Modeling Potential Distribution of Norway Maple (*Acer platanoides*) in Massachusetts, USA

Marlon Ortega & Tomáš Václavík
mortega@clarku.edu, tova@email.cz
Clark University, IDCE, Worcester, MA
April 2007

A comparative assessment of Mahalanobis typicality and ecological niche factor analysis

Predictive distribution modeling

Predictive Distribution Modeling is an innovative GIS method that allows us to create range maps of potential distribution of animals, plant assemblages or target habitats. The probability of distribution is calculated based on the known locations of species presence or absence and underlying environmental characteristics, i.e. precipitation, temperature, elevation, slope, etc. (Guisan and Thuiller 2005).
Various predictive models have been developed based on different conceptual considerations, statistical approaches and data availability. Past studies revealed that different methods produce incongruent results (Segurado and Araujo 2004). In this study, we assessed the performance of two models using presence data on *Acer platanoides* in Massachusetts.

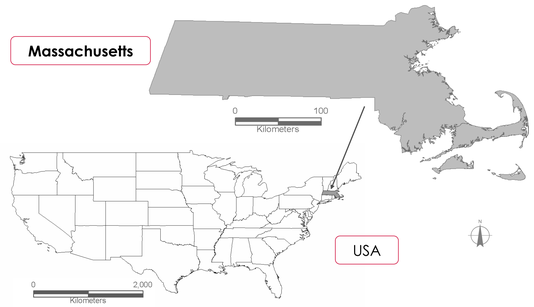
Mahalanobis typicality is derived from Mahalanobis distance which expresses the likelihood that a set of environmental variables at a specific location is typical to a known location of the species. The method tells us how typical is the area being analyzed compared to those used as a reference (Sangermano and Eastman 2007).

Environmental niche factor analysis compares, in the multidimensional space of ecological variables, the distribution of the localities where the focal species was observed to a reference set describing the whole study area. It computes factors (like Principal Component Analysis) that explain the major part of the ecological distribution of the species (Hirzel et al. 2002).

Norway maple

Acer platanoides is a deciduous tree native in continental Europe. It was introduced to the United States in the 18th century as an ornamental tree. It has naturalized to nearby woods, forest edges and fragmented forests, and from there has moved on to less disturbed habitats. *A. platanoides* is considered an invasive species in New England since it changes the original vegetation structure and out-competes native trees in forest canopy. Thus, it represents a threat to native species and forest habitats diversity (Anderson 1999).

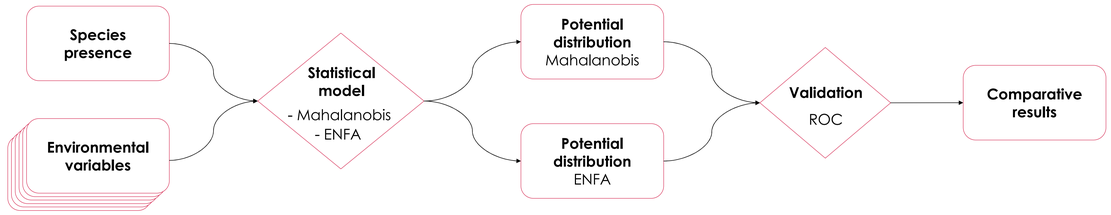
Study Area



Objectives

1. Create potential distribution maps of *A. platanoides* in the state of Massachusetts using two different statistical models (Mahalanobis and ENFA).
2. Compare the performance of both predictive models using Relative Operating Characteristics (ROC) assessment.

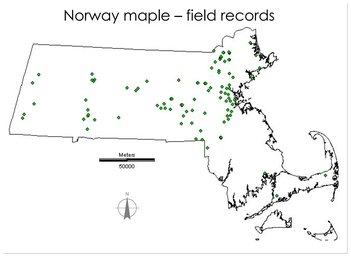
Methodology



Models inputs

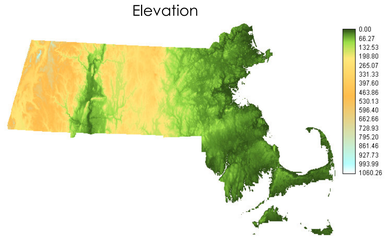
Species presence

Norway maple – field records

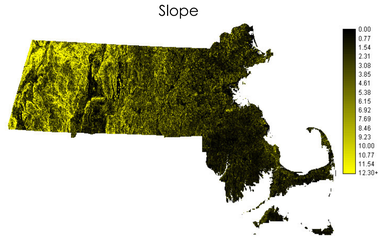


Environmental variables

Elevation

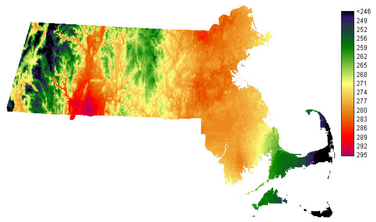


Slope

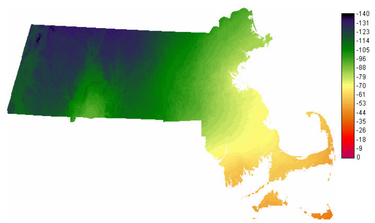


Average annual precipitation

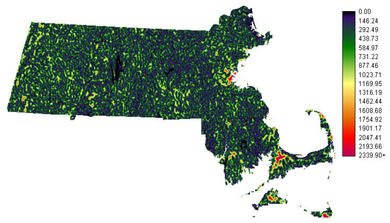
Maximum January temperatures



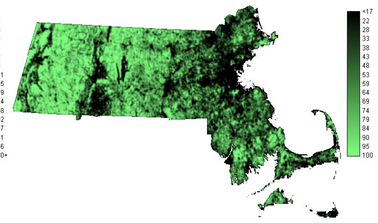
Minimum July temperatures



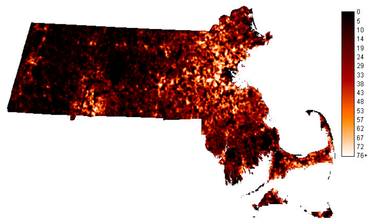
Distance to hydrological features



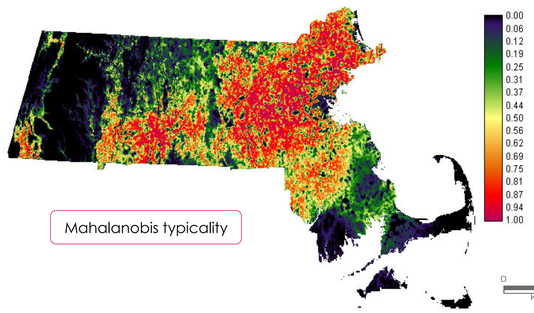
Frequency of forest



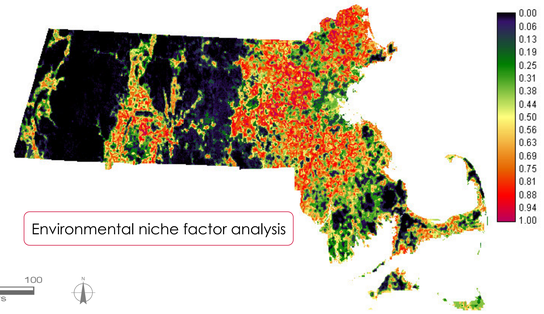
Frequency of residential development



Models outputs



Mahalanobis typicality



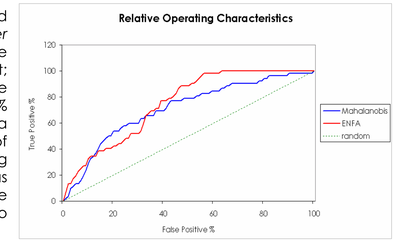
Environmental niche factor analysis

Data and software

Data on *Acer platanoides* presence (104 points) were downloaded from the Invasive Plant Atlas of New England [http://nbii-nin.ciesin.columbia.edu/ipane/index.htm].
Environmental variables of elevation and slope were derived from DEM obtained from the Clark University HERO database. Climate data used in the analysis were downloaded from the Worldclim database [http://www.worldclim.org]. The precipitation map represents the annual average precipitation for the period 1950-2000. The temperature data include average maximum July and minimum January temperatures for the same period. The source of hydrology and land cover data is MassGIS [http://www.mass.gov/mgis]. The distance to water bodies and frequency of forest and residential land cover categories were considered in the analysis. All the data were formatted to the resolution of 90x90 meters.
IDRISI 15 Andes Edition was used for data processing and Mahalanobis typicality analysis. Biomapper version 3.2 was employed for the ENFA analysis and the land cover frequency calculation.

Results

Both models produced suitability maps for *Acer platanoides*. The presence data were randomly split; 50% was used to calibrate the models, the other 50% to validate them. The Area Under the Curve (AUC) of the Relative Operating Characteristic (ROC) was calculated and the shape of the curve interpreted to compare both models.



The Mahalanobis results provided **ROC = 0.709**, ENFA results **ROC = 0.749**. The first and last 30% of the curve suggests that ENFA performs better in predicting the presence of *A. platanoides* as well as its absence within the study area. Mahalanobis provides better results in the mid-range of the predicted suitability.

Discussion and conclusions

ROC is a statistics frequently used for evaluation of species distribution models (Fielding and Bell 1997). It compares a rank map (e.g. predicted suitability map) against a Boolean map (e.g. species presence/absence). ROC indicates how well the events of the Boolean map falls within the high suitability values in the rank map (Pontius and Schneider 2001). The value of AUC is usually the only baseline to assess how well a model performs. However, ROC seems to have disadvantages for assessing models which use only presence data, such as Mahalanobis and ENFA. The ROC validation can be misinterpreted based on the definition of the pseudo-absence data for AUC calculation. Therefore, the validation should not be based solely on the absolute value of AUC. The shape of the curve needs to be taken into consideration, as it provides valuable information about the model behavior (Parmentier, *unpublished*).

Acknowledgements

Our thanks to Florencia Sangermano and Benoit Parmentier from Clark Labs, and professors Yelena Ogneva-Himmelberger, Gil Pontius and John Rogan from Clark University.

References

Anderson, R. (1999). Disturbance as a factor in the distribution of sugar maple and the invasion of Norway maple into a modified woodland. Rhodora 101(907): 264-273.
Hirzel, A.H., Hausser, J., Chessel, D. and Perrin, N. (2002). Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? Ecology, 83, 2027–2036.
Fielding. A. H. and Bell, J. F. (1997), A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24 (1): 38–49
Guisan, A. and Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models? Ecology Letters 8: 993-1009.
Parmentier, B. (2007). Improvements in the use of the Relative Operating Characteristic (ROC) Statistic. Clark University, Graduate School of Geography, Worcester, MA, USA, submitted for publication.
Pontius, R. G. and Schneider, L. C. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. Agriculture, Ecosystems and Environment 85: 239-248.
Sangermano F. and Eastman J. R. Linking GIS and Ecology. The use of Mahalanobis Tyipicalities to model species distribution. Proceedings of XI CONFIBSIG (Iberoamerican Conference of GIS). Buenos Aires, Argentina, May 2007.
Segurado, P. and Araujo, M. B. (2004). An evaluation of methods for modelling species distributions. Journal of Biogeography 31: 1555–1568.