Ecological Niche Modeling of the the invasive lionfish *Pterois miles* (Bennett, 1828) in the Mediterranean Sea

Poursanidis D.[1]

[1] *Foundation for Research and Technology - Hellas (FORTH), Institute of Applied and Computational Mathematics, N. Plastira 100, Vassilika Vouton, 70013, Heraklion, Greece, dpoursanidis@iacm.forth.gr*

Abstract

The lionfish *Pterois miles* (Bennett, 1828) is one of the most successful invasive aquatic species globally. Recently, few individuals have been reported from the East Mediterranean Sea. The potential distribution of the lionfish in the Mediterranean has been examined by using open access data from the GBIF, data from the new sites as well as environmental data from the BioOracle dataset. A maximum entropy model (MaxEnt) was developed based on these data, which was used to identify areas for the potential distribution of the species as well as to examine the factors that promote it.

Keywords: *Pterois miles*, Lionfish, Mediterranean, MaxEnt, BioOracle.

1. Introduction

In order to predict the distribution of species, the use of presence only data and environmental variables has become a diffusive method in order to address ecological and biogeographical questions (Peterson et al., 2011). Algorithms and models, distribution data and geophysical variables establish complex mathematical relationships and thus Species Distribution Models (SDMs) can be used to estimate the potential distribution of the studied species, the probability of presence under the studied conditions when these are projected into the geographical space (Soberon & Nakamura, 2009). One of the most important and useful application of this approach is the model projection to the future (application of climate change scenarios) or to new areas, for the prediction of the potential areas of invasion by non indigenous species in both aquatic and terrestrial environments (Thuiller et al., 2005; Wang et al., 2014). These potential areas are suitable areas yet not occupied by the species. Techniques to identify these areas need presence-only data (Jimenez-Valverde et al., 2008). Presence-only models include diverse methods and algorithms, among them MaxEnt (Phillips et al., 2006). The MaxEnt model approach is one of the most popular niche model applications in species distribution. It uses the principle of maximum entropy on presence-only data in order to estimate a set of functions that correlate habitat suitability and environmental variables with the potential geographic distribution.

Invasive species can be detrimental to ecosystems by altering habitats, competing with native organisms for limited resources, reducing biodiversity and causing extinctions of native species (Kalogirou, 2011; Katsanevakis et al., 2014). The lionfish *Pterois miles* (Bennett, 1828) and the congeneric *Pterois volitans* (Linnaeus, 1758) are among the most successful invasive aquatic species (Albins & Hixon, 2008). They have been introduced to the West Atlantic in the middle 90s' through aquarium trade and now they have been spread to the Caribbean Sea, the Sea of Mexico and the Florida coastal zone (Schofield, 2010). In the eastern Mediterranean, *P. miles* has been infrequently observed from 1992 until now (Golani & Sonin, 1992; Bariche et al., 2013; Evripidou, 2013). The aim of this work is to identify suitable areas for the lionfish in the Mediterranea Sea by incorporating presence data from open access sources like the Global Biodiversity Information Facility (GBIF) and marine environmental variables from the BioOracle database (Tyberghein et al., 2012) into the MaxEnt model. The results can support an early warning system for the identified areas as the species is highly venomous for the fishermen and the divers, persons that potentially could have a close encounter with it.

2. Materials and methods

2.1 Presence Data & Environmental Variables

Global Biodiversity Information Facility (GBIF, 2013) is one of the well-known open access data sources of the presence of animals and plants globally. It currently hosts >52.000.000 occurrences of >1.400.000 species globally. *Pterois miles* occurrence records have been extracted from the database in the form of *xml*. A total of 623 records are included in the database, including data from the invasive area of West Atlantic. Data for the Mediterranean have been extracted after the review of the aforementioned literature. Five records so far have been extracted and used. From the GBIF dataset, 503 georeferenced records have been used, as the rest didn't include location information. All data have been stored in a database for further use in the model process (Fig. 1).

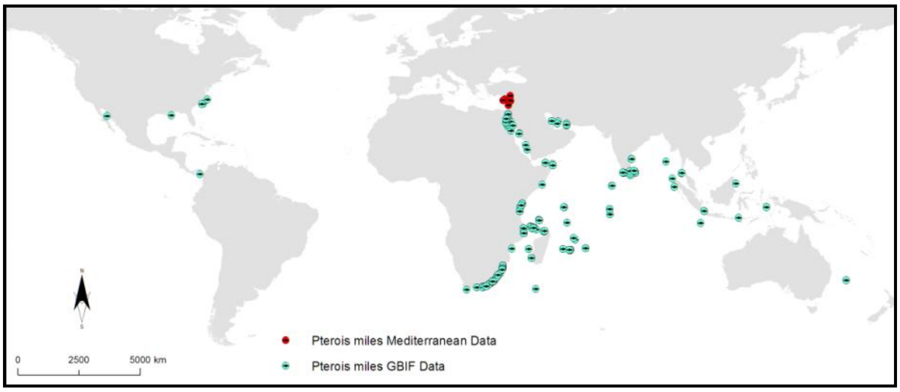

Fig. 1. Georeferenced data on the distribution of *Pterois miles* (Data for the Mediterranean Sea in red).

ORACLE dataset provides marine environmental information for global-scale applications. ORACLE offers an array of geophysical, biotic and climatic data at a spatial resolution 5 arcmin (9.2 km) in the ESRI ASCII format. Modelling techniques are sensitive to multicollinearity among the predictor variables used. Available predictor variables (23) were tested for multicollinearity based on Spearman's rank correlation (r.s, 0.7). This resulted in a subset of 9 uncorrelated predictor variables which were used as initial input to the models (Table 1). Presence data and environmental variables have been processed and analyzed by using the SDMtoolbox in ArcGIS 10.2 (Brown, 2014).

Table 1. Predictor variables used in the Model.

Name	Source	Unit	Abbr. name
Mean Sea Surface Temperature	BioOracle	^oC	sst_mean
Silicate	BioOracle	μmol/L	silica_mean
Salinity	BioOracle	PSS	sal_mean_an
Phosphate	BioOracle	μmol/L	phos_mean
Ph	BioOracle	unitless	ph_mean
Photosynthetically Available Radiation	BioOracle	Einstein/m^2/day	par_mean
Dissolved oxygen	BioOracle	ml/l	dissox
Chlorophyll A concentration	BioOracle	mg/m^3	chlo_mean
Calcite	BioOracle	mol/m^3	calcite_mean

2.2 MaxEnt.

MaxEnt is a machine-learning method that fits the probability distribution of maximum entropy for presences, constrained by the values of the pixels where the species has been found. The convergence to the optimal probability is guaranteed by the deterministic algorithms (Phillips et al., 2006; Elith et al., 2011). MaxEnt 3.3.3k (http://www.cs.princeton.edu/~schapire/maxent/) was run with default parameters. The prepared dataset has been split into two parts: (a) with (henceforth MOD1) and (b) without (MOD2) the Mediterranean records. This has been selected in order to use the Mediterranean records as validation data for the projection of the potential distribution in new Mediterranean areas.

3. Results.

When the 503 georeferenced data from GBIF have been used in the model, apart from the known places that *Pterois miles* occurs, including the expansion to the West Atlantic, new areas in the Mediterranean have been flagged as potential areas for future expansion (Fig. 2, Fig. 3). When the known locations from the Mediterranean Sea have been incorporated to the model analysis, the same areas in the Mediterranean as previous (MOD2) have been highlighted as potential areas for the expansion. By superimposing the known locations from the Mediterranean, these have been placed in areas of high probability of distribution (>0.8). Model performance has been tested by using the Area under the Receiver Operating Characteristic Curve, known as AUC. Is derived from the Receiver Operation Characteristics Curve (ROC) and is calculated by the MaxEnt app. In all models (MOD1 and MOD2) the value of AUC is > 0.9; the results are close to perfect prediction. Mean Temperature and Mean Chlorophyll have the most important contribution to the model.

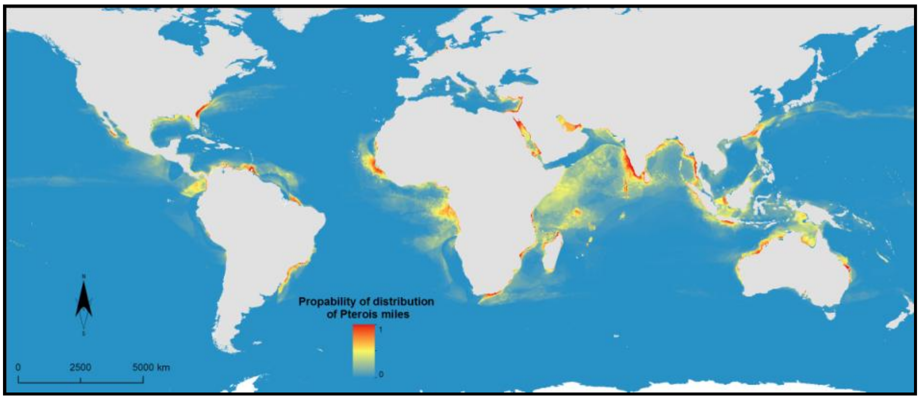

Fig. 2. Globally probability of the distribution of *Pterois miles* (MOD2).

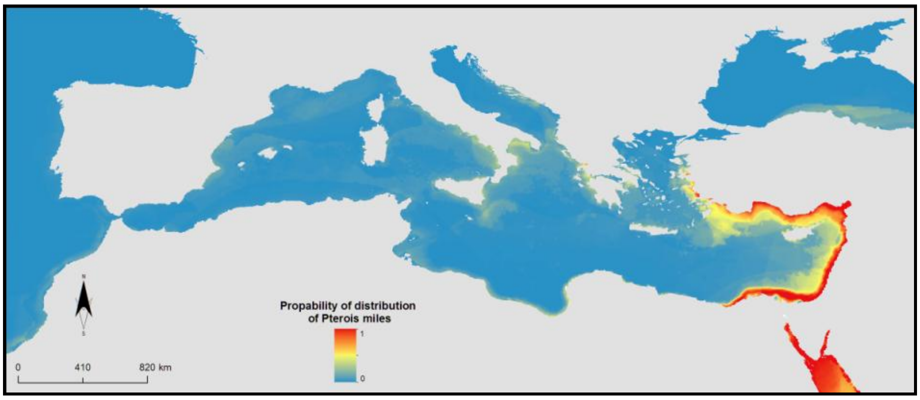

Fig. 3. Probability of the distribution of *Pterois miles* in the Mediterranean Sea (MOD2).

4. Conclusions/Discussion

The results provide an insight to the potential expansion of the lionfish in the Mediterranean Sea and the identification of the suitable areas for it to establish. By using the GBIF data only in order to train the model, excluding the Mediterranean records, the model show that for the certain species, the results are promising. Such approach should be used in high level planning in the scale of Mediterranean in the cases of newly invaders in order to identify the potential distribution. Also by using modelling techniques, the factors that will promote can be identified (based on the environmental data). The results can be used in an early warning system to inform the authorities in cases of species entrance which can cause problems to the human health, to high value biodiversity elements and to the aquaculture sector. Moreover, human activities like fishing can be affected by them, as in our studied species, the lionfish has strong poison in the spines and potentially can become harmful as well as due to the fact that are carnivores and feed on small crustaceans and fish, including the young of important commercial fish species. New records, if any, will be incorporated in the model in order to have a better understanding on the spread and the factors that promote it.

5. References

Albins, M.A. and Hixon, M.A. 2008. Invasive Indo-Pacific Lionfish *Pterois volitans* reduce recruitment of Atlantic coral-reef fishes. *Marine Ecology Progress Series*, 367, 233-238.

Bariche, M., Torres, M., Azzurro, E. 2013. The presence of the invasive lionfish *Pterois miles* in the Mediterranean Sea. *Mediterranean Marine Science*, 14(2): 292-294.

Brown, J.L. 2014. SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution*, 5, 694–700.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. *et al.* 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17, 43–57.

Evripidou, S. 2013. Toxic lionfish makes its way to Cyprus waters. http://www.cyprus-mail.com/cyprus/toxic-Lionfish-makes-its-way-cyprus waters/20130220 (Accessed 15 November 2014).

GBIF: Global Biodiversity Information Facility Backbone Taxonomy 2013. http://www.gbif.org/species/2334433 (Accessed 11 December 2014).

Golani, D. and Sonin, O. 1992. New records of the Red Sea fishes, *Pterois miles* (Scorpaenidae) and *Pteragogus pelycus* (Labridae) from the eastern Mediterranean Sea. *Japanese Journal of Ichthyology*, 39(2), 167-169.

Jimenez-Valverde, A., Lobo, J.M. and Hortal, J. 2008. Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions*, 14, 885–890.

Kalogirou, S. 2011. Alien fish species in the eastern Mediterranean Sea: Invasion biology in coastal ecosystems. p. 140. In: *Department of marine ecology*, University of Gothenburg, Gothenburg.

Katsanevakis, S., Wallentinus, I., Zenetos, A., Leppäkoski, E., Çinar, M.E. *et al.* 2014. Impacts of marine invasive alien species on ecosystem services and biodiversity: a pan-European review. *Aquatic Invasions*, 9(4), 391–423.

Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E. *et al.* 2011. *Ecological Niches and Geographic Distributions*. Princeton University Press, Princeton.

Phillips, S.J., Anderson, R.P. and Schapire, R.E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259.

Schofield, P.J. 2010. Update on geographic spread of invasive lionfishes (*Pterois volitans* [Linnaeus, 1758] and *P. miles* [Bennett, 1828]) in the Western North Atlantic Ocean, Caribbean Sea and Gulf of Mexico. *Aquatic Invasions*, 5 (1), S117–S122.

Soberon, J. and Nakamura, M. 2009. Niches and distributional areas: concepts, methods, and assumptions. p. 19644–19650. Proceedings of the National Academy of Sciences, USA.

Thuiller, W., Richardson, D.M., Pysek , P., Midgley, G.F., Hughes, G.O. *et al.* 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. Global Change Biology, 11 (12), 2234-2250.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F. *et al.* 2012. Bio-ORACLE: a global environmental dataset for marine species distribution modeling. *Global Ecology and Biogeography*, 21, 272–281.

Wang, W., Tang, X., Zhu, Q., Pan, K., Hu, Q., *et al.* 2014. Predicting the Impacts of Climate Change on the Potential Distribution of Major Native Non-Food Bioenergy Plants in China. PLoS ONE, 9 (11), 111587.