Botanical Province of Western Australia. *Aust. J. Bot.* 55, 225–238.
- Stukely, M., Webster, J., Ciampini, J., 2012. Results of *Phytophthora* sample testing. In: *Vegetation Health Service: Annual Report 2011–2012 Phytophthora Detection*.
- Syphard, A.D., Franklin, J., 2009. Differences in spatial predictions among species distribution modeling methods vary with species traits and environmental predictors. *Ecography* 32, 907–918.
- Thuiller, W., 2003. BIOMOD—optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biol.* 9, 1353–1362.
- Thuiller, W., Albert, C., Araujo, M.B., Berry, P.M., Cabeza, M., Guisan, A., Hickler, T., Midgley, G.F., Paterson, J., Schurr, F.M., Sykes, M.T., Zimmermann, N.E., 2008. Predicting global change impacts on plant species' distributions: future challenges. *Perspect. Plant Ecol. Evol. Syst.* 9, 137–152.
- Thuiller, W., Lafourcade, B., Engler, R., Araujo, M.B., 2009. BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography* 32, 369–373.
- Thuiller, W., Lavergne, S., Roquet, C., Boulangeat, I., Lafourcade, B., Araujo, M.B., 2011. Consequences of climate change on the tree of life in Europe. *Nature* 470, 531–534.
- Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* 46, 234–240.
- Trevor, H., Robert, T., Andreas, B., 1994. Flexible discriminant analysis by optimal scoring. *J. Am. Stat. Assoc.* 89, 1255–1270.
- Václavík, T., Kanaskie, A., Hansen, E.M., Ohmann, J.L., Meentemeyer, R.K., 2010. Predicting potential and actual distribution of sudden oak death in Oregon: prioritizing landscape contexts for early detection and eradication of disease outbreaks. *For. Ecol. Manage.* 260, 1026–1035.
- Václavík, T., Meentemeyer, R.K., 2009. Invasive species distribution modeling (iSDM): are absence data and dispersal constraints needed to predict actual distributions? *Ecol. Modell.* 220, 3248–3258.
- van Gils, H., Batsukh, O., Rossiter, D., Munthali, W., Liberatoscioli, E., 2008. Forecasting the pattern and pace of *Fagus* forest expansion in Majella National Park, Italy. *Appl. Veg. Sci.* 11, 539–546.
- van Gils, H., Conti, F., Ciaschetti, G., Westinga, E., 2012. Fine resolution distribution modelling of endemics in Majella National Park, Central Italy. *Plant Biosyst.* 146, 276–287.
- van Gils, H., Westinga, E., Carafa, M., Antonucci, A., Ciaschetti, G., 2014. Where the bears roam in Majella National Park, Italy. *J. Nat. Conserv.* 22, 276–287.
- Vayssières, M.P., Plant, R.E., Allen-Diaz, B.H., 2000. Classification trees: an alternative non-parametric approach for predicting species distributions. *J. Veg. Sci.* 11, 679–694.
- Veloz, S.D., 2009. Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *J. Biogeogr.* 36, 2290–2299.
- Verbruggen, H., Tyberghein, L., Belton, G.S., Mineur, F., Jueterbock, A., Hoarau, G., Gurgel, C.F.D., De Clerck, O., 2013. Improving transferability of introduced species' distribution models: new tools to forecast the spread of a highly invasive seaweed. *PLoS ONE* 8, e68337.
- Wenger, S.J., Olden, J.D., 2012. Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods Ecol. Evol.* 3, 260–267.
- Wenger, S.J., Som, N.A., Dauwalter, D.C., Isaak, D.J., Neville, H.M., Luce, C.H., Dunham, J.B., Young, M.K., Fausch, K.D., Rieman, B.E., 2013. Probabilistic accounting of uncertainty in forecasts of species distributions under climate change. *Global Change Biol.* 19, 3343–3354.
- Weste, G., Marks, G.C., 1987. The biology of *Phytophthora cinnamomi* in Australasian forests. *Annu. Rev. Phytopathol.* 25, 207–229.
- Wiens, J.A., Stralberg, D., Jongsomjit, D., Howell, C.A., Snyder, M.A., 2009. Niche, models, and climate change: assessing the assumptions and uncertainties. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19729–19736.
- Yu, F., Wang, T., Groen, T.A., Skidmore, A.K., Yang, X., Geng, Y., Ma, K., 2015. Multi-scale comparison of topographic complexity indices in relation to plant species richness. *Ecol. Complexity* 22, 93–101.
- Zentmyer, G.A., 1988. Origin and distribution of four species of *Phytophthora*. *Trans. Br. Mycol. Soc.* 91, 367–378.