The Hyper-Envelope Modeling Interface (HEMI): A Novel Approach Illustrated Through Predicting Tamarisk (*Tamarix* spp.) Habitat in the Western USA

Jim Graham · Nick Young · Catherine S. Jarnevich ·
Greg Newman · Paul Evangelista ·
Thomas J. Stohlgren

Abstract Habitat suitability maps are commonly created by modeling a species' environmental niche from occurrences and environmental characteristics. Here, we introduce the hyper-envelope modeling interface (HEMI), providing a new method for creating habitat suitability models using Bezier surfaces to model a species niche in environmental space. HEMI allows modeled surfaces to be visualized and edited in environmental space based on expert knowledge and does not require absence points for model development. The modeled surfaces require relatively few parameters compared to similar modeling approaches and may produce models that better match ecological niche theory. As a case study, we modeled the invasive species tamarisk (*Tamarix* spp.) in the western USA. We compare results from HEMI with those from existing similar modeling approaches (including BioClim, BioMapper, and Maxent). We used synthetic surfaces to create visualizations of the various models in environmental space and used modified area under the curve (AUC) statistic and akaike information criterion (AIC) as measures of model performance. We show that HEMI produced slightly better AUC values, except for Maxent and better AIC values overall. HEMI created a model with only ten parameters while Maxent produced a model with over 100 and BioClim used only eight. Additionally, HEMI allowed visualization and editing of the model in environmental space to develop alternative potential habitat scenarios. The use of Bezier surfaces can provide simple models that match our expectations of biological niche models and, at least in some cases, out-perform more complex approaches.

Keywords: Habitat suitability modeling · Species distribution modeling · *Tamarix* · Tamarisk · Species niche · Bezier curves

Introduction

Habitat suitability modeling, which is closely related to species distribution modeling, is used to understand and predict the potential distribution of a species and is used in many areas of conservation and restoration ecology (Elith and Leathwick 2009). Habitat suitability modeling can provide critical insight into the potential spatial and temporal distribution of a species' population (Franklin 2009; Penman and others 2010), how species respond to novel environments (Hinojosa-Diaz and others 2009), locating rare and endangered species (Guisan and others 2006), exploring the distribution of invasive species (Morisette and others 2006; Evangelista and others 2008; Jarnevich and others 2011), and defining conservation priority areas (Fuller and others 2008).

In most cases, habitat suitability models integrate data on the observed location (i.e., presence points) of a species with environmental data (e.g., temperature, precipitation, slope, soil type) to create a model of the species' niche in environmental space and then convert the results to

J. Graham (✉)
Department of Environmental Science & Management,
Humboldt State University, Arcata, CA 95521, USA
e-mail: James.Graham@humboldt.edu

N. Young · G. Newman · P. Evangelista
Natural Resource Ecology Laboratory, Colorado State
University, Fort Collins, CO 80523-1499, USA

C. S. Jarnevich · T. J. Stohlgren
Fort Collins Science Center, U.S. Geological Survey,
2150 Centre Ave. Building. C, Fort Collins, CO 80526, USA

geographic space as a raster (or grid), which can be used to generate a final suitability map (Franklin 2009). This modeling approach assumes that the variables chosen for the environmental data are the most influential ones.

Based on fundamental ecological niche theory (Hutchinson 1957; MacArthur and Wilson 1963), we could create a niche model using a continuous envelope that inside contains the environmental conditions under which a species could potentially exist. Outside the envelope, conditions are unsuitable for the species to persist. We would typically expect a species to be absent outside its niche in environmental space, and the habitat quality for that species to increase moving from the boundaries toward the optimal point within the niche (Lomolino and others 2006). At least for plant habitat suitability, we might expect the shape of this envelope to be a relatively continuous curve, such as the relationship between seed germination and water stress (Maraghni and others 2010).

Many modeling packages do not provide visualization of the model in environmental space, but only in geographic space (Graham and Hijmans 2006). This can prevent experts from evaluating whether the model responds as expected based on their knowledge of the species. Additionally, many habitat suitability modeling algorithms do not allow modification of the niche model based on expert knowledge, but rely solely on auto-detected correlations between species' presence observations and environmental variables. This relies heavily on the assumption that species are inhabiting the full extent of their potential habitat, an assumption that is often violated with invasive species (Elith and Leathwick 2009). The ability to modify these models given expert information would allow researchers to refine their predictions to better reflect knowledge about what environmental factors limit a species distribution.

Our objective was to devise a new modeling approach that (1) automatically provides model optimization and allows for model editing; (2) provides a visualization of the model in environmental space; (3) uses only presence points; (4) creates the simplest (i.e., most parsimonious) model to effectively describe a target species' relationship with the environment; and (5) provides models that are continuous in environmental space. In this paper, we introduce the hyper-envelope modeling interface (HEMI), an approach that uses Bezier surfaces in a hyper-volume to model a species niche. We provide an overview of the HEMI modeling approach and then compare its performance with three other similar approaches: BioClim, BioMapper, and Maxent. We developed habitat suitability models for tamarisk (*Tamarix* sp.), a plant that is invasive to the western USA. Tamarisk is a facultative phreatophyte that desiccates flood plains and water tables (Blackburn and others 1982; Pinay and others 1992). Because it is a specialist species limited largely by water availability and

cool temperatures, tamarisk distributions have been successfully modeled at multiple spatial and temporal scales (Morisette and others 2006; Evangelista and others 2008; Jarnevich and others 2011).

Methods

Hyper-Envelope Modeling Interface (HEMI)

HEMI is similar to other habitat suitability modeling approaches in that it first extracts environmental values for each occurrence point, then creates a model within the environmental space, and finally allows the creation of predictive maps based on this model. However, HEMI creates the model in an interactive manner that allows the modeler to view and modify the process at each step. HEMI provides two-dimensional histograms of the interaction between N environmental characteristics from the area sampled for the species of interest, where N can range from two to five environmental layers. The histograms capture the available habitat, such as that for tamarisk, based on environmental characteristics such as minimum monthly temperature and mean annual precipitation (Fig 1). The histograms show the relative availability of combinations of environmental characteristics in the sampled area using a color gradient. The environmental characteristics are then overlaid with a two-dimensional histogram showing the density of occurrence points in environmental space
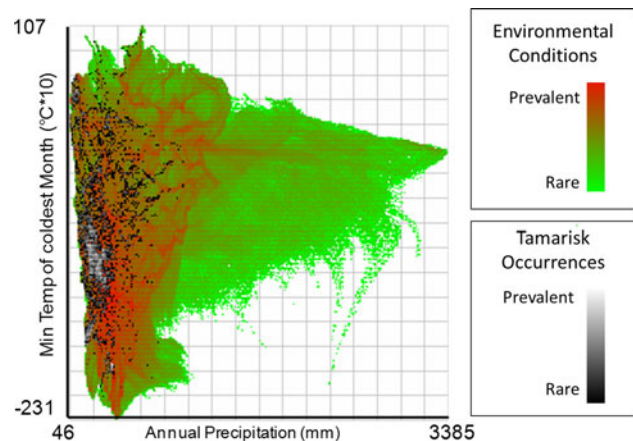

Fig. 1 A two-dimensional histogram showing the amount of area within the western USA that contains each combination of the environmental characteristics. *Green* represents that a small amount of area is available those combinations of characteristics, such as the small number of areas with over 3,000 mm of annual precipitation, and *white* and *red* indicate a large amount of area is available such as the large areas of relatively dry and cold conditions. Similarly, *white* represents environmental conditions where there were a large number of tamarisk occurrences while *black* represents a small number of occurrences

(Figs. 1; 2a). To find a rough approximation for the model, HEMI first uses a kernel density smoother with a linear kernel function to compute an overall potential density estimate for each pixel in the environmental space (Fig. 2b, c). Contour lines are added at different densities that can be specified by the user. For this study, we used contour lines at 10 and 90 % of the maximum density (Fig. 2d). The habitat suitability model is then described by Bezier curves with control points based on the contour lines (Fig. 2e). To find the final model, a function is run to search for new positions of the control points that improve the model using the area under the curve (AUC) metric (Fig. 2f). At this point, the user can create a map of the predicted habitat suitability or edit the model. Feedback to the user is provided through display of the receiver operator curve (ROC) and associated AUC values. The entire process can be completed one step at a time or automatically. HEMI was implemented in the Java programming language as an extension for the software package BlueSpray from SchoonerTurtles, Inc. (see SchoonerTurtles.com).

The key to HEMI's ability to model habitat suitability is the use of Bezier curves. Bezier curves can be used for modeling geometric surfaces in two- and three-dimensional spaces and are related to spline curves (Galvez and others 2007). Bezier curves allow control points to define the shape of the curves with greater flexibility than traditional polynomials. These curves use an independent parameter, t, which ranges from 0 to 1 as the curve moves from a starting control point to an ending control point. Additional control points are added to change the shape of the curve. A third-order Bezier curve is defined by four control points: the two end points of the curve (P_0 and P_3) and two control points that are not on the curve (P_1 and P_2) (Eq. 1). t then acts to move the coordinates from the starting control point (P_0) to the ending control point (P_3). The additional two control points that are not on the curve provide a "pull" on the curve to produce continuous, third-order curves. For two-dimensional Bezier curves, there is an equation for the x coordinate values (Eq. 2) and a separate equation for the y coordinate values (Eq. 3).
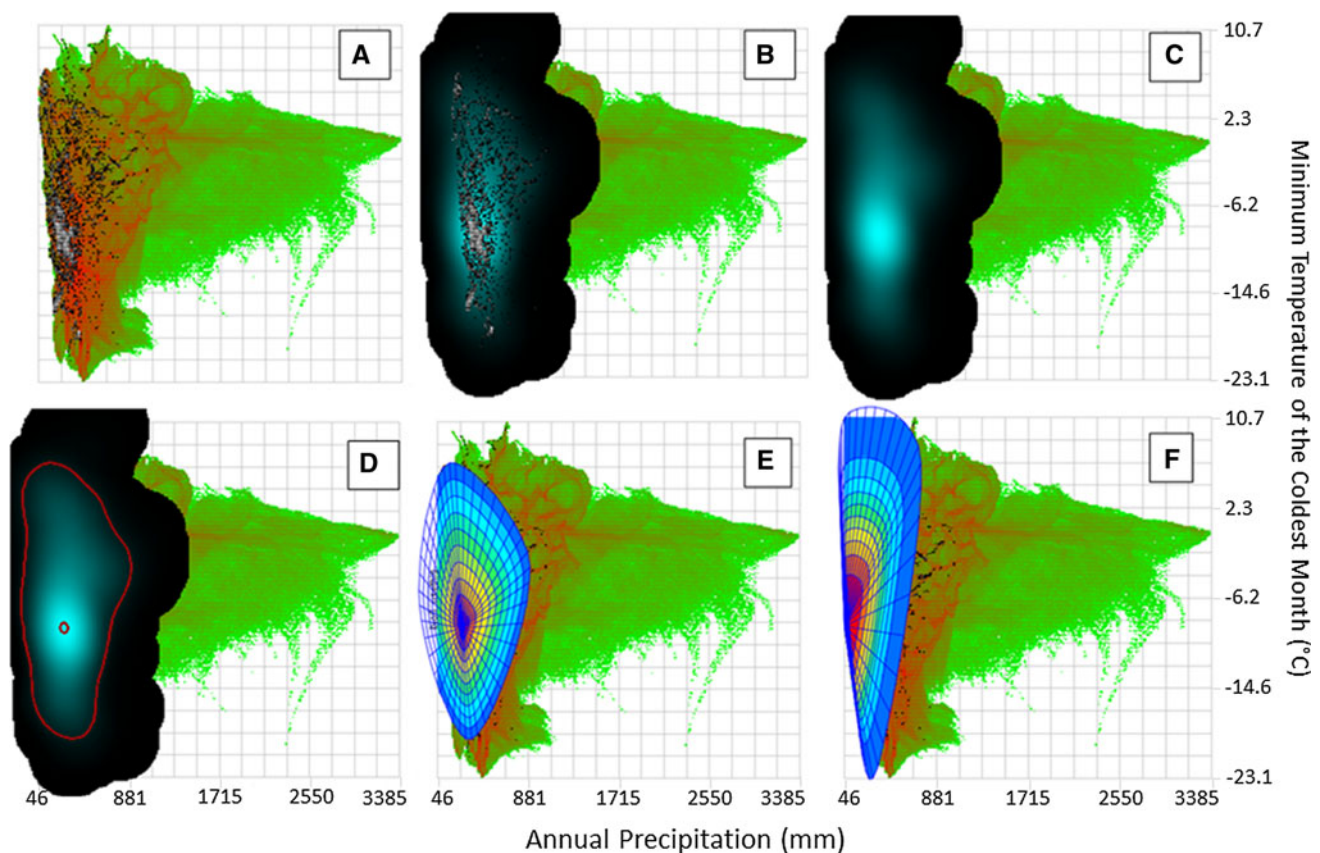

Fig. 2 Stages of developing a niche model for tamarisk in the western USA. In all graphs, the vertical axis is the minimum temperature of the coldest month, and the horizontal axis is the annual precipitation while the *red areas* in the histogram represent the most common combination of the environmental parameters with *green* being less common. **a** A two-dimensional histogram of the occurrence points with *black* being individual points and *white* being the highest density of points. **b** A density surface for the occurrences with the occurrences overlaid. **c** The density surface without occurrences. **d** Contour lines added to the density surface at 10 and 90 %. **e** A niche model with one control point at the center of the inner contour line and four points around the outer contour line. **f** An optimized model

$$B(t) = (1-t)^3 P_0 + 3(1-t)^2 t P_1 + 3(1-t)t^2 P_2 + t^3 P_3 \quad (1)$$

$$X(t) = (1-t)^3 X_0 + 3(1-t)^2 t X_1 + 3(1-t)t^2 X_2 + t^3 X_3 \quad (2)$$

$$Y(t) = (1-t)^3 Y_0 + 3(1-t)^2 t Y_1 + 3(1-t)t^2 Y_2 + t^3 Y_3 \quad (3)$$

Bezier curves within HEMI are specifically designed to provide curvilinear envelopes defined by control points on the curve. A set of Bezier curves are connected to form a complete polygon with the two end points connecting the curves (which are not really part of the curve) effectively removed, so they do not influence the envelope. Mathematically, if a Bezier curve is defined by four points, P_0 before the curve, P_1 and P_2 as the end points of the curve, and P_3 beyond the curve, then slope at P_1 is then defined by the slope of a line through P_0 and P_2 (S_1) while the slope at P_2 is defined by a line through P_1 and P_3 (S_2). Solving for the slopes and substituting them into the original Bezier function (Eq. 1) gives a new Bezier curve function (Eq. 4) with two functions for computing the third- and second-order factors (Eqs. 5 and 6). The final result is a single continuous curve with control points that can be placed automatically or manually. HEMI has the ability to produce envelopes with more than four points, but we found that the additional points did not improve the model performance.

$$B(t) = at^3 + bt^2 + S_1 t + P_1 \quad (4)$$

$$a = 2(P_1 - P_2) + S_1 + S_2 \quad (5)$$
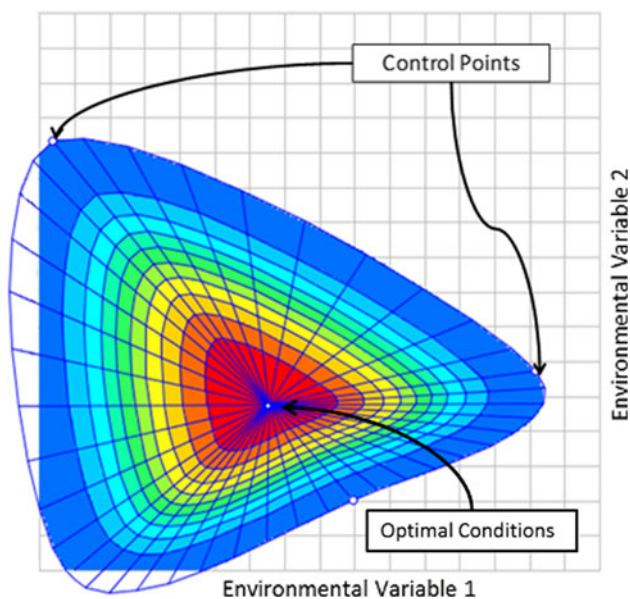
$$b = 3(P_2 - P_1) - 2S_1 - S_2 \quad (6)$$

To complete the surface, a point at the optimal environment conditions within the envelope was set to one while the boundary of the niche envelope was set to zero (Fig. 3). The transition between these two values was controlled by a traditional Bezier curve with a sigmoidal shape (Fig. 4). The shape of the curve used for this study was based on the premises that (1) changes in species response to environmental change near the optimal environmental conditions are small and (2) the area along the boundary may include microhabitats that the environmental data are unable to capture.

Habitat suitability maps in geographic space were produced by extracting the environmental predictor values for each pixel and determining where pixels fall within environmental space in relation to the niche envelope. Pixels falling outside the model would be set to zero while those at the center of the niche would be set to one. Values along the contours of the model were given values based on the transition curve (Fig. 4).

Visualizing the Niche

For habitat suitability models, the ability to view the environmental niche defined by the model is important for model generation and evaluation (Elith and Graham 2009). HEMI and BioMapper provide the ability to view the niche model in environmental space. To view the environmental niche for the other modeling methods, we created synthetic environmental layers based on the full range of values of the environmental layer. One layer changed vertically for one predictor variable while the other was changing horizontally for the second predictor variable. The models were
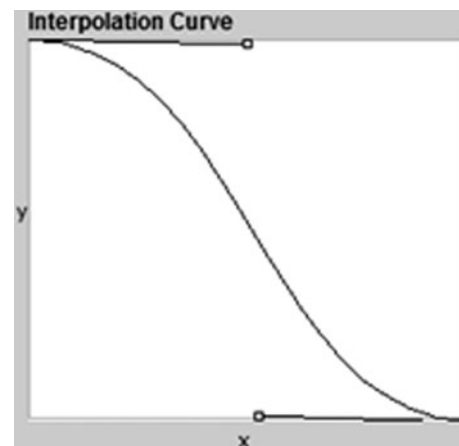

Fig. 3 A sample niche envelope in environmental space where the *blue areas* are marginal habitat and the *red area* represents optimal habitat. Four control points at the edges control the shape of the model. These control points and the center point can be moved by the modeler to alter the shape of the environmental space model

Fig. 4 The interpolation curve between the center of the niche and the boundary is controlled by traditional Bezier control points. The starting point on the *left* represent the values at the optimal habitat, or 1, with the ending point on the *right* representing the boundary of habitat, or zero. The two additional points control the shape of the curve but were fixed at $y = 1$, $x = 0.5$ for the first point and $y = 1$, $x = 0.5$ for the second point for this study

applied to the synthetic layers to generate a visual of the niche model in environmental space (Elith and Graham 2009).

Study Sites

We tested HEMI, BioClim, BioMapper, and Maxent within the western USA where tamarisk habitat suitability has been modeled successfully in the past (Jarnevich and others 2011). We defined the western USA to include the states from North Dakota south to Texas and west to the Pacific Ocean. Tamarisk infests over one million acres in North America (Brotherson 1987) and is currently the third most frequently occurring woody riparian plant in the western USA (Friedman and others 2005).

Occurrence Data

Tamarisk presence points were obtained from multiple independent sources including data available from the National Institute of Invasive Species Science (www.niiss.org). We used 11,159 presence points that were freely available and are the same as those used by Jarnevich in 2001, with the exception of the Kerns dataset, (Jarnevich and others 2011), randomly divided into 7,814 training points and 3,345 test points.

Environmental Variables

To develop our models, we selected two environmental predictors identified as significant in previous studies: annual temperature and precipitation (Friedman and others 2005; Kerns and others 2009; Jarnevich and others 2011). While additional variables may have provided a better model, the goal of this study was to demonstrate the abilities of HEMI rather than provide the optimal model. The uncertainties within these predictors may affect the outcome of the model (Soria-Auza and others 2010), but since the primary purpose of this study was to compare outputs from different methods using the same predictor variables, these uncertainties could be ignored. We used annual precipitation and minimum temperature of the coldest month acquired from the WorldClim database v1.4 (Hijmans and Graham 2006).

Modeling Algorithms

We compared the HEMI model algorithm with three commonly used habitat suitability model algorithms: BioClim, BioMapper, and Maxent. All these model techniques are correlative and do not require absence data. The BioClim model (Busby 1986) operates by placing a series of simple thresholds around each of the presence points for each environmental variable to define the environmental space occupied by a species, forming a rectangle in N-dimensional space. Each pixel in the output layer is ranked according to where its environmental characteristics fall in relation to this hyper-rectangle to produce a habitat suitability map. We used DIVA-GIS (www.diva-gis.org) to implement BioClim (www.diva-gis.org/docs/DIVA-GIS5_manual.pdf). We used BioClim's default percentile (i.e., 95 %) to develop our models, eliminating the most environmentally extreme 5 % of the presence points.

BioMapper (Hirzel and others 2002) uses ecological niche factor analysis (ENFA) to develop a model. This approach compares the mean of the environmental variables at the presence points to the mean across the sampled region, and produces marginality and specialization values to define a species niche (Hirzel and others 2002; Hirzel and Arlettaz 2003). Marginality indicates how different the environmental conditions are where the species occurs compared to the average environmental conditions available, while specialization values indicate how restricted the environmental conditions are where the species occurs compared to the range of environmental conditions available. We used the default median algorithm. The geometric mean algorithm is more similar to how HEMI calculates the realized niche, but could not be used as the software crashed repeatedly due to the large dataset.

Maxent uses maximum entropy modeling and background points representing the available environment to fit a combination of linear and second-order polynomials to a dataset (Phillips and others 2006). We used version 3.3.3e with the default settings except we increased the maximum iterations to 5,000 to allow it to reach convergence and restricted background point selection to US counties that contained tamarisk presence locations to decrease sampling bias effects. We also used the multidimensional environmental similarity surface (MESS) analysis in Maxent to highlight locations with novel environments compared to sampled environments used to develop the model.

Evaluation

To compute the same measures of model predictive capability across all the modeling approaches examined, we computed AUC from the predicted habitat map and a set of occurrence points. AUC is typically computed by moving a threshold from one to zero and measuring the proportion of points that are true presences against the proportion of points that are true absences. Since we did not use absence points, we computed the AUC based on the proportion of the sample area that was contained within the environmental envelope versus the number of occurrences that were contained within the environmental envelope (Fig. 5). This calculation is performed directly on a histogram of the
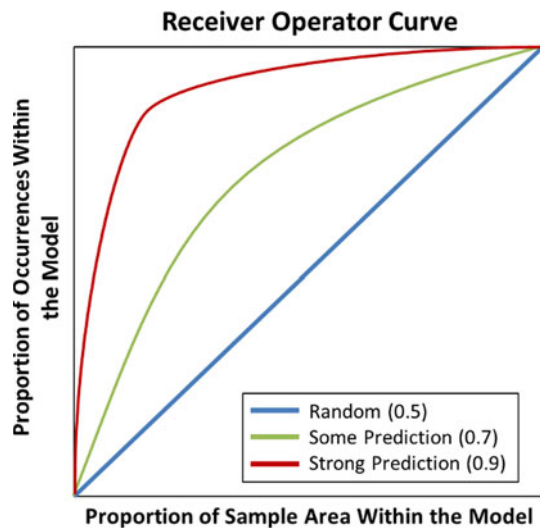
Fig. 5 An illustration of how area under the curve (AUC) was computed. AUC is computed by finding the area within the model as a threshold is changed from 1 (very little of the model) to near 0 (the entire model). The proportion of the number of occurrences within the modeled area is plotted versus the proportion of the sample area within the modeled area

environmental factors, using the number of occurrences found for each unique combination of environmental factors against the number of locations in the sample area that contain the same environmental factors. This keeps the model from being biased toward environmental conditions that are widespread versus those that are rare.

We withheld a random 30 % of the data for testing and used the other 70 % to train the models. AUCs were computed for both the train and test data. We ran additional models using HEMI to see whether manually editing the model in environmental space improved model results.

We also used an independent dataset of estimated acreage of tamarisk at the quarter quad level from a survey of county weed coordinators by the Western Weeds Coordinating Committee to evaluate the models using the PresenceAbsence library in R v2.14.2. To evaluate model performance, we calculated the sensitivity equals specificity threshold value (where the chance of correctly predicting an occurrence is the same as the chance of correctly predicting an absence) to calculate binary prediction maps (suitable and unsuitable habitat). These were used to calculate percent correctly classified, sensitivity, specificity, and Cohen's kappa. We also calculated the AUC vales for the independent dataset.

Akaike information criterion (AIC) was computed for each model against the training dataset. This was completed by treating each of the model outputs as a probability surface and multiplying the modeled probability of each occurrence together to obtain an overall likelihood (Warren and Seifert 2011). AIC was then computed in the traditional manner. These calculations were executed in

BlueSpray. We should note that while the same dataset was used to compute AIC, Maxent uses an internal random selection of background points, which will change the AIC values for each model run. We also calculated a Spearman's rank correlation coefficient in BlueSpray to evaluate the level of agreement between each pair of predicted suitabilities.

Results

HEMI Performance

HEMI successfully created a model for habitat suitability of tamarisk using a fundamentally different modeling approach based on Bezier curves. The environmental space available in the western USA generally included areas with relatively low precipitation and moderate low temperature environmental conditions (Fig. 1); tamarisk appears to prefer drier regions of the western USA. The occurrence points overlaying this environmental space show first that the occurrence data are not continuous, with one large concentration of occurrences at about a minimum temperature of $-20\,^{\circ}C$ and two smaller concentrations at lower temperatures (Fig. 1). That the occurrence data are non-continuous could be due to environmental conditions not being prevalent, tamarisk not yet invading these areas, or certain environmental conditions not being sampled as intensively as others.

The four models covered similar areas of environmental space, but appear different when examined in the environmental space (Fig. 6). The automated HEMI model produced geographic maps that visually appear similar to the existing approaches (Fig. 7). Maxent produced the highest test AUC (0.77), while HEMI had the second highest test AUC (0.75) followed by BioMapper (0.74) and BioClim (0.67). The manually created HEMI model included all of the available occurrence points in environmental space and produced a map that predicted tamarisk over a wider area (test AUC = 0.73; Table 1; Fig. 8). BioClim and HEMI were the most parsimonious models. HEMI required one control point for the center of the niche and four control points describing the boundary of the niche. Since there were two environmental predictor layers, two values were required for each point resulting in ten parameters to fully represent the niche model. The Maxent output file containing the coefficients included 182 parameters. BioClim uses a minimum, 5, 95 %, and a maximum parameter for each environmental layer giving eight parameters for the two environmental layers used in this study (Busby 1986). BioMapper uses a grid of 25×25 values for the model resulting in 625 parameters (Hirzel and others 2002).

Fig. 6 Predicted habitat suitability (displayed in terms of potential density in environmental space) for tamarisk within the western USA for **a** HEMI, **b** BioMapper, **c** BioClim, and **d** Maxent models
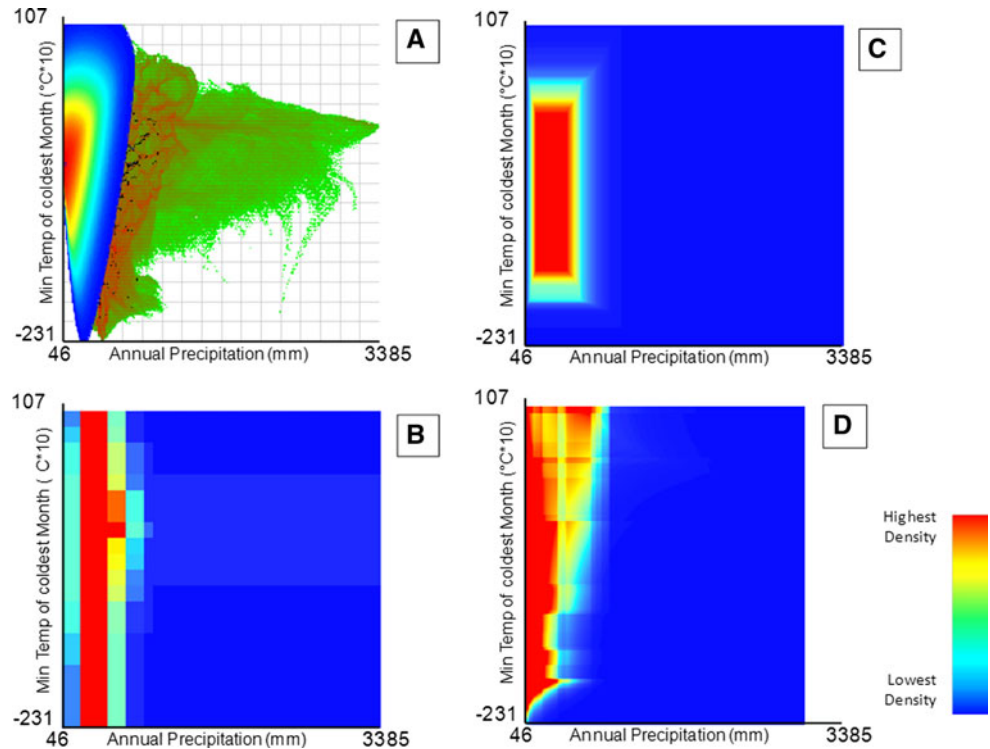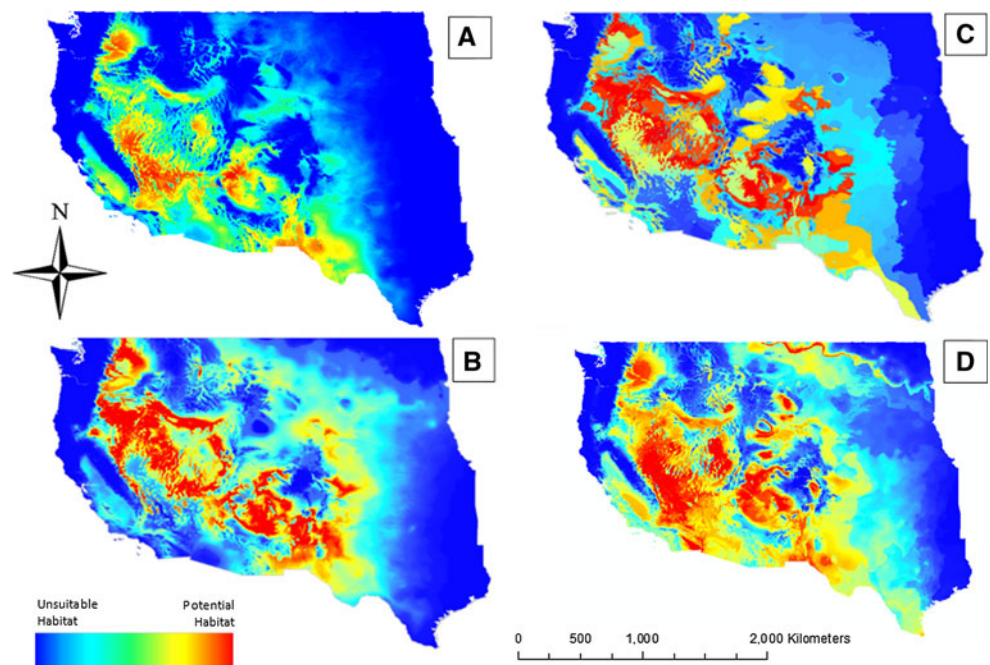

Fig. 7 Predicted suitable habitat (displayed in terms of potential habitat suitability) for tamarisk in the western USA with **a** HEMI, **b** BioMapper, **c** BioClim, and **d** Maxent modeling approaches

The independent quarter quad data evaluation produced AUC values somewhat lower than predicted from the training and test datasets, but following the same trend as before (Table 2). The automated HEMI model had the lowest AIC value (98190) followed by Maxent (102297) and the manually edited HEMI model (102457) (Table 1).

In terms of the Spearman's correlation evaluation, BioClim and BioMapper had the most similar predictions ($P = 0.82$) and Maxent and BioClim had the most dissimilar predictions. HEMI's predictions agreed the most with BioMapper ($P = 0.81$) and agreed the least with Maxent ($P = 0.75$; Fig. 9).

Discussion

The HEMI was able to model potential species habitat by using relatively simple Bezier functions to create environmental niche envelopes. This framework allowed explicit visualization and modification of the species' environmental niche space. All of the models, except BioClim, were within 4 % of each other when AUC was used as a comparative metric. HEMI produced the lowest AIC values and was able to achieve this with an order-of-magnitude fewer parameters than Maxent and BioMapper (Table 1). Since all of the performance measures were within 5 % of each other for HEMI, Maxent, and BioMapper, HEMI's greater parsimony may be its primary advantage.

The performance of HEMI compared to more complicated modeling algorithms is surprising and bears further examination. HEMI uses a series of Bezier functions to describe the niche space of a species and produce a habitat suitability model. Bezier curves provide continuous model surfaces with an order-of-magnitude fewer parameters than some other methods. However, since AUC is computed without explicitly accounting for the number of parameters

in the model, it appears that simple curves may actually reflect a more accurate way to describe species distribution in environmental niche space. Our results highlight the difficulty in comparing existing presence-only habitat suitability models and suggest that perhaps new metrics are needed for presence-only modeling algorithms that consider model fit, number of parameters (parsimony), niche theory, and even expert knowledge in model evaluation.

The BioClim's "box-model" approach is apparent in the visualization of the niche (Fig. 6). This approach included only eight parameters. BioMapper produced a relatively continuous model but used 625 parameters. Both BioMapper and HEMI produced niche models with the relatively continuous surfaces that might be expected from ecological theory for plant species (Fig. 6).

The Maxent model produced a higher AUC value (though HEMI was within 0.02 and BioMapper within 0.03), and its map more closely modeled the data provided, including the Missouri River corridor into Montana as potential tamarisk habitat. This result, combined with the larger number of parameters for Maxent, may be an indication of over-fitting the data, especially when using the default settings (Anderson and Gonzalez 2011). HEMI provided a simpler model with fewer parameters but may be under-fitting the data. Both Maxent and HEMI allow the

Table 1 Area under the curve (AUC) values (i.e., training and test datasets), the number of parameters, and estimated AIC values for each model

Method	Train AUC	Number of parameters	Test AUC	AIC
HEMI automatic	0.75	10	0.76	98,190
HEMI edited	0.73	10	0.73	102,457
BioClim	0.67	8	0.68	103,752
BioMapper median	0.74	625	0.72	103,606
Maxent	0.77	182	0.78	102,297

AIC values were completed based on the test dataset

Table 2 Quarter quad tamarisk survey evaluation of models including the calculated sensitivity equals specificity threshold used to calculate the other metrics, the percent correctly classified (PCC), sensitivity, specificity, Cohen's kappa, and area under the curve (AUC)

Method	PCC	Sensitivity	Specificity	Kappa	AUC
HEMI automatic	66	0.66	0.66	0.33	0.70
HEMI edited	63	0.64	0.62	0.27	0.66
BioClim	58	0.59	0.57	0.16	0.62
BioMapper median	64	0.64	0.64	0.28	0.67
Maxent	65	0.65	0.64	0.30	0.70

Fig. 8 The edited model from HEMI showing the model in environmental space on the *left* and geographic space on the *right*
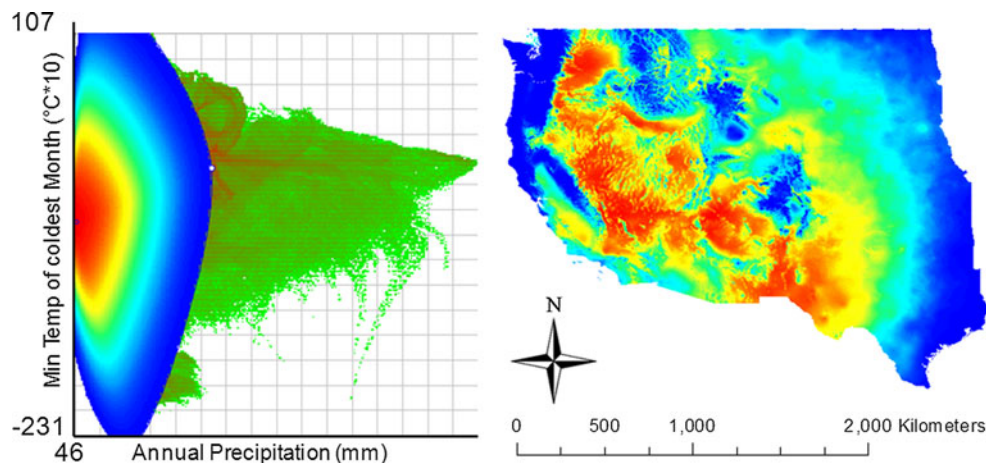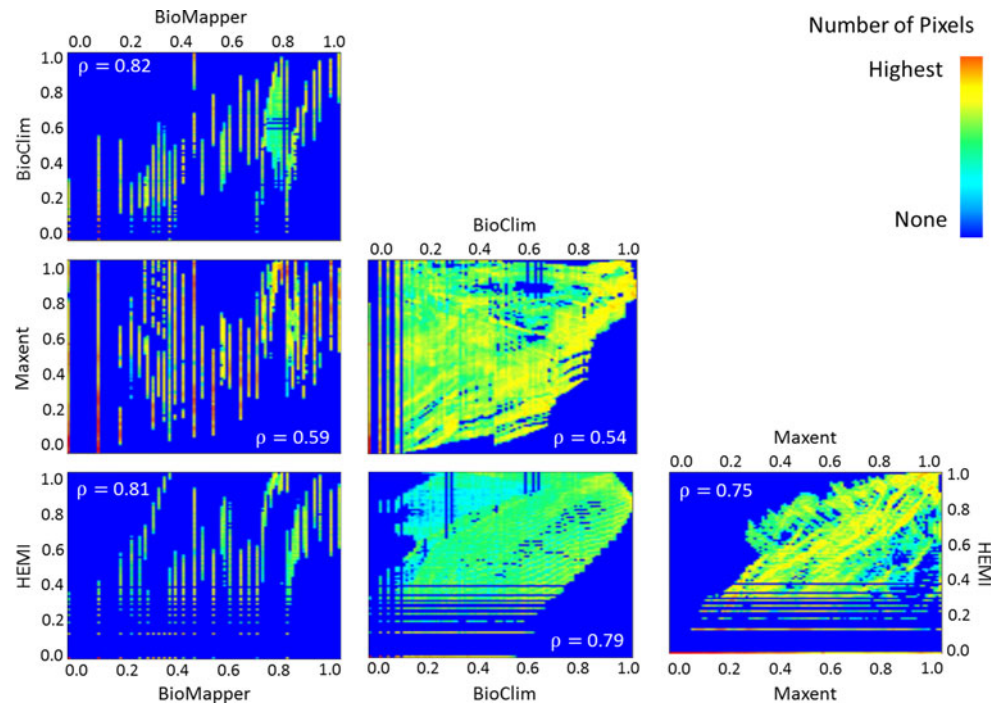
Fig. 9 Two-dimensional histograms of correlation between pixels in each of the model outputs and Spearman's rank correlation coefficients

ability to change the complexity of the envelopes created, and additional research into the effects of varying the settings for both approaches is warranted.

The manually created HEMI model has a larger extent in both environmental and geographic space and a slightly lower AUC of 0.73 compared to the automated model's AUC of 0.75. The manually created model may over-predict the potential habitat of tamarisk. However, since tamarisk is a riparian plant and our environmental layers are at 1 km, there may be areas where tamarisk is surviving in small drainages that are poorly represented in the environmental data. Over-predicting the habitat can help guide field monitoring efforts to ensure that areas where invasive species may occur in microhabitats are monitored. Regardless, HEMI provides the user with control on how well to fit the data to modify the model based on expert knowledge and the needs of their application. As an additional note, HEMI is an interactive program allowing the user to modify models in environmental space and then visualize a potential habitat in geographic space in a few seconds.

The question of whether these types of models follow what would be expected from the physiological requirements of a species could only be explained by large-scale experiments to determine the environmental needs of each species being modeled. The logistics of these experiments, including the varieties of genotypes that would be required, make these types of experiments impossible. There is also a question as to whether these experiments would capture the complexities of a natural landscape. Instead, we are using field data to develop models of the environmental needs of species. The field data will have its own set of problems including uncertainty and bias. We feel the best we can do at this stage in the evolution of these models is to strive to develop approaches that create models that match current ecological theory. The visualizations provided by HEMI allowed us to gain additional insight into the environmental needs of tamarisk and also raised questions about our field data.

HEMI was built to create simple models using Bezier surfaces with control points. This limits HEMI to continuous environmental variables such as temperature and precipitation, but does not allow for the use of categorical variables such as soil type. This could be overcome by modeling each segment of a discontinuous variable and then combining results. While this study used only two environmental variables, HEMI supports up to five. Future extensions to HEMI are planned to allow for the incorporation of uncertainty from occurrences and environmental layers and addition of other algorithms to determine the optimal niche model. While this is the first introduction of HEMI, we hope to model a variety of species in different geographic regions with different datasets to further investigate its applicability.

or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government. We also wish to thank Tom Hobbs for his instruction on the goals and methods for ecological modeling. HEMI was created within the BlueSpray application which is the property of SchoonerTurtles, Inc. To obtain a copy of BlueSpray, please contact Jim Graham at jimg@schoonerturtles.com.

References

Anderson RP, Gonzalez I (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with Maxent. Ecol Model 222(15):2796–2811. doi:10.1016/j.ecolmodel.2011.04.011

Blackburn WH, Knight RW, Schuster JL (1982) Saltcedar Influence on sedimentation in the Brazos River. J Soil Water Conserv 37(5):298–301

Brotherson JD, Dean F (1987) Tamarix: impacts of a successful weed. Rangelands 9:110–112

Busby JR (1986) A biogeoclimatic analysis of Nothofagus cunninghamii (hook) oerst in southeastern Australia. Aust J Ecol 11(1):1–7

Elith J, Graham CH (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. Ecography 32(1):66–77. doi:10.1111/j.1600-0587.2008.05505.x

Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. Annu Rev Ecol Evol Syst 40:677–697

Evangelista PH, Kumar S, Stohlgren TJ, Jarnevich CS, Crall AW, Norman JB, Barnett DT (2008) Modelling invasion for a habitat generalist and a specialist plant species. Divers Distrib 14(5):808–817

Franklin J (2009) Mapping species distributions: spatial inference and prediction. Cambridge University Press, Cambridge

Friedman JM, Auble GT, Shafroth PB, Scott ML, Merigliano MF, Preehling MD, Griffin EK (2005) Dominance of non-native riparian trees in western USA. Biol Invasions 7(4):747–751. doi:10.1007/s10530-004-5849-z

Fuller T, Morton DP, Sarkar S (2008) Incorporating uncertainty about species' potential distributions under climate change into the selection of conservation areas with a case study from the Arctic Coastal Plain of Alaska. Biol Conserv 141(6):1547–1559. doi:10.1016/j.biocon.2008.03.021

Galvez A, Iglesias A, Cobo A, Puig-Pey J, Espinola J (2007) Bezier curve and surface fitting of 3D point clouds through genetic algorithms, functional networks and least-squares approximation. In: Gervasi O, Gavrilova ML (eds) Proceedings of computational science and its applications—ICCSA 2007, Pt 2, vol 4706. Lecture Notes in Computer Science. Springer, Berlin, pp 680–693

Graham CH, Hijmans RJ (2006) A comparison of methods for mapping species ranges and species richness. Glob Ecol Biogeogr 15(6):578–587. doi:10.1111/j.1466-822x.2006.00257.x

Guisan A, Broennimann O, Engler R, Vust M, Yoccoz NG, Lehmann A, Zimmermann NE (2006) Using niche-based models to improve the sampling of rare species. Conserv Biol 20(2):501–511. doi:10.1111/j.1523-1739.2006.00354.x

Hijmans RJ, Graham CH (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. Glob Change Biol 12(12):2272–2281. doi:10.1111/j.1365-2486.2006.01256.x

Hinojosa-Diaz IA, Feria-Arroyo TP, Engel MS (2009) Potential distribution of orchid bees outside their native range: the cases of Eulaema polychroma (Mocsary) and Euglossa viridissima (Friese) in the USA (Hymenoptera: Apidae). Divers Distrib 15(3):421–428. doi:10.1111/j.1472-4642.2008.00549.x

Hirzel AH, Arlettaz R (2003) Modeling habitat suitability for complex species distributions by environmental-distance geometric mean. Environ Manage 32(5):614–623

Hirzel AH, Hausser J, Chessel D, Perrin N (2002) Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? Ecology 83(7):2027–2036

Hutchinson GE (1957) Population studies—animal ecology and demography—concluding remarks. Cold Spring Harbor Symp Quant Biol 22:415–427

Jarnevich CS, Evangelista P, Stohlgren TJ, Morisette J (2011) Improving national-scale invasion maps: tamarisk in the Western United States. Western North Am Naturalist 71(2):164–175

Kerns BK, Naylor BJ, Buonopane M, Parks CG, Rogers B (2009) Modeling tamarisk (Tamarix spp.) habitat and climate change effects in the northwestern United States. Invasive Plant Sci Manag 2(3):200–215

Lomolino MV, Riddle BR, Brown JH (2006) Biogeography, 3rd edn. Sinauer Associates, Sunderland

MacArthur RH, Wilson EO (1963) Equilibrium-theory of insular zoogeography. Evolution 17(4):373

Maraghni M, Gorai M, Neffati M (2010) Seed germination at different temperatures and water stress levels, and seedling emergence from different depths of Ziziphus lotus. S Afr J Bot 76(3):453–459. doi:10.1016/j.sajb.2010.02.092

Morisette JT, Jarnevich CS, Ullah A, Cai WJ, Pedelty JA, Gentle JE, Stohlgren TJ, Schnase JL (2006) A tamarisk habitat suitability map for the continental United States. Front Ecol Environ 4(1):11–17

Penman TD, Pike DA, Webb JK, Shine R (2010) Predicting the impact of climate change on Australia's most endangered snake, Hoplocephalus bungaroides. Divers Distrib 16(1):109–118. doi:10.1111/j.1472-4642.2009.00619.x

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model 190(3–4):231–259. doi:10.1016/j.ecolmodel.2005.03.026

Pinay G, Fabre A, Vervier P, Gazelle F (1992) Control of C, N, P distribution in soils of riparian forests. Landscape Ecol 6:121–132

Soria-Auza RW, Kessler M, Bach K, Barajas-Barbosa PM, Lehnert M, Herzog SK, Bohner J (2010) Impact of the quality of climate models for modelling species occurrences in countries with poor climatic documentation: a case study from Bolivia. Ecol Model 221(8):1221–1229. doi:10.1016/j.ecolmodel.2010.01.004

Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecol Appl 21(2):335–342. doi:10.1890/10-1171.1