

## Research Article

## Modeling suitable habitat of invasive red lionfish *Pterois volitans* (Linnaeus, 1758) in North and South America's coastal waters

Paul H. Evangelista<sup>1</sup>, Nicholas E. Young<sup>1\*</sup>, Pamela J. Schofield<sup>2</sup> and Catherine S. Jarnevich<sup>3</sup>

<sup>1</sup>Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO 80523, USA

<sup>2</sup>US Geological Survey, Gainesville, FL 32653, USA

<sup>3</sup>US Geological Survey, Fort Collins Science Center, Fort Collins, CO 80526, USA

\*Corresponding author

E-mail: [Nicholas.Young@Colostate.edu](mailto:Nicholas.Young@Colostate.edu)

Received: 9 October 2015 / Accepted: 11 April 2016 / Published online: 5 May 2016

Handling editor: Charles Martin

### Abstract

We used two common correlative species-distribution models to predict suitable habitat of invasive red lionfish *Pterois volitans* (Linnaeus, 1758) in the western Atlantic and eastern Pacific Oceans. The Generalized Linear Model (GLM) and the Maximum Entropy (Maxent) model were applied using the Software for Assisted Habitat Modeling. We compared models developed using native occurrences, using non-native occurrences, and using both native and non-native occurrences. Models were trained using occurrence data collected before 2010 and evaluated with occurrence data collected from the invaded range during or after 2010. We considered a total of 22 marine environmental variables. Models built with non-native only or both native and non-native occurrence data outperformed those that used only native occurrences. Evaluation metrics based on the independent test data were highest for models that used both native and non-native occurrences. Bathymetry was the strongest environmental predictor for all models and showed increasing suitability as ocean floor depth decreased, with salinity ranking the second strongest predictor for models that used native and both native and non-native occurrences, indicating low habitat suitability for salinities < 30. Our model results also suggest that red lionfish could continue to invade southern latitudes in the western Atlantic Ocean and may establish localized populations in the eastern Pacific Ocean. We reiterate the importance in the choice of the training data source (native, non-native, or native/non-native) used to develop correlative species distribution models for invasive species.

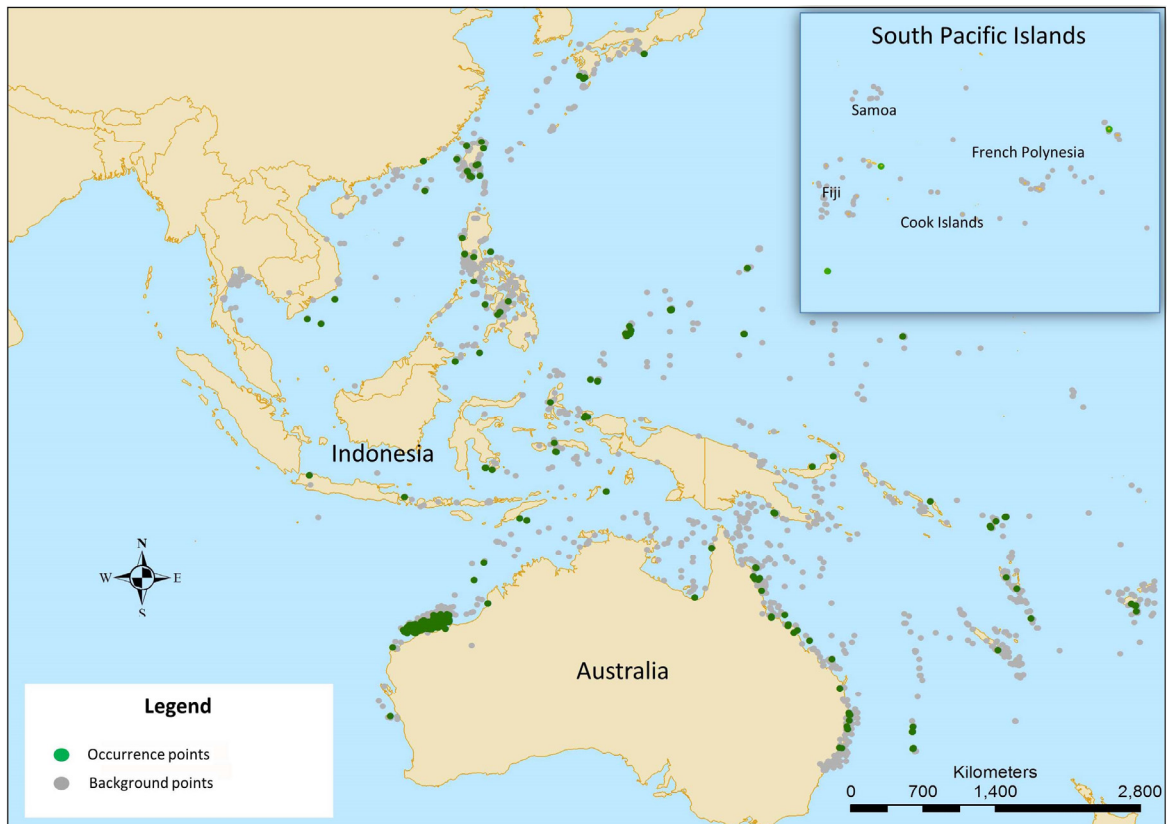
**Key words:** lionfish, species distribution model, Maxent, Generalized Linear Model, native and non-native species

### Introduction

Coastal marine systems are among the most heavily-invaded habitats on Earth (Grosholz 2002). Invasive marine species are of concern because they can cause a variety of negative environmental, economic and social impacts by threatening environmental integrity and biodiversity, marine industries such as fishing and tourism, and human health (Bax et al. 2003). On a global scale, most non-native marine species are transferred via shipping (e.g., ballast, hull fouling), although aquaculture, canal construction and the aquarium trade are also viable pathways for introduction (Ruiz et al. 1997; Molnar et al. 2008). For marine fishes, the aquarium trade is a major pathway for introduction, with more than 11 million individuals imported into the USA annually (Rhyne et al. 2012). The release of non-native aquarium

ornamental fishes into suitable coastal habitat in southeast Florida has resulted in a hotspot of non-native fish diversity (Semmens et al. 2004; Schofield et al. 2009).

Understanding the environmental conditions that allow or regulate a species' ability to persist may help provide insight into invasive potential and geographic spread, thus allowing early detection and rapid response. In terrestrial ecosystems, correlative species distribution models (SDM) have become popular and reliable tools for mapping habitat suitability and invasive potential of a new species (Guisan and Thuiller 2005; Elith et al. 2006; Evangelista et al. 2008; Elith and Leathwick 2009). Many SDMs are designed to statistically relate species occurrences to a variety of geospatial environmental data sets to identify habitat characteristics that define a species' range and potential distribution (Guisan et al. 2000;



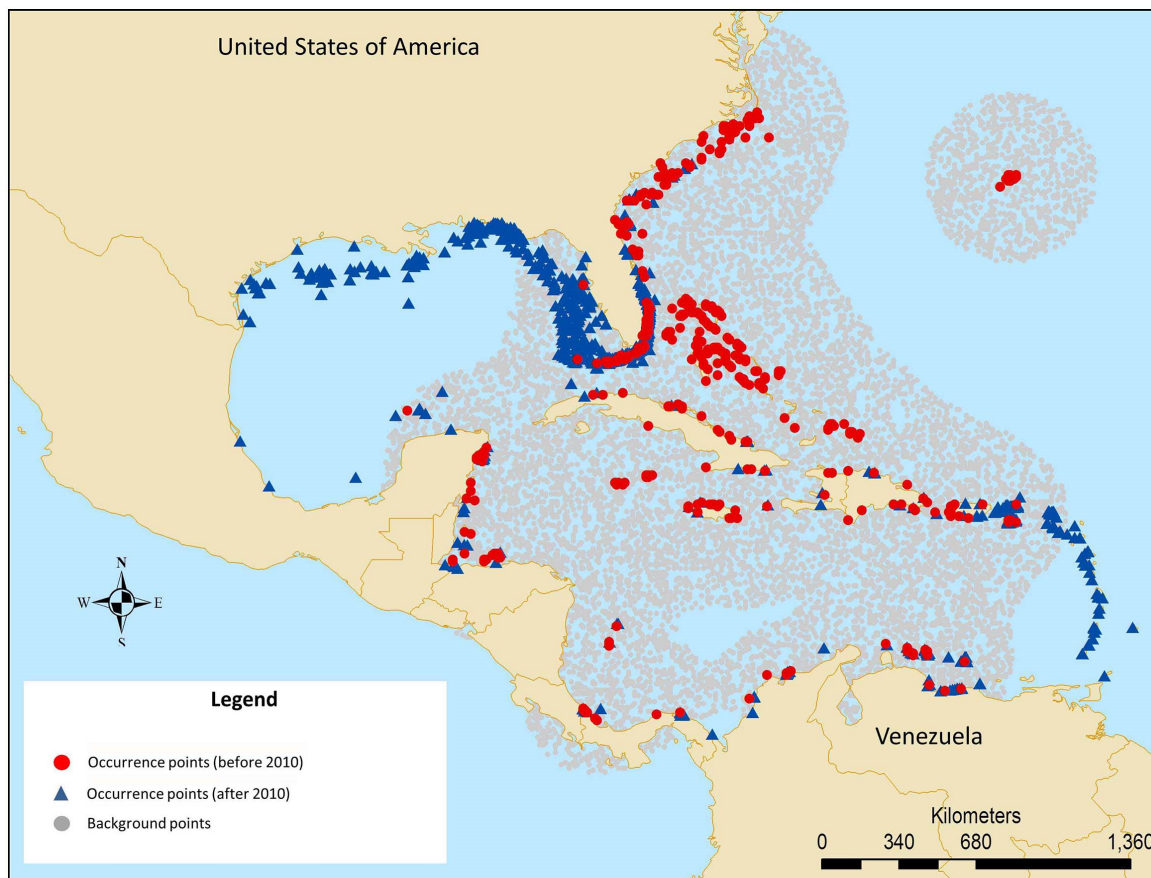
**Figure 1.** Native range of red lionfish (*Pterois volitans*) including occurrence data ( $n = 249$ ) and background points ( $n = 1,301$ ) used to build GLM and Maxent models. Occurrence data and background points were collected from the Global Biological Information Facility (GBIF 2015).

Phillips et al. 2004; Elith and Leathwick 2009). These models allow users the opportunity to visualize spatial predictions given multiple environmental variables, extrapolate information beyond observed occurrences, and evaluate model performance.

Applications of SDMs in marine environments remain relatively new, despite widespread use for terrestrial species (Robinson et al. 2011). This is due in part to sparse geospatial information on species occurrences and environmental attributes of marine systems (Kaschner et al. 2006; Bentlage et al. 2013). Furthermore, marine environments are complex systems with physical and biological variables that are distributed within a three-dimensional space that spans the ocean floor to the ocean surface. However, data on the availability of marine species occurrences are increasing with online databases such as the Ocean Biogeographic Information System (OBIS), the Global Biodiversity Information Facility (GBIF) and the US Geological Survey's Nonindigenous Aquatic Species (USGS-NAS) database. The quantity and quality of spatial marine environmental data sets has

recently improved enabling the application of SDMs in marine systems (Bučas et al. 2013; Jones et al. 2013).

Lionfish (*Pterois miles* [J.W. Bennett, 1828] and *P. volitans* [Linnaeus, 1758]) have become prolific in the northwestern Atlantic Ocean, Caribbean Sea and Gulf of Mexico where they are ecologically and economically harmful (Morris and Whitfield 2009; Albins and Hixon 2013). Lionfish are native to tropical coral reefs in the Indian Ocean and throughout the Indo-West Pacific region (Schultz 1986; Schofield 2009). First confirmed off the Florida coast in 1985 (Morris and Akins 2009), lionfish are popular in the aquatic pet trade, which is believed to be the source of their introductions (Semmens et al. 2004; Rhyne et al. 2012). Lionfish are the first marine fish to establish and spread rapidly in the region, and have done so at an alarming pace (Morris and Akins 2009; Schofield 2009, 2010). They appear to have few predators in their invaded range, limiting biotic control of their spread (Hackerott et al. 2013). Known to be prolific breeders in their invaded range, females can release up to 40,000+ eggs at one



**Figure 2.** Non-native range of red lionfish (*Pterois volitans*) including occurrence data reported before 2010 ( $n = 399$ ), occurrence data reported on or after 2010 ( $n = 522$ ), and background points ( $n = 10,000$ ) used to build GLM and Maxent models. Occurrence data were from the US Geological Survey’s Nonindigenous Aquatic Species database (Schofield et al. 2015) and background data were generated in the Software for Assisted Habitat Modeling (Morissette et al. 2013).

time and may breed three or more times a month (Morris and Whitfield 2009; Gardner et al. 2015). The duration of the planktonic larvae stage is estimated to last about a month, allowing larvae to disperse over large distances with ocean currents (Ahrenholz and Morris 2010; Morris et al. 2011). As their non-native range continues to expand, scientists question how far south lionfish may reach in the western Atlantic, and if there is potential for invasion in the eastern Pacific (Luiz et al. 2013; Johnston and Purkis 2014; Albins and Hixon 2013).

In this study, we used two SDMs, Generalized Linear Model (GLM) and Maximum Entropy (Maxent), to examine different approaches for modeling the distribution of *Pterois volitans* (hereafter referred to as red lionfish). The objectives of our study were to; (1) compare model results when developed with the use of native and/or non-native occurrence data and (2) determine the potential risk of new invasions of red lionfish within the non-native range.

## Methods

### Study areas

The native range of red lionfish is in the eastern Indo-Pacific region, stretching east-west from the French Polynesian islands to Thailand, and north-south from Japan to southern Australia (Figure 1). Its range differs from the closely-related *P. miles* whose native distribution is found further west in the Indian Ocean (Schofield et al. 2015). While there is some overlap between the native ranges of the two, they appear to be distinct species (Schultz 1986; Freshwater and Hamner 2009). Both species have established on the Atlantic coast of North, Central and South America — our primary study area. This includes Bermuda, the Gulf of Mexico and the Caribbean Sea, where lionfish have been ephemerally observed as far north as Newport, Rhode Island in the USA and as far south as Caracas, Venezuela (Figure 2; Schofield et al.

**Table 1.** Environmental predictor variables, including units of measure and sources, considered for modeling distribution of red lionfish.

Environmental variables	Units	Source
Calcite concentration	mol/m <sup>3</sup>	Tyberghein et al. 2012
Maximum chlorophyll A concentration	mol/m <sup>3</sup>	Tyberghein et al. 2012
Mean chlorophyll A concentration	mol/m <sup>3</sup>	Tyberghein et al. 2012
Minimum chlorophyll A concentration	mol/m <sup>3</sup>	Tyberghein et al. 2012
Chlorophyll A concentration range	mol/m <sup>3</sup>	Tyberghein et al. 2012
Maximum diffuse attenuation	m <sup>-1</sup>	Tyberghein et al. 2012
Mean diffuse attenuation	m <sup>-1</sup>	Tyberghein et al. 2012
Minimum diffuse attenuation	m <sup>-1</sup>	Tyberghein et al. 2012
Dissolved oxygen	ml/l	Tyberghein et al. 2012
Nitrate	μ mol/l	Tyberghein et al. 2012
Maximum photosynthetically available radiation	Einstein/m <sup>2</sup> /day	Tyberghein et al. 2012
Mean photosynthetically available radiation	Einstein/m <sup>2</sup> /day	Tyberghein et al. 2012
pH	none	Tyberghein et al. 2012
Phosphate concentration	μ mol/l	Tyberghein et al. 2012
Salinity	none	Tyberghein et al. 2012
Silicate concentration	μ mol/l	Tyberghein et al. 2012
Maximum sea surface temperature	Celsius	Tyberghein et al. 2012
Mean sea surface temperature	Celsius	Tyberghein et al. 2012
Minimum sea surface temperature	Celsius	Tyberghein et al. 2012
Sea surface temperature range	Celsius	Tyberghein et al. 2012
Global primary productivity averages	Pg	Behrenfeld and Falkowski 1997
Bathymetry	Km	Amante and Eakins 2009

2009; Schofield 2010; Schofield et al. 2015). Recently, one specimen was found off the southern coast of Brazil (Ferreira et al. 2015); however, it is unclear whether that individual constitutes a range extension or an independent introduction. While both species of lionfish have been documented in the non-native range, > 93% of the individuals identified to the species level have been red lionfish (Hamner et al. 2007; Freshwater and Hamner 2009; Betancur-R et al. 2011). Non-invaded waters of the Americas' Atlantic and Pacific coasts were of particular interest in our study.

### Data

We obtained 554 native-range red lionfish occurrence records from the Global Biological Information Facility (GBIF 2015). These data were vetted by removing records: a) that had a reported spatial resolution accuracy error greater than 10 km; b) which occurred more than once within our 10 km<sup>2</sup> sample unit, retaining only one record per environmental variable cell to remove pseudo-replication; c) without an observation date; and d) collected outside the red lionfish home range defined by Schofield et al. (2015). Following the vetting process, there were 249 occurrences remaining (Figure 1). Occurrence

data for red lionfish from its non-native range were acquired from the US Geological Survey's Nonindigenous Aquatic Species database (Schofield et al. 2015) which has been shown to be a valuable dataset for addressing pressing conservation issues (Scyphers et al. 2014). These data include contributions from a variety of sources and, as a whole, can be considered opportunistic observations (for details of this data aggregation, see Schofield et al. 2009). Each record in this database has a spatial location accuracy assessment of "approximate" or "accurate"; only occurrences recorded as accurate were used in model development and testing. Occurrences were further reduced to one sample per 10 km<sup>2</sup> to match the spatial resolution of the environmental variable. Following this process, 971 non-native occurrences for red lionfish remained for our analyses. These occurrence data were divided into two datasets; those reported before 2010 (n=399) were used to train models, while those reported during or after 2010 (n=572) were used to evaluate models (Figure 2). For models that used native/non-native occurrences, we combined the native and pre-2010 non-native data, resulting in 648 occurrence points for model training.

We compiled 22 environmental variables representing marine attributes that might explain red lionfish distribution (Table 1). We acquired bathymetry

(ocean floor depth) from the National Oceanic and Atmospheric Administration (Amante and Eakins 2009). BioOracle provided nutrient content, productivity metrics and annual surface temperatures (Tyberghein et al. 2012). Primary productivity averages were acquired from Oregon State University (Behrenfeld and Falkowski 1997). All environmental variables had global extents and, when necessary, were resampled using bilinear interpolation to a 10 km<sup>2</sup> grid for a uniform resolution across all layers. This resolution aligns with the environmental and occurrence data of our study and captures the general patterns at the hemispherical and global scales.

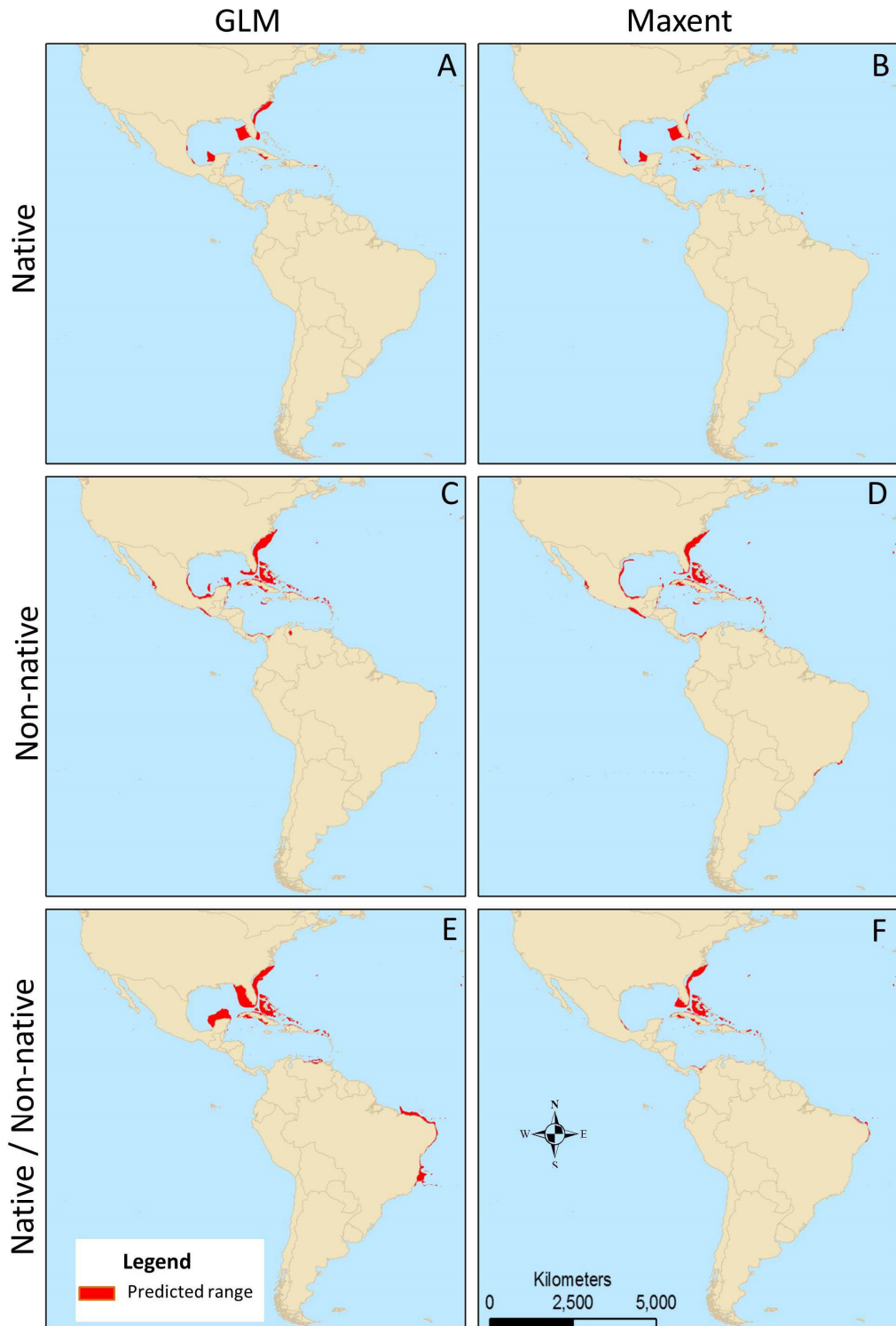
### Analysis

We developed six models, three using GLM and three using Maxent, to predict suitable habitat for red lionfish in the Atlantic and Pacific coasts of the Americas. Generalized Linear Models are a common regression-based method that use maximum likelihood to estimate model parameters and can generate non-normal response variable distributions (Guisan et al. 2002). Maxent is a machine learning statistical algorithm that employs the maximum entropy principle to address complex problems by making inferences from available data while avoiding constraints from the unknown (Phillips et al. 2004; Phillips et al. 2006). Both model approaches are well-documented in the literature and are regularly used for modeling distributions of invasive species in terrestrial systems, especially when extrapolating to novel areas since they can be parameterized to be more simplistic compared to other models (Evangelista et al. 2008; Evangelista et al. 2009; Václavík et al. 2012). The models were developed using red lionfish occurrence data from its native range, non-native range, and combined native and non-native ranges. We used the Software for Assisted Habitat Modeling (SAHM) within the Vistrails workflow software (Freire et al. 2006) to develop our GLM and Maxent models. The SAHM system was created to increase the efficiency of running multiple SDMs, and to store pre- and post-modeling processes to accurately reproduce analyses and results (Morissette et al. 2013). Data preprocessing tasks, executing model algorithms, evaluating models, and displaying spatial and textual results can all be performed in SAHM. For the models that used the GLM, we used the default settings in SAHM which included using the binomial model family and Akaike Information Criterion (AIC) for the simplification method, and not allowing squared terms for environmental variables. We also used the default settings in SAHM for Maxent models except for increasing the

betamultiplier (i.e., regularization value) to 3.0 to control for overfitting while allowing more generalized response curves to improve extrapolation (Elith et al. 2010). A Multivariate Environmental Similarity Surface (MESS) was generated for each model to identify novel regions outside the environmental range of model training data (Phillips et al. 2009). This provides a spatial reference of where the model is extrapolating beyond the data used to develop the model. Model predictions in novel areas should be interpreted with more caution than those areas within the defined environmental ranges (Elith et al. 2010).

To reduce the set of environment variables, we examined correlations among variables (i.e., Spearman, Pearson, and Kendall correlation coefficients) using the SAHM *CovariateCorrelationAndSelection* module. One of each pair of highly correlated variables ( $|r| > 0.7$  for any of the correlation coefficients; Dormann et al. 2013) was removed before model development. We based selections on known or probable ecological relevance to the species based on the literature and our own experiences with lionfish and invasive species ecology. In addition to potential correlation differences in environmental variables between the native and the non-native range, the drivers of a species' distribution can be different between ranges (Broennimann et al. 2007; Fitzpatrick et al. 2007). Therefore, we performed a separate correlation test and environmental variable selection for each range model (i.e., for the native range, the non-native range and the combined native/non-native range models).

Presence data used in the models were as described above in the data section. Because we did not have explicit absence locations, we ran GLM and Maxent models using background locations to capture the environmental characteristics available to lionfish (Elith and Leathwick 2009). To generate background locations only in areas accessible to the species, while also accounting for potential sample bias in the training data, we used the target guild background approach by obtaining GBIF data for the family Scorpaenidae from within the defined red lionfish native range as suggested by Phillips et al. (2009). These data were vetted as described above in the data section for the native occurrences and resulted in 1,301 background locations. To model the non-native range, we used the SAHM *BackgroundSurfaceGenerator* module to generate 10,000 background sample locations within a 95% isopleth Kernel Density Estimate (KDE) surface around the training locations that were in areas with ocean floor depths <2,000 m to attempt to match the sampling effort, the environment available to lionfish, the resolution of our environmental data, and the state of



**Figure 3.** Categorical model results showing the predicted distribution of red lionfish (*Pterois volitans*) in our study area. Models comparing GLM and Maxent models were constructed using native occurrence (A, B), non-native occurrence (C, D), and combined native/non-native occurrence (E, F) datasets. Predictions made outside the environmental range of each occurrence dataset were masked and are not included.

spread of the red lionfish (following Elith et al. 2010). It is important to note that the depth of the ocean floor does not necessarily represent the depth at which a red lionfish occurrence was found. In other words, the fish could be anywhere in the water column over the ocean floor. Further, if depth is indeed a barrier to red lionfish, having background locations beyond the physiological limits of lionfish are useful in defining these boundaries within the model. For our final models, using both the native and non-native occurrences, we randomly removed 7,931 background samples from the non-native occurrences so that each dataset had the same ratio of training samples to background samples for our combined native/non-native models which resulted in 645 training samples and 3,370 background samples. Final models were produced in two formats. The first with predictions displayed in a binary format (i.e., suitable or unsuitable habitat) and second in a continuous format (i.e., low to high suitability).

#### *Model evaluation*

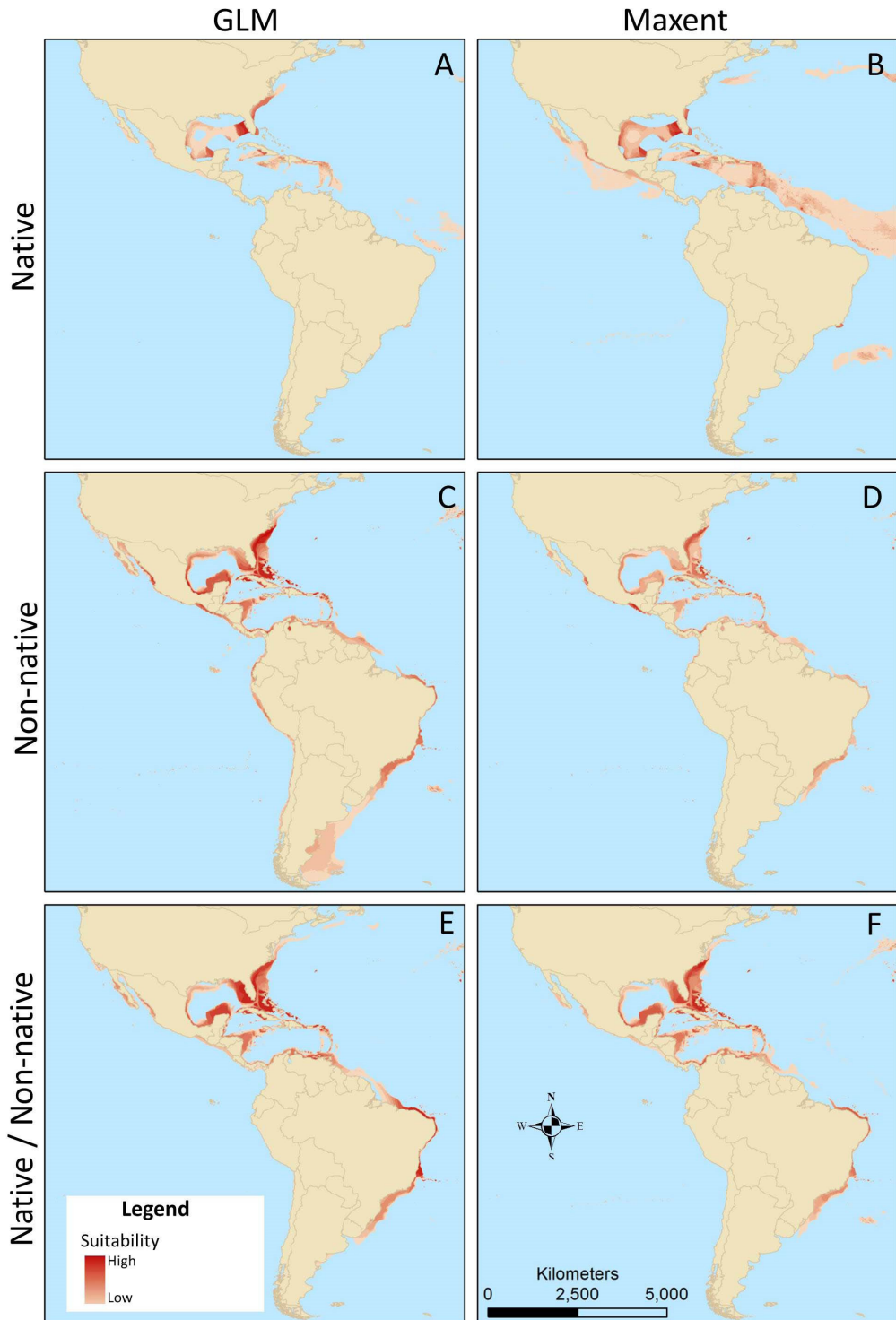
We evaluated all models using a 10 fold cross-validation of the training data and an independent test data set from our study area. The independent non-native occurrence points of lionfish locations were collected on or after 2010 and represent the existing and newly invaded areas, while background locations were generated based on a 95% isopleth KDE surface of these records to match the training data format. The maximum of sensitivity plus specificity divided by two was used as the threshold for model evaluations because it has been shown to perform well across different training data sets (Liu et al. 2013). We evaluated the models using threshold-dependent and threshold-independent evaluation metrics to test model performance. For each model, threshold-dependent evaluations were measured by specificity (true negative rate), sensitivity (true positive rate), and percent correctly classified. The Receiver Operating Characteristic Area Under the Curve (AUC) was used as the threshold-independent evaluation where sensitivity is plotted against 1-specificity for all possible thresholds (Pearce and Ferrier 2000) using the *ApplyModel* module in SAHM (Fielding and Bell 1997). The AUC is calculated using presence and absence (or in this case background locations) observations to measure the probability that a random present location has a higher predicted value than a random absence or background location (Fielding and Bell 1997). An AUC value of 0.5 indicates a model no better than random while a value of 1.0 indicates a model with a perfect fit to the data.

We also compared a measure of variable importance and response curves across models. SAHM measures variable importance using an algorithm independent method by calculating the change in AUC when the values for a predictor are permuted between the presence and background data. If this permutation causes a large change, then the predictor has more influence on the model. These AUC differences were then transformed into relative values of importance. SAHM also produces marginal response curves for each predictor in a model, where the response is calculated when all other predictors are held at their mean.

#### **Results**

Generally, the GLM and Maxent models had comparable spatial predictions for red lionfish habitat with each of the occurrence data sets we tested (Figures 3 and 4). For both models, the selected settings created models that performed well and did not show evidence of overfitting (i.e., no large difference between training and test evaluation metrics and relatively smooth response curves). The AUC evaluations and the percent correctly classified indicated good performance for all models (Table 2). Model evaluations on the independent test data were variable. For the independent test data, the GLM and Maxent models built with both native and non-native occurrences had the highest AUC values (0.88 and 0.86, respectively) and percent correctly classified (72.8 and 68.8, respectively). However, the models that used non-native occurrences also performed well with the independent test data. The models that used non-native and native/non-native occurrences had higher sensitivity when compared to specificity (Table 2). The models that used only native occurrences were the exception, suggesting that they were better at distinguishing where lionfish may occur rather than where they are likely not to occur (i.e., background environment). Based on these evaluation metrics, the models built with non-native and native/non-native occurrence data generally outperformed those that used only native occurrences.

Environmental variable selection and importance was variable among all models, however all models identified bathymetry as the largest contributing variable for predicting red lionfish habitat (Table 3). Ocean floor depths below 1500 m had negligible suitability followed by steeply increasing suitability to the shallower ocean floor depths (Figure 5). When only non-native occurrences were used in our models, the percent contribution of bathymetry for the GLM and Maxent models were 52.6% and 44.2%, respectively. The second highest variable contribution



**Figure 4.** Continuous model results showing the predicted distribution of red lionfish (*Pterois volitans*) in our study area. Models comparing GLM and Maxent models were constructed using native occurrence (A, B), non-native occurrence (C, D), and combined native/non-native occurrence (E, F) datasets. Predictions made outside the environmental range of each occurrence dataset were masked and are not included.

**Table 2.** Evaluation metrics for GLM and Maxent models using native, non-native and combined native/non-native occurrence data. Area Under the receiver operating Curve (AUC), percent correctly classified, sensitivity and specificity model evaluations are provided for each model. Evaluation test data includes training data (train), 10 fold cross-validation of training data (CV train) and non-native occurrences from 2010 or later (test).

Training data source	Model	AUC			% Correctly classified		Sensitivity		Specificity	
		Train	CV train	Test	Train	Test	Train	Test	Train	Test
Native occurrences	GLM	0.79	0.78	0.72	86.1	50.8	0.60	0.90	0.91	0.49
Native occurrences	Maxent	0.84	0.82	0.77	81.3	59.3	0.72	0.87	0.83	0.58
Non-native occurrences	GLM	0.82	0.81	0.85	73.6	68.8	0.79	0.96	0.73	0.67
Non-native occurrences	Maxent	0.88	0.87	0.85	77.8	65.8	0.86	0.97	0.77	0.64
Native/Non-native occurrences	GLM	0.79	0.79	0.88	70.9	72.8	0.75	0.92	0.70	0.72
Native/Non-native occurrences	Maxent	0.84	0.83	0.86	81.4	68.8	0.73	0.91	0.83	0.68

**Table 3.** Percent contribution of environmental variables for each model and dataset tested. Only variables selected for at least one model are listed. The highest (\*), second highest (†) and third highest (‡) contributing variables for each model are marked.

Environmental variables	Native occurrences		Non-native occurrences		Native/non-native occurrences	
	GLM	Maxent	GLM	Maxent	GLM	Maxent
Bathymetry	49.7*	24.7*	52.6*	44.2*	51.8*	48.4*
Chlorophyll A concentration range	–	–	–	–	–	1.6
Maximum chlorophyll A concentration	–	–	–	–	–	2.5
Maximum photosynthetically available radiation	3.4	–	–	–	–	5.9
Maximum sea surface temperature	9.4‡	6.3	–	–	–	0.1
Mean chlorophyll A concentration	–	14.8‡	–	–	–	–
Mean photosynthetically available radiation	–	4.5	13.1	10.9	3.3	1.6
Minimum sea surface temperature	0.5	5.8	13.9‡	14.2‡	–	2.3
Nitrate	4.3	4.8	–	–	21.7‡	–
pH	3.2	2.1	18.2†	21.1†	–	11.3‡
Phosphate concentration	0.2	12.2	–	6.3	–	6.0
Salinity	29.3†	21.1†	2.2	3.3	23.3†	18.3†
Silicate	–	3.7	–	–	–	2.0

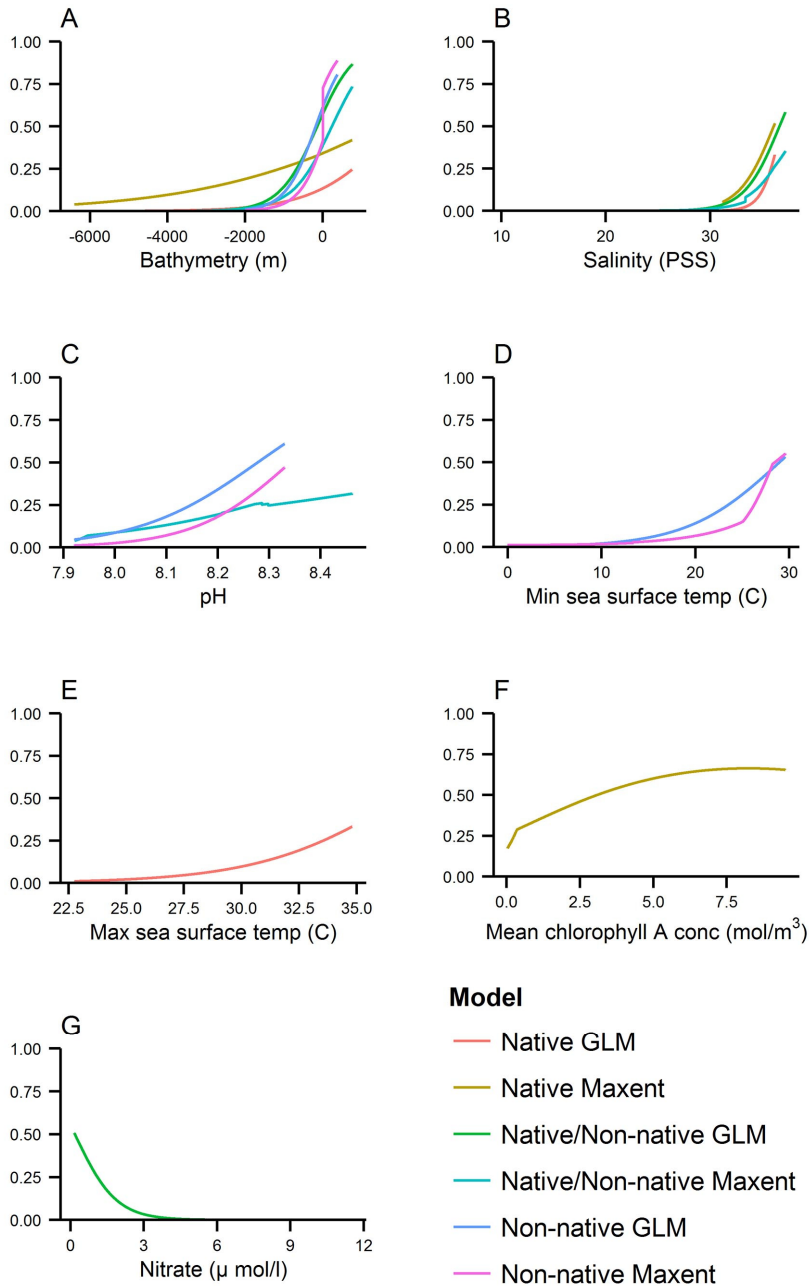
for models using non-native occurrence was pH, which was closely followed by minimum sea surface temperature. Salinity was the second highest contributor for the models that used native and native/non-native occurrences (Table 3) with suitability values increasing for salinity > 30 (Figure 5).

The MESS analysis showed the amount of novel area was higher for the models developed using the non-native range occurrences than those developed using the native or the native/non-native occurrences (Appendix 1). Non-native GLM and Maxent models had 92% and 95% of the total area classified as novel, respectively, based on the extent shown in Figure 2 and Figure 3, while native GLM and Maxent models had 73% and 76%, respectively, and GLM and Maxent for the native/non-native had 17% and 54% of the extent classified as novel, respectively.

Regions where red lionfish populations are known to be established in the non-native range were identified as novel from the MESS analysis of the native models (Appendix 2). This suggests that red lionfish are expanding into environments that are beyond the environmental range found in their native distribution. As expected, the native/non-native models had the least amount of novel area.

### Discussion

All of our models, whether they used native, non-native or the combined native/non-native occurrence data, predicted latitudinal expansion of red lionfish range along the Atlantic coast of North and South America. Models developed from non-native and native/non-native occurrences predict that red lionfish



**Figure 5.** Predictor variable marginal response curves for the top three contributing variables for each red lionfish (*Pterois volitans*) model (Native GLM, Native Maxent, Non-native GLM, Non-native Maxent, Native/non-native GLM, and Native/non-native Maxent). The environmental variables included: (A) bathymetry, (B) salinity, (C) pH, (D) minimum sea surface temperature, (E) maximum seas surface temperature, (F) mean chlorophyll A concentration, and (G) nitrate.

may, at least seasonally, occur as far north as New Jersey. There are reports of juvenile red lionfish occurring as far north as Rhode Island; however, these sightings are believed to be the result of larvae swept north by the Gulf Stream during summer

months. Those individuals are not expected to be able to overwinter and establish at that latitude because winter temperatures are thought to be too low for their survival (Kimball and Miller 2004). Models using native occurrences do not predict

significant expansion along the Atlantic coast of South America; however, the models built from native and native/non-native occurrences show potential invasions along Brazil's coast and as far south as the Gulf of San Jorge off the shores of Argentina. The predicted habitat from the native/non-native occurrence data support earlier predictions by Morris and Whitfield (2009) which implicate cold water temperatures off the coast of Uruguay as being the primary barrier to expansion (Kimball and Miller 2004). Based on the literature and current patterns of invasion, the models that use non-native and native/non-native occurrences seem to be more ecologically plausible than the models generated with only native occurrence data. These results are also supported by consistently higher evaluation metrics (Table 2). However, the GLM non-native model predicted marginally suitable habitat along the southern coast of Argentina where temperatures are likely to limit red lionfish persistence in those regions. The models that used only native occurrences appear to have under-predicted the distribution of red lionfish within the most invaded waters of Bermuda, the Gulf of Mexico, and the Caribbean Sea (Figure 3 and Appendix 1) and over-predict the potential spread in novel environments (Figure 4 and Appendix 2). All models under-predicted the suitable habitat in the Gulf of Mexico where many of post-2010 red lionfish occurrences have been documented (Figure 2), although the non-native models were better at capturing these new invasion than the other models (Figures 3 and 4).

Our model results for the eastern Pacific Ocean predict only a few areas that could be suitable for red lionfish. Although new invasions along the Pacific coast of the Americas are a serious concern, some scientists suggest environmental conditions are not as conducive to lionfish establishing in these regions (e.g., Johnston and Purkis 2014). The strongest predictions for red lionfish habitats along the Pacific coast were from our models that used non-native occurrences (Figure 4). Despite concerns of over-predicting, a closer examination of our results reveals that all of our models agree to some extent that red lionfish could potentially establish in the eastern Pacific Ocean and potentially cause declines of native fishes in these regions (Green et al. 2012). Of particular concern are the Gulf of California, Central American coastline from Guatemala to Panama, Peruvian coast, and the reefs around the Galapagos Islands. While there have been no documented observations of red lionfish in the western Pacific Ocean to date, the species is prevalent in the aquarium pet industry and could potentially be introduced along the Pacific coast.

The marginal response curves for environmental variables that were ranked in the top three contributing variables across multiple models displayed similar relationships (Figure 5). While recent studies related to salinity tolerance of lionfish have indicated lionfish can survive for a period of time at low salinities (Jud and Layman 2012; Schofield et al 2015), our models showed very low habitat suitability for salinities  $< 30$  (Figure 5). This pattern suggests that although lionfish can tolerate low salinities, they primarily survive and persist in higher salinity environments. Minimum temperature has also been suggested as a limiting factor to the range of lionfish, especially in the non-native range (Kimball and Miller 2004; Johnston and Purkis 2014). Laboratory studies found temperatures  $< 10^{\circ}\text{C}$  to be lethal to lionfish and lionfish stopped feeding at temperatures  $< 16.1^{\circ}\text{C}$  (Kimball and Miller 2004). Our results compliment these findings, showing low suitability between  $10$  and  $20^{\circ}\text{C}$  but steeply increasing above  $20^{\circ}\text{C}$  for minimum sea surface temperature. Interestingly, pH was a top contributing variable for three models where pH values at 7.9 represented lower suitability than values at 8.4. However, there has been little research to evaluate how lionfish respond to pH in laboratory or field settings, and the response we observed in the results may be indirectly related to an environmental variable not considered in the model.

The GLM and Maxent models performed similarly when tested with all three types of occurrence data. Not only were the models' spatial predictions comparable, but their evaluation metrics and environmental variable importance were similar as well. This suggests that the differences in our model results are largely attributed to the types of occurrence data used. The predicted suitable habitat of red lionfish was clearly different between the models that used native occurrences and those that used non-native and native/non-native occurrences. Similar studies for terrestrial species found that native occurrences alone were poor predictors for new invasions (Broennimann and Guisan 2008). This should not be surprising given that biotic constraints on a species within its native range are often absent in new environments (e.g., competition, predation). Evidence of red lionfish expanding beyond the known environmental limits of its native range has already been documented. For example, Schultz (1986) and Kulbicki et al. (2011) reported that red lionfish inhabit water column depths up to 50 m in its native Indo-Pacific range; however, the species has been found in much greater water column depths in its invaded range in the Atlantic Ocean (Schofield et al. 2015; Kulbicki et al. 2011;

Switzer et al. 2015). Spatial predictions and evaluation metrics for models that used both non-native and native/non-native occurrences were very similar, though GLM and Maxent models using native/non-native occurrences had better evaluation metrics and more consistent results between GLM and Maxent models than the non-native occurrences. A number of studies suggest that the combination of native and non-native occurrences produces the best models for predicting habitats at risk to invasion (Welk 2004; Mau-Crimmins et al. 2006; Broennimann and Guisan 2008; Jimenez-Valverde et al. 2011). Collectively, they capture a broader range of environmental conditions that are suitable for red lionfish populations. This is demonstrated by the reduction of novel environments produced by our MESS maps when comparing the environmental parameters defined by the native/non-native occurrences to those that only used native or non-native occurrence data by themselves (Appendix 1 and 2).

In conclusion, our study demonstrates that the potential suitable habitat for red lionfish in its invaded range extends beyond its current distribution. Coastlines in South America and perhaps in select areas along the eastern Pacific are most at risk for future lionfish invasions. Using an independent test dataset, we accurately predicted new invasion of lionfish in the western Atlantic. Although our study examines the relationship of a species with its habitat in a three-dimensional environment, our results are confined to a two-dimensional spatial output. This in part is due to the limited data in our environmental predictor variables throughout the water column. As with all models, there remains some level of uncertainty when trying to predict invasion risk. Our models assume there is an opportunity for introduction and do not account for local biological interactions such as predators (e.g., Hackerott et al. 2013) or fine-scale environmental features. Dispersal via ocean currents has been shown to be an important driver to where lionfish have been able to invade (Johnston and Purkis 2011; Johnston and Purkis 2015); however, this information was not included in our models. Our models reflect habitat suitability of lionfish given any means of introduction or transport, and do not directly account for dispersal barriers or specific facilitators. Models that used both native and non-native occurrences (Figures 3 and 4) provide the best predictions for future invasions of red lionfish. Those results have high evaluation metrics and better reflect the environmental extent of the species in both its native and non-native ranges.

## Acknowledgements

The authors thank the US Geological Survey (USGS) Invasive Species Program for funding this work. We also thank our colleagues at the USGS Fort Collins Science Center, the USGS Gainesville Center and the Natural Resource Ecology Laboratory at Colorado State University for their support, facilities and expertise. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government or Colorado State University.

## References

- Albins MA, Hixon MA (2013) Worst case scenario: potential long-term effects of invasive predatory lionfish (*Pterois volitans*) on Atlantic and Caribbean coral-reef communities. *Environmental Biology of Fishes* 96: 1151–1157, <http://dx.doi.org/10.1007/s10641-011-9795-1>
- Amante C, Eakins BW (2009) BW ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA, <http://dx.doi.org/10.7289/V5C8276M> (accessed January 2014)
- Ahrenholz DW, Morris JA (2010) Larval duration of the lionfish, *Pterois volitans* along the Bahamian Archipelago. *Environmental Biology of Fishes* 88: 305–309, <http://dx.doi.org/10.1007/s10641-010-9647-4>
- Bax N, Williamson A, Agüero M, Gonzalez E, Geeves W (2003) Marine invasive alien species: a threat to global biodiversity. *Marine Policy* 27: 313–323, [http://dx.doi.org/10.1016/S0308-597X\(03\)00041-1](http://dx.doi.org/10.1016/S0308-597X(03)00041-1)
- Behrenfeld M, Falkowski P (1997) Photosynthetic rates derived from satellite based chlorophyll concentration. *Limnology and Oceanography* 42: 1–20, <http://dx.doi.org/10.4319/lo.1997.42.1.0001>
- Bentlage B, Peterson AT, Barve N, Cartwright P (2013) Plumbing the depths: extending ecological niche modelling and species distribution modelling in three dimensions. *Global Ecology and Biogeography* 22: 952–961, <http://dx.doi.org/10.1111/geb.12049>
- Betancur-R, R, Hines A, Acero-P A, Orti G, Wilbur AE, Freshwater DW (2011) Reconstructing the lionfish invasion: insights into Greater Caribbean biogeography. *Journal of Biogeography* 38: 1281–1293, <http://dx.doi.org/10.1111/j.1365-2699.2011.02496.x>
- Broennimann O, Treier UA, Muller-Scharer H, Thuiller W, Peterson AT, Guisan A (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters* 10: 701–709, <http://dx.doi.org/10.1111/j.1461-0248.2007.01060.x>
- Broennimann O, Guisan A (2008) Predicting current and future biological invasions: both native and invade ranges matter. *Biology Letters* 11: 585–589, <http://dx.doi.org/10.1098/rsbl.2008.0254>
- Bučas M, Bergström U, Downie AL, Sundblad G, Gullström M, Von Numers M, Šiaulyš A, Lindegarth M (2013) Empirical modelling of benthic species distribution, abundance, and diversity in the Baltic Sea: evaluating the scope for predictive mapping using different modelling approaches. *ICES Journal of Marine Science* 70: 1233–1243, <http://dx.doi.org/10.1093/icesjms/fts036>
- Dormann CF, Elith J, Bacher, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ, Münkemüller T, McClean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurellm D, Lautenbach S (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36: 027–046, <http://dx.doi.org/10.1111/j.1600-0587.2012.07348.x>
- Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve

- prediction of species' distributions from occurrence data. *Ecography* 29: 129–151, <http://dx.doi.org/10.1111/j.2006.0960-7590.04596.x>
- Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics* 40: 677–697, <http://dx.doi.org/10.1146/annurev.ecolsys.110308.120159>
- Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution* 1: 330–342, <http://dx.doi.org/10.1111/j.2041-210X.2010.00036.x>
- Evangelista P, Kumar S, Stohlgren TJ, Jarnevich CS, Crall AW, Norman III JB, Barnett DT (2008) Modelling invasion for a habitat generalist and a specialist plant species. *Diversity and Distributions* 14: 808–817, <http://dx.doi.org/10.1111/j.1472-4642.2008.00486.x>
- Evangelista P, Stohlgren TJ, Morisette JT, Kumar S (2009) Mapping invasive tamarisk (*Tamarix*): a comparison of single-scene and time-series analyses of remotely sensed data. *Remote Sensing* 1: 519–533, <http://dx.doi.org/10.3390/rs1030519>
- Ferreira CEL, Luiz OJ, Floeter SR, Lucena MB, Barbosa MC, Rocha CR, Rocha LA (2015) First record of invasive lionfish (*Pterois volitans*) for the Brazilian Coast. *PLoS ONE* 10: e0123002, <http://dx.doi.org/10.1371/journal.pone.0123002>
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24: 38–49, <http://dx.doi.org/10.1017/S0376892997000088>
- Fitzpatrick MC, Weltzin JF, Sanders NJ, Dunn RR (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography* 16: 24–33, <http://dx.doi.org/10.1111/j.1466-8238.2006.00258.x>
- Freire J, Silva C, Callahan S, Santos E, Scheidegger C, Vo H (2006) Managing rapidly-evolving scientific workflows. In: Moreau L, Foster I (eds), Provenance and Annotation of Data. International Provenance and Annotation Workshop, IPAW 2006, Chicago, IL, USA, May 3–5, 2006, Revised Selected Papers. Springer, Berlin Heidelberg, pp 10–18, [http://dx.doi.org/10.1007/11890850\\_2](http://dx.doi.org/10.1007/11890850_2)
- Freshwater D, Hamner R (2009) Molecular evidence that the lionfishes *Pterois miles* and *Pterois volitans* are distinct species. *Journal of the North Carolina Academy of Science* 125: 39–46
- Gardner PG, Frazer TK, Jacoby CA, Yanong RP (2015) Reproductive biology of invasive lionfish (*Pterois* spp.). *Frontiers in Marine Science* 2: 1–10, <http://dx.doi.org/10.3389/fmars.2015.00007>
- Global Biological Information Facility (GBIF) (2015) <http://www.gbif.org/> (12 January 2014)
- Green SJ, Akins JL, Maljković A, Côté IM (2012) Invasive lionfish drive Atlantic coral reef fish declines. *PLoS ONE* 7: e32596, <http://dx.doi.org/10.1371/journal.pone.0032596>
- Grosholz E (2002) Ecological and evolutionary consequences of coastal invasions. *Trends in Ecology & Evolution* 17: 22–27, [http://dx.doi.org/10.1016/S0169-5347\(01\)02358-8](http://dx.doi.org/10.1016/S0169-5347(01)02358-8)
- Guisan A, Zimmermann NE, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147–186, [http://dx.doi.org/10.1016/S0304-3800\(00\)00354-9](http://dx.doi.org/10.1016/S0304-3800(00)00354-9)
- Guisan A, Edwards TCJ, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling* 157: 89–100, [http://dx.doi.org/10.1016/S0304-3800\(02\)00204-1](http://dx.doi.org/10.1016/S0304-3800(02)00204-1)
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8: 993–1009, <http://dx.doi.org/10.1111/j.1461-0248.2005.00792.x>
- Hackerott S, Valdivia A, Green S, Côté I (2013) Native predators do not influence invasion success of Pacific lionfish on Caribbean reefs. *PLoS ONE* 8: e68259, <http://dx.doi.org/10.1371/journal.pone.0068259>
- Hamner RM, Freshwater DW, Whitfield PE (2007) Mitochondrial cytochrome b analysis reveals two invasive lionfish species with strong founder effects in the western Atlantic. *Journal of Fish Biology* 71: 214–222, <http://dx.doi.org/10.1111/j.1095-8649.2007.01575.x>
- Jimenez-Valverde A, Peterson AT, Soberon J, Overton JM, Aragon P, Lobo JM (2011) Use of niche models in invasive species risk assessments. *Biological Invasions* 13: 2785–2797, <http://dx.doi.org/10.1007/s10530-011-9963-4>
- Johnston MW, Purkis SJ (2011) Spatial analysis of the invasion of lionfish in the western Atlantic and Caribbean. *Marine Pollution Bulletin* 62: 1218–1226, <http://dx.doi.org/10.1016/j.marpolbul.2011.03.028>
- Johnston MW, Purkis SJ (2014) Lionfish in the eastern Pacific: a cellular automaton approach to assessing invasion risk. *Biological Invasions* 16: 2681–2695, <http://dx.doi.org/10.1007/s10530-014-0696-z>
- Johnston MW, Purkis SJ (2015) A coordinated and sustained international strategy is required to turn the tide on the Atlantic lionfish invasion. *Marine Ecology Progress Series* 533: 219–235, <http://dx.doi.org/10.3354/meps11399>
- Jones MC, Dye SR, Warren R, Cheung WW (2013) Applying distribution model projections for an uncertain future: the case of the Pacific oyster in UK waters. *Aquatic Conservation: Marine and Freshwater Ecosystems* 23: 710–722, <http://dx.doi.org/10.1002/aqc.2364>
- Jud ZR, Layman, CA (2012) Site fidelity and movement patterns of invasive lionfish, *Pterois* spp., in a Florida estuary. *Journal of Experimental Marine Biology and Ecology* 414: 69–74, <http://dx.doi.org/10.1016/j.jembe.2012.01.015>
- Kaschner K, Watson R, Trites A, Pauly D (2006) Mapping worldwide distributions of marine mammal species using a relative environmental suitability (RES) model. *Marine Ecology Progress Series* 316: 285–310, <http://dx.doi.org/10.3354/meps316285>
- Kimball M, Miller J (2004) Thermal tolerance and potential distribution of invasive lionfish (*Pterois volitans/miles* complex) on the east coast of the United States. *Marine Ecology Progress Series* 283: 269–278, <http://dx.doi.org/10.3354/meps283269>
- Kulbicki M, Beets J, Chabanet P, Cure K, Darling E, Floeter SR, Galzin R, Green A, Harmelin-Vivien M, Hixon M, Letourneur Y, Lison de Loma T, McClanahan T, McIlwain J, MouTham G, Myers R, O'Leary JK, Planes S, Vigliola L, Wantiez L (2011) Distributions of Indo-Pacific lionfishes *Pterois* spp. in their native ranges: implications for the Atlantic invasion. *Marine Ecology Progress Series* 446: 189–205, <http://dx.doi.org/10.3354/meps09442>
- Liu C, White M, Newell G (2013) Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography* 40: 778–789, <http://dx.doi.org/10.1111/jbi.12058>
- Luiz OJ, Floeter SR, Rocha LA, Ferreira CE (2013) Perspectives for the lionfish invasion in the South Atlantic: are Brazilian reefs protected by the currents. *Marine Ecology Progress Series* 485: 1–7, <http://dx.doi.org/10.3354/meps10383>
- Mau-Crimmins TM, Schussman HR, Geiger EL (2006) Can the invaded range of a species be predicted sufficiently using only native-range data?: Lehmann lovegrass (*Eragrostis lehmanniana*) in southwestern United States. *Ecological Modelling* 193: 736–746, <http://dx.doi.org/10.1016/j.ecolmodel.2005.09.002>
- Molnar JL, Gamboa RL, Revenga C, Spalding MD (2008) Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment* 6: 485–492, <http://dx.doi.org/10.1890/070604>
- Morisette JT, Jarnevich CS, Holcombe TR, Talbert CB, Ignizio D, Talbert MK, Silva C, Koop D, Swanson A, Young NE (2013) VisTrails SAHM: visualization and workflow management for species habitat modeling. *Ecography* 36: 129–135, <http://dx.doi.org/10.1111/j.1600-0587.2012.07815.x>
- Morris JA, Akins JL (2009) Feeding ecology of invasive lionfish (*Pterois volitans*) in the Bahamian archipelago. *Environmental Biology of Fishes* 86: 389–398, <http://dx.doi.org/10.1007/s10641-009-9538-8>
- Morris JA, Whitfield P (2009) Biology, ecology, control and management of the invasive Indo-Pacific lionfish: an updated integrated assessment. NOAA/National Ocean Service Center for Coastal Fisheries and Habitat Research. NOAA Technical Memorandum NOS NCCOS, 99, 65 pp

- Morris JA, Shertzer KW, Rice JA (2011) A stage-based matrix population model of invasive lionfish with implications for control. *Biological Invasions* 13: 7–12, <http://dx.doi.org/10.1007/s10530-010-9786-8>
- Pearce J, Ferrier S (2000) An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecological Modelling* 128: 127–147, [http://dx.doi.org/10.1016/S0304-3800\(99\)00227-6](http://dx.doi.org/10.1016/S0304-3800(99)00227-6)
- Phillips SJ, Dudik M, Schapire RE (2004) A maximum entropy approach to species distribution modeling. In: Greiner R, Schuurmans D (eds) (2004) Twenty-first International Conference on Machine Learning – ICML '04. ACM Press, New York, New York, USA, p 83, <http://dx.doi.org/10.1145/1015330.1015412>
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231–259, <http://dx.doi.org/10.1016/j.ecolmodel.2005.03.026>
- Phillips SJ, Dudik M, Elith J, Graham CH, 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, <http://dx.doi.org/10.1890/07-2153.1>
- Rhyné AL, Tlustý MF, Schofield PJ, Kaufman L, Morris JA Jr, Bruckner AW (2012) Revealing the appetite of the marine aquarium fish trade: the volume and biodiversity of fish imported into the United States. *PLoS ONE* 7: e35808, <http://dx.doi.org/10.1371/journal.pone.0035808>
- Robinson LM, Elith J, Hobday AJ, Pearson RG, Kendall BE, Possingham HP, Richardson AJ (2011) Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. *Global Ecology and Biogeography* 20: 789–802, <http://dx.doi.org/10.1111/j.1466-8238.2010.00636.x>
- Ruiz GM, Carlton JT, Grosholz ED, Hines AH (1997) Global invasions of marine and estuarine habitats by non-indigenous species: mechanisms, extent, and consequences. *American Zoologist* 37: 621–632, <http://dx.doi.org/10.1093/icb/37.6.621>
- Schofield PJ (2009) Geographic extent and chronology of the invasion of non-native lionfish (*Pterois volitans* [Linnaeus 1758] and *P. miles* [Bennett 1828]) in the Western North Atlantic and Caribbean Sea. *Aquatic Invasions* 4: 473–479, <http://dx.doi.org/10.3391/ai.2009.4.3.5>
- Schofield PJ, Morris JA Jr, Akins L (2009) Field guide to nonindigenous marine fishes of Florida. NOAA Technical Memorandum NOS NCCOS 92, [http://fl.biology.usgs.gov/Marine\\_Fish\\_ID/index.html](http://fl.biology.usgs.gov/Marine_Fish_ID/index.html)
- Schofield P (2010) Update on geographic spread of invasive lionfishes (*Pterois volitans* and *P. miles*) in the Western North Atlantic Ocean, Caribbean Sea and Gulf of Mexico. *Aquatic Invasions* 5: S117–S122, <http://dx.doi.org/10.3391/ai.2010.5.S1.024>
- Schofield PJ, Morris JA, Langston JN, Fuller PL (2015) *Pterois volitans/miles* FactSheet. USGS Nonindigenous Aquatic Species Database, Gainesville, FL. <http://nas.er.usgs.gov/queries/factsheet.aspx?SpeciesID=963> (accessed 15 June 2014)
- Scyphers SB, Powers SP, Akins JL, Drymon JM, Martin CW, Schobernd ZH, Schofield PJ, Shipp RL, Switzer TS (2014) The role of citizens in detecting and responding to a rapid marine invasion. *Conservation Letters* 8: 242–250, <http://dx.doi.org/10.1111/conl.12127>
- Schultz E (1986) *Pterois volitans* and *Pterois miles*: two valid species. *Copeia* 1986: 686–690, <http://dx.doi.org/10.2307/1444950>
- Semmens B, Buhle ER, Salomon AK, Pattengill-Semmens CV (2004) A hotspot of non-native marine fishes: evidence for the aquarium trade as an invasion pathway. *Marine Ecology Progress Series* 266: 239–244, <http://dx.doi.org/10.3354/meps266239>
- Switzer TS, Tremain DM, Keenan SF, Stafford CJ, Parks SL, McMichael Jr, RH (2015) Temporal and Spatial Dynamics of the Lionfish Invasion in the Eastern Gulf of Mexico: Perspectives from a Broad-scale Trawl Survey. *Marine and Coastal Fisheries* 7: 10–17, <http://dx.doi.org/10.1080/19425120.2014.987888>
- Tyberghein L, Verbruggen H, Pauly K, Troupin C, Mineur F, Clerck O (2012) Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography* 21: 272–281, <http://dx.doi.org/10.1111/j.1466-8238.2011.00656.x>
- Václavík T, Kupfer JA, Meentemeyer RK (2012) Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). *Journal of Biogeography* 39: 42–55, <http://dx.doi.org/10.1111/j.1365-2699.2011.02589.x>
- Welk E (2004) Constraints in range predictions of invasive plant species due to non-equilibrium distribution patterns: purple loosestrife (*Lythrum salicaria*) in North America. *Ecological Modelling* 179: 551–567, <http://dx.doi.org/10.1016/j.ecolmodel.2004.04.020>

## Supplementary material

The following supplementary material is available for this article:

**Appendix 1.** Categorical model predictions with novel areas shown from MESS analysis.

**Appendix 2.** Continuous model predictions with novel areas shown from MESS analysis.

This material is available as part of online article from:

[http://www.aquaticinvasions.net/2016/Supplements/AI\\_2016\\_Evangelista\\_et\\_al\\_Supplement.pdf](http://www.aquaticinvasions.net/2016/Supplements/AI_2016_Evangelista_et_al_Supplement.pdf)