Research Article

Characterizing the suitable habitat of *Miconia calvescens* in the East Maui Watershed

Niels Jorgensen¹, James Leary²*, Mark Renz¹ and Brooke Mahnken³

¹University of Wisconsin-Madison, 1575 Linden Drive, Madison, WI 53706, USA
²University of Florida, 7922 NW 71st Street, Gainesville, Florida 32653, USA
³Maui Invasive Species Committee, P.O. Box 983, Makawao, Hawaii 96768, USA

Author e-mails: njorgensen@wisc.edu (NL), learyj@ufl.edu (JL), mrenz@wisc.edu (MR), mahnken@hawaii.edu (BM)

*Corresponding author

Abstract

The East Maui Watershed (EMW) is a > 60,000-ha forested watershed with wide temperature and precipitation gradients being invaded by miconia (*Miconia calvescens* DC.). Current miconia management efforts focus on protecting important watershed and critical habitat areas from miconia invasion. Herein, we report on a miconia species distribution model to predict unoccupied areas that may be still vulnerable to invasion. This suitable habitat model was developed from an ensemble of five algorithms associating five physical features of EMW with miconia occurrence data from a 26-yr management history (1991–2016; n = 114,953). All of the algorithms performed well based on model evaluation statistics (e.g. AUC ≥ 0.83; TSS ≥ 0.36). Elevation, slope and rainfall were consistently important predictors, while aspect indices were non-contributors. The binary ensemble model suggests a total of ~ 56.9% of the area of interest is susceptible to invasion by miconia. An independent dataset collected in 2017–2018 (n = 5,222) was used to field validate the ensemble habitat suitability model (EHSM) and found that the model could correctly predict suitable habitat 94% of the time. All five of the model algorithms were updated using this new management data, and the predicted suitable area decreased 2.3%. While binary models are useful for risk assessment, the classification of an area as suitable or not suitable has limitations for land managers adopting for management activities. Utilizing the mean weighted consensus probability surface representation of the model allows for more scrutiny of potential suitable habitat. We suggest using this approach when planning future monitoring efforts, especially if specific areas have a higher prioritization for conservation than others.

Key words: boosted regression trees, generalized linear models, multivariate adaptive regression splines, MaxEnt, random forests, Ensemble Habitat Suitability Model

Introduction

Naturalization of an introduced species is achieved when abiotic and biotic barriers to survival and reproduction are overcome, whereas a species becomes invasive when barriers to dispersal are overcome (Colautti and MacIsaac 2004; Richardson et al. 2000; Theoharides and Dukes 2007). While impacts of invasive species can vary, the most concerning are species that modify the landscape and have environmental (Crooks 2002;
D’Antonio and Vitousek 1992; Mack and D’Antonio 2003; Vilà et al. 2011) and economic consequences (Vitousek et al. 1997; Levine and D’Antonio 2003; Mack et al. 2000; Pimentel et al. 2005; Wilcove et al. 1998). While these species are a global issue, they disproportionately affect small islands (Wilcove et al. 1998). Hawaii’s geographical isolation, mild sub-tropical climate and evolved endemism have likewise made these ecosystems highly vulnerable to non-native invasions (Gillespie et al. 2008).

Miconia (*Miconia calvescens* DC.) is a mid-story canopy woody species native to South and Central America that is invading Maui, Hawaii. It is classified as one of “100 of the World’s Worst Invasive Alien Species” (Lowe et al. 2000) as it is a forest ecosystem modifier to tropical regions of the Pacific Rim (Medeiros et al. 1997; Meyer and Florence 1996; Nanko et al. 2015). Miconia displaces native plant communities (Meyer and Florence 1996) which results in accelerated erosion (Nanko et al. 2013, 2015; Giambelluca et al. 2009) and increased potential for soil sloughing (Gagné et al. 1992).

Miconia was introduced to Hana, Maui (20.7575°N; 155.9884°W) in the early 1970s as a botanical specimen (i.e. founder population) (Gagné et al. 1992) and only recognized two decades later to be a major invader of the north and east (windward) watersheds of Haleakala Volcano, also known as the East Maui Watershed (EMW) (Medeiros et al. 1997). Control efforts were initiated in 1991 (Gagné et al. 1992) and have continued with > 1.5M miconia eliminated to date (Leary et al. 2018). Despite these efforts, eradication of the population has not been achieved with ~ 30% of the EMW invaded by miconia.

Due to the breadth of the invasion and lack of matching resources to counter, eradication of miconia from the EMW is not feasible. The current strategy instead concentrates limited resources on the invasion front to mitigate encroachment into watershed and critical habitat areas (Leary et al. 2018; Meyer et al. 2011). In these areas, there is a lack of information to predict if and where management is required. For instance, unnecessary deployment to areas where miconia may not be able to establish, or the alternative of not searching areas that are unknowingly vulnerable to miconia invasion, may reduce success of this effort. Spatial assignment of tactics is inherently key to broad-scale strategy for management of a dynamic species invasion on a landscape level (Grice et al. 2011). Unfortunately, this can be time consuming as Leary et al. (2018) estimated up to 90% of effort can be spent searching for 1% of the population.

Management can be optimized by targeting areas with a high likelihood of species occurrence (Hauser et al. 2016). Habitat suitability models (HSMs) are created from species occurrence data and can inform management efforts to areas predicted to be high suitability (Crall et al. 2013; Wang et al. 2014). These spatial outputs are created by associating these occurrences with predictors of the physical environment to explain drivers of suitability...
for a species to invade, and then extrapolating to broader areas of interest (Elith and Leathwick 2009). There are many approaches to modeling suitable habitat, and some algorithms will have greater predictive capacity than others given the same inputs (Elith and Graham 2009). Combining model outputs into an ensemble has been found to enhance predictive ability by minimizing the shortcomings of any one approach (Stohlgren et al. 2010).

Here we present the development of an ensemble of habitat suitability models (EHSM) of five algorithm outputs associating features of the physical environment and climatic factors of the EMW with miconia occurrence data from a 28-yr management history (1991–2018). These probabilistic models allow us to explore variable influence on the overall predictions and can be represented visually as probability density functions or by splitting the model into a binary representation. Our objectives were to: (1) evaluate the response of each of the predictor layers in explaining suitable habitat for miconia in the EMW; (2) field validate the binary ensemble model’s ability to correctly classify suitable habitat; and, (3) use a mean weighted probability representation to suggest prioritization of future monitoring and management.

**Materials and methods**

**Study location**

Research was conducted on the windward slope of the 60,000-ha watershed of the Haleakala Volcano in Maui Hawaii. This watershed provides critical habitat to over one hundred threatened and endangered species and collects over 9 billion liters of precipitation daily (Anonymous 2019; Shade 1999). This study focused on 49,136-ha (Figure 1a) confined to the NW slope of Maliko watershed wrapping around to the SE slope of Manawainui watershed, with an upper elevation of 1220 m. The AOI is climatically diverse, with wide temperature lapse rate (mean annual, mid-day 14.9–26.2 °C) and orographic precipitation (652–9621 mm) gradients (Giambelluca et al. 2014; Frazier et al. 2016).

**Miconia occurrence data**

Occurrence records (Figure 1a; Supplementary material Table S1 for availability) were derived from a 26-yr history of miconia management in the EMW with time-stamped, georeferenced locations of miconia eliminated by ground and aerial operations from 1991 to 2016 (n = 114,953 representing 1,516,727 miconia individuals) (Leary et al. 2018). Data with poor spatial accuracy were discarded from the analysis to avoid complications in the model predictions (Graham et al. 2008). Aerial operations deploying Herbicide Ballistic Technology (HBT; methods described by Leary et al. 2014) were conducted in 2017–2018 eliminating more miconia (n = 5,222; Figure 1b) and used as a post-validation test of the ensemble model (see details below).
Characterizing suitable habitat of *M. calvescens* in the EMW


**Figure 1.** Occurrence data from 1991–2016 (a) used to develop habitat suitability models for *Miconia calvescens* in the East Maui Watershed (n = 114,953). The area of interest (AOI) is delineated by the white outline. Additional occurrence data from 2017–2018 (n = 5,222) (b) were used to update the initial models.

**Habitat suitability models**

All habitat suitability models were developed using the Software for Assisted Habitat Modeling (SAHM) within the Vistrails scientific workflow management system (Freire et al. 2012; Morisette et al. 2013). SAHM is open-source software that builds spatial models using algorithms by associating occurrence data with predictor layer inputs. Background points
Characterizing suitable habitat of *M. calvescens* in the EMW

Table 1. Candidate predictor layers and their native spatial resolution used for modeling suitable habitat of *Miconia calvescens* on the island of Maui, HI. Temperature was ultimately not used for model development due to high correlation with elevation. All data layers available through University of Hawaii – Mānoa. Justification for usage of each layer is based on previous studies modeling suitable habitat for miconia: 1Pouteau et al. 2011; 2Vorsino et al. 2014; 3Libeau et al. 2019.

<table>
<thead>
<tr>
<th>Predictor Layer</th>
<th>Spatial Resolution</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Annual Precipitation (mm)</td>
<td>250 m</td>
<td>Frazier et al. 2016</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>250 m</td>
<td>Frazier et al. 2016</td>
</tr>
<tr>
<td>Elevation (from DEM) (m)</td>
<td>10 m</td>
<td>University of Hawaii – Mānoa</td>
</tr>
<tr>
<td>Slope (derived from DEM) (°)</td>
<td>10 m</td>
<td>University of Hawaii – Mānoa</td>
</tr>
<tr>
<td>% Northness (cosine of Aspect)</td>
<td>10 m</td>
<td>University of Hawaii – Mānoa</td>
</tr>
<tr>
<td>% Eastness (sine of Aspect)</td>
<td>10 m</td>
<td>University of Hawaii – Mānoa</td>
</tr>
</tbody>
</table>

were derived from randomly selected points (10% of the total 1,131,840 points in the dataset, or 41,828 points) along recorded surveillance routes of helicopter search operations as described by Leary et al. (2014). These data represent absence at a given time period, but do not necessarily constitute true absence; an approach referred to as presence – targeted background habitat modeling (Phillips et al. 2009). Miconia monitoring efforts occurred over much of the AOI since the inception of HBT, thus the pseudo-absence points were geographically widespread across the EMW. We eliminated any duplicate (i.e. observations taken at different times) occurrence records, and SAHM further thinned the dataset by eliminating multiple occurrences in each pixel for a final 46,291 (~ 40% of the total) occurrences. Similarly, SAHM gives preference to presence over absence points (a cell cannot be both a presence and an absence), which further thinned our pseudo-absence dataset from an original 41,828 points to 16,794 used by the algorithms for model development in SAHM.

Six predictor layers of the physical environment were considered for model development based on prior efforts to model miconia (Table 1). Predictor layers related to climate were obtained from the University of Hawaii-Mānoa and were derived from the work by Frazier et al. (2016). Topographic predictors were calculated from a 10 m digital elevation model (DEM) using the Surface Toolbox in ArcMap (version 10.5; ESRI Co., Redlands, CA). All predictor layers were projected (NAD 1983 HARN), aggregated (mean values), resampled (bilinear interpolation), and clipped to the island of Maui at a final 10m spatial resolution in the SAHM “PARC” module. The PARC module in SAHM requires a template layer that defines the spatial reference system, the spatial extent and spatial resolution of the final model (Morisette et al. 2013). Correlation between the predictor layers was assessed in SAHM in the “CovariateCorrelationAndSelection” module, which generated a matrix of correlation coefficients from three different indices (Pearson, Spearman, and Kendall). As expected, temperature and elevation (lapse rate = 0.6 °C/100 m; Baruch and Goldstein 1999) were highly correlated (r = 0.96). Elevation was chosen for the model, because it had finer native spatial resolution than temperature (10 m vs. 250 m) while
also explaining more deviance (see Morisette et al. 2013). No other predictors were highly correlated ($r < |0.5|$) and thus used in model development. A 10-fold cross-validation (Hijmans 2012) was employed to test each model by randomly withholding 10% of the data and testing it against the trained model, iterating this process 10 times, each with a different subset of the data.

Five different algorithms were selected in SAHM to develop models, which included boosted regression trees (BRT), generalized linear models (GLM), multivariate adaptive regression splines (MARS), maximum entropy (MaxEnt), and random forests (RF). Each algorithm has an intrinsic method for developing probability density functions, resulting in variations of model outputs under the same conditions; thus, avoiding random or bias preference for one model over the others (Qiao et al. 2015). The performance of each output was evaluated with probability threshold independent (e.g. AUC) and dependent metrics (e.g. sensitivity, specificity). Models with area under the curve (AUC) values > 0.5 for the receiver operating characteristic (ROC) were considered better than random chance of accurately discriminating between presence and absence. For this study, an AUC of ≥ 0.8 was considered acceptable (Hosmer and Lemeshow 1989; Anderson et al. 2003). Threshold dependent metrics percent correctly classified (PCC) and true skills statistic (TSS) were calculated to complement sensitivity and specificity (Allouche et al. 2006).

After assessing these metrics, several of the model algorithms were tuned to reduce discrepancies between training and test results (i.e. overfitting) and improve predictability. The GLM, MARS and MaxEnt models did not exhibit any discernable overfitting between the training and test datasets, and therefore, maintained default parameters. The BRT model was tuned for improvements by setting the tree complexity to 2, and the learning rate to 0.005 (Elith et al. 2008). The random forest model was also tuned by setting the number of trees created to 1000 and the “mtry” parameter (number of predictors randomly sampled at each split in the tree) to the square root of the number of candidate predictor layers (Belgiu and Drăguţ 2016).

Probability rasters were produced for each of the five model algorithms. First, binary suitability maps for each algorithm were created from probabilities ≥ 0.5; a commonly selected probability threshold cutoff used in habitat models (Buckley et al. 2010; Liu et al. 2005, Freeman and Moisen 2008). These were combined into a single ensemble with a binary output of suitable and unsuitable based on the “wisdom of the crowd” described by Stohlgren et al. (2010), where only one out of the five outputs is needed to be suitable as criteria for the ensemble to be suitable. Each of the binary layers were summed using the “EnsembleBuilder” module within SAHM insofar as the training AUC value was at least 0.8, and the difference between the training and cross-validation AUC values was not greater than 0.05 (Hahn et al. 2016). Finally, the island-level model was clipped to the extent of the AOI (see Figure 1).
Predictor variable importance and marginal response curves were calculated in SAHM to examine model contribution and dependence of predicted suitability (Jarnevich et al. 2018; Kesler and Walker 2015; Lötter and le Maitre 2014; Pearson 2007). Variable importance was determined by the change in AUC when predictor values are permuted between the occurrence records and background data points. The larger the change produced by the permutation, the greater influence the predictor has on the model. The predictors selected by each of the five models were relativized as a weighted mean. Marginal response curves were calculated for each model predictor with all other predictors held constant at their mean value to qualify the ecological plausibility of the suitable grid values (Jarnevich et al. 2018).

Our modeling approach seeks to maximize transparency and statistical rigor as outlined by Sofaer et al. (2019). Throughout, we employ their framework to adhere to the “Acceptable” or “Ideal” level for each attribute. Our assessment of performance is summarized in the supplementary section (Table S1).

Post-validation and model reiteration

While spatial autocorrelation was assessed and attempted to be curtailed in our study, we found no improvements to the overall performance of the model. To test whether our model could generalize to new data we chose to rigorously test the original model with independently collected data. Occurrence data from HBT management in 2017–2018 were used to determine the rate of the binary model to correctly classify new miconia targets (sensitivity; Type I Error). Following the post-validation, the models were reiterated with 2017–2018 validation data added. Background points were again derived from locations along helicopter routes during search efforts. Raster outputs were deductively compared for changes in pixel suitability assignments.

Mean weighted probability model

A mean weighted consensus probability model was created by weighting the individual probability density functions of each algorithm by their respective training set AUC values (see Equation 1 in Marmion et al. 2009) using the Raster Calculator tool in ArcMap (version 10.5; ESRI Co., Redlands, CA). The validation dataset (occurrences from 2017–2018) were used to explore the differences between the original model’s mean weighted probability values of occurrences that were correctly or incorrectly classified as suitable by our original binary EHSM occurrences from 1991–2016). This continuous, consensus probability model was used to explore potential scenarios for future monitoring efforts.
Figure 2. Binary ensemble habitat suitability model (EHSM) predictions of suitable habitat pixels for *Miconia calvescens* on the island of Maui, Hawaii. The area of interest (AOI) is delineated by the white outline. Models were developed using known occurrence records of miconia combined with five geospatial predictor variables representing the physical environment. Areas of red are pixels deemed suitable for miconia by at least one of the five model algorithms used in the ensemble. Data from 1991–2016 were used to train an initial model, and additional data from 2017–2018 management operations were used to validate the initial model and update to incorporate this new data. Yellow indicates areas on the landscape where the updated model predicted new pixels as suitable, while cyan indicates pixels where the predictions were reduced.

Table 2. ESHM evaluation metrics including area under the receiver operating characteristic curve (AUC), percent correctly classified (PCC), sensitivity, specificity, and true skill statistic (TSS) for five different modeling approaches. Presented are values for both the training set (e.g. BRT<sub>Train</sub>) and 10-fold cross-validation test set with standard deviation (e.g. BRT<sub>CV</sub>). Occurrence data was taken during monitoring efforts from 1991–2016. Five model algorithms were used in this study: boosted regression trees (BRT), generalized linear models (GLM), multivariate adaptive regression splines (MARS), maximum entropy (MaxEnt), and random forests (RF).

<table>
<thead>
<tr>
<th>Model</th>
<th>AUC</th>
<th>PCC</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>TSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRT&lt;sub&gt;CV&lt;/sub&gt;</td>
<td>0.83 ± 0.01</td>
<td>78.73 ± 0.43</td>
<td>0.91 ± 0.003</td>
<td>0.45 ± 0.01</td>
<td>0.36 ± 0.01</td>
</tr>
<tr>
<td>BRT&lt;sub&gt;Train&lt;/sub&gt;</td>
<td>0.84</td>
<td>78.98</td>
<td>0.91</td>
<td>0.45</td>
<td>0.36</td>
</tr>
<tr>
<td>GLM&lt;sub&gt;CV&lt;/sub&gt;</td>
<td>0.87 ± 0.01</td>
<td>79.22 ± 0.73</td>
<td>0.77 ± 0.01</td>
<td>0.84 ± 0.01</td>
<td>0.62 ± 0.01</td>
</tr>
<tr>
<td>GLM&lt;sub&gt;Train&lt;/sub&gt;</td>
<td>0.87</td>
<td>79.23</td>
<td>0.77</td>
<td>0.84</td>
<td>0.62</td>
</tr>
<tr>
<td>MARS&lt;sub&gt;CV&lt;/sub&gt;</td>
<td>0.85 ± 0.01</td>
<td>78.48 ± 0.58</td>
<td>0.89 ± 0.004</td>
<td>0.50 ± 0.01</td>
<td>0.39 ± 0.02</td>
</tr>
<tr>
<td>MARS&lt;sub&gt;Train&lt;/sub&gt;</td>
<td>0.85</td>
<td>78.50</td>
<td>0.89</td>
<td>0.50</td>
<td>0.39</td>
</tr>
<tr>
<td>MaxEnt&lt;sub&gt;CV&lt;/sub&gt;</td>
<td>0.85 ± 0.01</td>
<td>73.18 ± 0.46</td>
<td>0.67 ± 0.01</td>
<td>0.91 ± 0.005</td>
<td>0.58 ± 0.01</td>
</tr>
<tr>
<td>MaxEnt&lt;sub&gt;Train&lt;/sub&gt;</td>
<td>0.86</td>
<td>73.22</td>
<td>0.67</td>
<td>0.92</td>
<td>0.58</td>
</tr>
<tr>
<td>RF&lt;sub&gt;CV&lt;/sub&gt;</td>
<td>0.93 ± 0.003</td>
<td>85.36 ± 0.41</td>
<td>0.90 ± 0.01</td>
<td>0.72 ± 0.01</td>
<td>0.62 ± 0.01</td>
</tr>
<tr>
<td>RF&lt;sub&gt;Train&lt;/sub&gt;</td>
<td>0.93</td>
<td>85.45</td>
<td>0.90</td>
<td>0.72</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Results

Model development and assessment

The binary ensemble model developed with the original occurrence dataset predicted that 64.6% of the entire area of interest was suitable for miconia (Figure 2). The initial five models had AUCs > 0.80 for both training and cross-validation testing (Table 2) exceeding the threshold of good performance.
Characterizing suitable habitat of *M. calvescens* in the EMW


**Figure 3.** Marginal probability response plots for elevation (a), mean annual precipitation (b) and slope (c) for initial and updated models (grey and black lines, respectively).

(Hosmer and Lemeshow 1989). The difference in AUC values between training and test-validation models did not exceed 0.01, well below the threshold (e.g. 0.05), that would indicate model overfitting (Hahn et al. 2016). Model overfitting was also determined by assessing the marginal response curves for each predictor variable used to construct the models (Figure 3). All five of the binary ensemble models performed well on the other evaluation metrics (Table 2). TSS values of the models ranged 0.36–0.62, exceeding an acceptable threshold of 0.35 (Jarnevich et al. 2018).

Mean annual precipitation, elevation and slope were relatively equal contributors across the five different algorithms, while the two aspect indices (percent northness and eastness) were not significant contributors to any of the algorithms (Table 3). Among these three predictor variables, slope had the most varied importance across models; it was the highest contributor for the BRT algorithm, but only modestly important or not
imported for the other four models. In contrast, elevation and precipitation variables were consistently high contributors. Precipitation was excluded from the BRT model but was ~30% of relative importance across the other four algorithms.

The marginal probability response curves for both elevation (Figure 3a) and precipitation (Figure 3b) show monomodal distributions whereby suitability is confined to a range of each of these variables. Areas from sea level to just over 600 m asl and mean annual precipitation between 1400 mm to 5500 mm had high levels of suitable habitat in the EMW. Slope (Figure 3c), on the other hand, contributed to suitability across all values found within the EMW, with more influence on the overall model predictions in flatter areas on the landscape.

**Validation of binary EHSM**

Operations conducted in 2017–2018 collected additional occurrence data (n = 5,222) that were independent of the data (see Figure 1b for distribution of new data) used to construct the initial model (Figure 2). In the updated model output, the ensemble predicted that 2.5% of pixels originally predicted as suitable were no longer suitable given the new information, while 1.6% of unsuitable pixels were considered suitable with the updated data inputs; a net change of roughly 0.9% of the watershed (yellow and cyan pixels in Figure 2). Applying our binary representation of the model to this dataset, 93.9% of the total occurrences occurred in classified suitable pixels. The updated model, with a 4.5% increase in occurrence data (1991–2018), produced nearly identical assessment metrics (e.g., AUC > 0.84, TSS > 0.4) (Table S2). BRT improved for both AUC (0.83 to 0.87) and TSS (0.36 to 0.47).

Similar to the initial predictor importance results, slope was the most important predictor of suitability for the MARS and MaxEnt models, precipitation was the most important predictor for the GLM model, and elevation was the most influential variable in the RF model (Table S3). BRT displayed a shift in the two variables it found most important from slope and elevation in the initial model to elevation and precipitation for the updated model. Overall, the average marginal probability response curves displayed nearly identical distributions in these updated models as the original (Figure 3; black lines).

**Table 3.** Relativized predictor importance for explaining suitable habitat of *Miconia calvescens* in the EMW (data from 1991–2016) for five different modeling approaches: boosted regression trees (BRT), generalized linear models (GLM), multivariate adaptive regression splines (MARS), maximum entropy (MaxEnt), and random forests (RF).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>BRT</th>
<th>GLM</th>
<th>MARS</th>
<th>MaxEnt</th>
<th>RF</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Eastness</td>
<td>N/A</td>
<td>&lt; 1%</td>
<td>N/A</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>% Northness</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>Elevation</td>
<td>38%</td>
<td>39%</td>
<td>34%</td>
<td>35%</td>
<td>42%</td>
</tr>
<tr>
<td>Precipitation</td>
<td>N/A</td>
<td>40%</td>
<td>21%</td>
<td>22%</td>
<td>37%</td>
</tr>
<tr>
<td>Slope</td>
<td>62%</td>
<td>21%</td>
<td>45%</td>
<td>43%</td>
<td>20%</td>
</tr>
</tbody>
</table>

Characterizing suitable habitat of *M. calvescens* in the EMW

Figure 4. The weighted mean probability surface of the EHSM predictions of suitable habitat for *Miconia calvescens* on the island of Maui, Hawaii. The area of interest (AOI) is delineated by the white outline. Occurrence data from 1991–2016.

**Mean weighted probability model**

The weighted mean probability model (Figure 4) resulted in complete coverage of the AOI with a range of suitability grid values from 0.01–0.90. While the binary models performed at or above acceptable assessment metrics, the binary output was less than 100% accurate (93.9%) in correctly predicting suitable pixels. Comparison of mean probabilities from the validation dataset, allows for contrasting the differences in probabilities between correctly (n = 4095) and incorrectly (n = 317) classified occurrences. Mean weighted probabilities of correctly classified occurrences had a median value of 0.55 and an interquartile range between 0.48–0.64 (Figure 5). Incorrectly classified occurrences had lower probabilities with median value of 0.21 and an interquartile range from 0.12–0.31.

**Discussion**

*Miconia calvescens* represents a considerable threat to critical habitat for species endemic to the Hawaiian Islands and other island nations of the Pacific Rim. Over the last 28 years, management of miconia in the East Maui Watershed (EMW) has consisted of both ground and aerial control. With the extent of the problem beyond eradication, land managers are now focusing management to higher elevations and steeper slopes to control nascent populations before they can colonize and naturalize (Leary et al. 2013, 2014). Helicopter operations are the only alternative to managing miconia in most of the EMW, but cost-prohibitive to comprehensive surveillance of
the entire watershed. Strategically concentrating limited management efforts on suitable landscapes is the best alternative towards optimizing target interdiction and spread containment. Due to the extensive mapping efforts over the last 28 years, we were able to develop an ensemble habitat suitability model that determined where miconia has the potential to establish in the EMW and field validate its performance. We found a robust model was created that correctly identified over 93% of miconia occurrences from an independent dataset. As much of the currently uninfested areas at higher elevations are being prioritized for conservation of wildlife and ecosystem services, this information can help improve efficiency of monitoring and management efforts in these critical areas.
Drivers of suitability

Suitable habitat for miconia in the EMW was driven by elevation, precipitation, and slope. Others have identified precipitation and temperature, a proxy for elevation (Giambelluca et al. 2013; Morán-Tejeda et al. 2016) to be highly influential to miconia occurrence (Libeau et al. 2019; Vorsino et al. 2014). While windwardness (i.e. aspect) was also found important by Libeau et al. (2019) our models did not identify aspect indices (percent northness or eastness) as important in describing suitable habitat. This could be due to the difference in the modeled AOI (EMW vs. Islands of the Pacific Rim) and that the entire AOI is confined to the windward slope of Haleakala and with exceptionally high precipitation (see site description above). Decades of expert knowledge managing this entire mountain beyond the AOI infers an understanding of habitat confined to wet forest.

Even with the focus of our AOI, results generally agreed with others with respect to directionality and limits of miconia habitat. Miconia preferred elevations between ~100–600 m, but the highest recorded miconia was at 886 m in the EMW. While others have reported it to establish at elevations in excess of 1350 m (Wurdack 1980; Pouteau et al. 2011), it is unclear if miconia can establish at higher elevations of the EMW as temperature and precipitation are largely related to elevation. Mean annual precipitation totals between 1440–5500 mm were also found to be highly probable for miconia invasion. Note the cutoff at 1440 mm represents the lowest value observed in the modeled AOI. Areas adjacent to the AOI are more mesic which may not support miconia invasion. Observations on NW Oahu (1500 mm) (Mederios et al. 1997) and Tahiti (2000 mm) (Meyer 1996, 1998) corroborate this cutoff and suggest precipitation limits the spread of miconia in the EMW. The EHSM models suggest miconia is suitable to a range of terrain with highest probabilities on slopes < 30°. While this is dissimilar to Pouteau et al. (2011) (~15–60°), our model results still suggest high probability of invading steep terrain, just marginally less than in flatter areas. These differences are likely due to the spatial resolution of our EHSM (10 m) compared to others (5 m), which could decrease the accuracy of the values obtained at the 10 m scale in the EMW.

Using mean weighted probabilities to improve prioritization

While the binary ensemble model was shown to be effective in predicting presence of miconia, some occurrences were misclassified (7%). Over 80% of these observations were found within 30 m of the nearest suitable pixel (i.e. 2–3 pixels from the suitable habitat edge) and likely would have been detected. That being said, misclassifications can still occur when these habitat models are simplified to a binary representation, regardless of the probability cutoff value. An alternative option for prioritization is utilizing the probability surfaces provided by each model as these offer a higher
resolution gradient of suitability for land managers than a binary decision (Morisette et al. 2013). We used mean weighted probabilities from each of the five models and present averaged results in bins from 0.01–0.90 (Figure 4). This approach allows the land manager to view an average probability in less suitable areas and use their knowledge of the land to make an expert decision if monitoring and management should be conducted. This could be especially useful where areas on the landscape have different priorities for conservation. For instance, high priority areas would warrant monitoring efforts, even if suitability was low for miconia. The value of this approach is demonstrated when comparing the mean weighted probabilities of the correctly and incorrectly classified validation occurrences (Figure 5). While on average the incorrectly classified values had one half the predicted probability, they still had values ranging from 0.05–0.41. Over two-thirds of the misclassified occurrences from the binary model have probability values that fall within the range of the correctly classified occurrences (0.16–0.88). Therefore, while the binary ensemble model was able to minimize Type I error, a gradient for suitability clearly exists across the AOI.

If land managers utilized expert knowledge in combination with the probability values, most of these misclassifications with the binary approach could be avoided. Thus, this information can improve decisions with respect to monitoring and management. Use of mean weighted probabilities also allows for integration with other biological information associated with invasive species biology. Specifically, associating life history traits (e.g. recruitment, survivorship, fecundity, etc.), dispersal, and seed bank longevity to suitability beyond just occurrence can tailor monitoring and management with better anticipation of impact (see Figure 8 in Leary et al. 2018). Future work should seek to integrate these models to develop a better understanding of where miconia may spread to suitable locations, and how management should be applied to best meet the management goals of protecting critical assets in the EMW.

Conclusions

The invasion of miconia on the windward slope of Haleakala Volcano has not yet reached full occupancy of its suitable habitat. It is further encroaching on the upper elevation watershed with critical habitat that require protection. While the binary ensemble model was robust and field validated, incorrect classifications were observed 7% of the time. Further improvement can occur when utilizing mean weighted probabilities of models as they allow more flexibility for land managers to utilize while prioritizing monitoring and management. The approach highlighted here shows promise in prioritizing future monitoring and management efforts, especially when resources are limited. If utilized correctly, land managers
can apply this information to exploit environmental constraints to enhance management, outpacing the biology of this invasive species in cost-effective ways.

Acknowledgements

This study was supported in part by the USDA Forest Service Special Technology Development Program Award R5-2012-01, USDA Hatch Act Formula Grant project 132H, Hawaii Invasive Species Council Research and Technology Grant Program for FY18, Maui County Department of Water Supply and the Office of Economic Development via collaboration with the Maui Invasive Species Committee. Major funding and data archives for this management program provided by the Pacific Island Exotic Plant Management Team and Haleakala National Park of the National Park Service. We would like to acknowledge Christine Sheppard and two anonymous reviewers for very helpful and insightful suggestions and edits.

References


Characterizing suitable habitat of *M. calvescens* in the EMW


Characterizing suitable habitat of *M. calvescens* in the EMW


**Supplementary material**

The following supplementary material is available for this article:

**Table S1.** Individual attributes of the habitat suitability modeling process.

**Table S2.** Model evaluation metrics including area under the receiver operating characteristic curve (AUC), percent correctly classified (PCC), sensitivity, specificity, and true skill statistic (TSS) for five different modeling approaches.

**Table S3.** Relativized predictor importance for explaining suitable habitat of *Miconia calvescens* on the island of Maui, HI for five different modeling approaches for data from 1991–2018.

This material is available as part of online article from:

[http://www.reabic.net/journals/mbi/2021/Supplements/MBI_2021_Jorgensen_etal_SupplementaryTables.xlsx](http://www.reabic.net/journals/mbi/2021/Supplements/MBI_2021_Jorgensen_etal_SupplementaryTables.xlsx)