CSIRO PUBLISHING

International Journal of Wildland Fire
http://dx.doi.org/10.1071/WF11178

Wildfire ignition-distribution modelling: a comparative study in the Huron–Manistee National Forest, Michigan, USA

Avi Bar Massada[A,D], *Alexandra D. Syphard*[B], *Susan I. Stewart*[C] *and Volker C. Radeloff*[A]

[A]Department of Forest and Wildlife Ecology, University of Wisconsin – Madison, 1630 Linden Drive, Madison, WI 53706, USA.
[B]Conservation Biology Institute, 10423 Sierra Vista Avenue, La Mesa, CA 91941, USA.
[C]Northern Research Station, USDA Forest Service, 1033 University Avenue, Suite 360, Evanston, IL 60201, USA.
[D]Corresponding author. Email address: barmassada@wisc.edu

Abstract. Wildfire ignition distribution models are powerful tools for predicting the probability of ignitions across broad areas, and identifying their drivers. Several approaches have been used for ignition-distribution modelling, yet the performance of different model types has not been compared. This is unfortunate, given that conceptually similar species-distribution models exhibit pronounced differences among model types. Therefore, our goal was to compare the predictive performance, variable importance and the spatial patterns of predicted ignition-probabilities of three ignition-distribution model types: one parametric, statistical model (Generalised Linear Models, GLM) and two machine-learning algorithms (Random Forests and Maximum Entropy, Maxent). We parameterised the models using 16 years of ignitions data and environmental data for the Huron–Manistee National Forest in Michigan, USA. Random Forests and Maxent had slightly better prediction accuracies than did GLM, but model fit was similar for all three. Variables related to human population and development were the best predictors of wildfire ignition locations in all models (although variable rankings differed slightly), along with elevation. However, despite similar model performance and variables, the map of ignition probabilities generated by Maxent was markedly different from those of the two other models. We thus suggest that when accurate predictions are desired, the outcomes of different model types should be compared, or alternatively combined, to produce ensemble predictions.

Additional keywords: GLM, Maxent, Random Forests.

Received 18 December 2011, accepted 16 July 2012, published online 18 September 2012

Introduction

The increase in wildfire frequency and associated risks to the environment, houses and people calls for better understanding of the processes that control wildfire activity. This is especially important in areas where human settlements are interspersed with wildland fuels (i.e. the wildland–urban interface) (Radeloff *et al.* 2005), because this is where the majority of human wildfire-ignitions occur (Sturtevant and Cleland 2007; Syphard *et al.* 2007, 2008), and where the risk to human lives and property is highest (Bar Massada *et al.* 2009). Understanding the driving forces of ignitions and predicting where fires are likely to ignite are core elements in devising better strategies to mitigate wildfire initiation and identifying areas at risk (Finney 2005).

Wildfire ignitions are either human or natural. Both types of ignitions are essentially non-random processes that depend on top-down and bottom-up drivers that have explicit spatial patterns (Krawchuk *et al.* 2006; Parisien and Moritz 2009; Moritz *et al.* 2011; Narayanaraj and Wimberly 2011). Top-down drivers reflect broad-scale spatial variability of weather and climate, which in turn affect natural ignitions (e.g. lightning) directly, and also change fuel moisture thus changing the likelihood of an ignition (Latham and Williams 2001). Bottom-up drivers consist of local-scale variables that affect fuel combustibility, and proximity to human ignition sources (Krawchuk *et al.* 2006; Sturtevant and Cleland 2007; Syphard *et al.* 2007, 2008).

The most common approach to understanding the spatial patterns of wildfire ignitions and their driving forces is distribution modelling based on historical ignition locations (Sturtevant and Cleland 2007; Syphard *et al.* 2008; Bar Massada *et al.* 2011). Some studies looked solely at anthropogenic (Yang *et al.* 2007; Syphard *et al.* 2008) or lightning (Diaz-Avalos *et al.* 2001; Krawchuk *et al.* 2006) ignitions, whereas others examined both

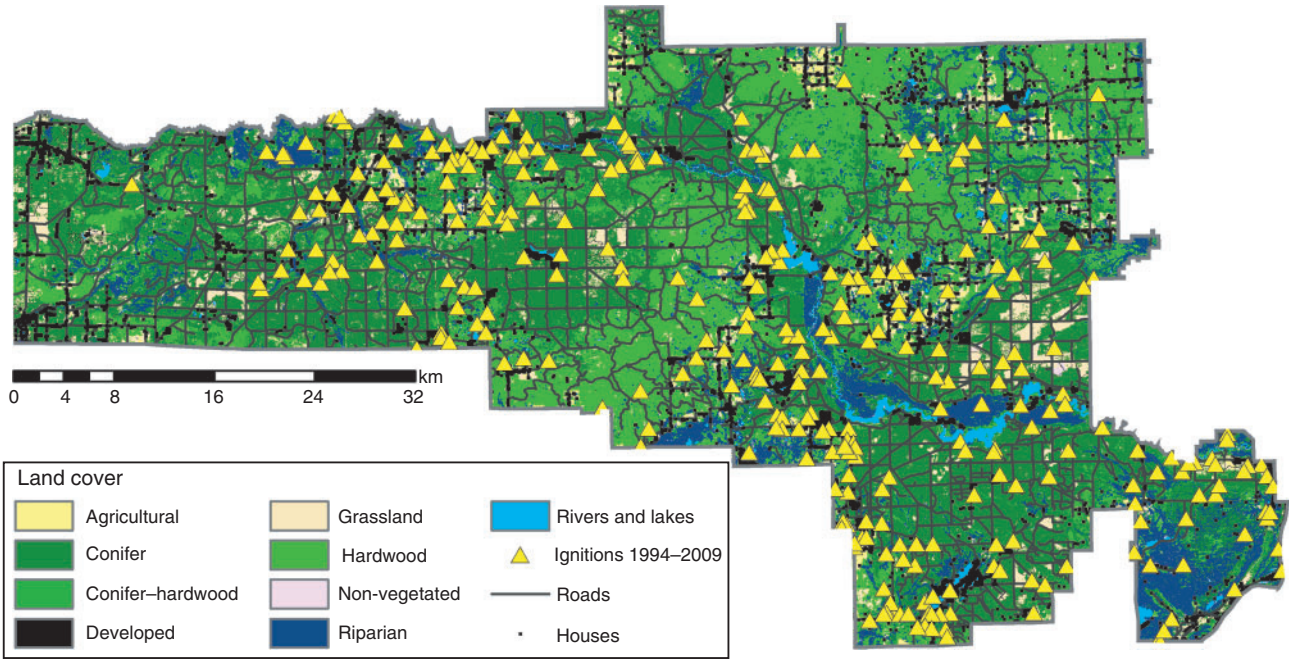

Fig. 1. Wildfire ignitions and landcover in the study area, Huron National Forest in northern Lower Michigan, USA.

(Sturtevant and Cleland 2007; Parisien and Moritz 2009). The general finding is that human and natural wildfire ignitions have distinctive spatial patterns that can be readily quantified using human and biophysical explanatory variables. However, it is important to distinguish the two goals of ignition modelling: (1) explanatory modelling to test hypotheses about the role that different factors play in causing ignitions, (2) predictive modelling to identify those areas that are most prone to ignitions, and where fire prevention or fuel reduction treatments can be targeted.

Conceptually and methodologically, ignition-distribution modelling is closely related to species-distribution modelling (SDM, Franklin 2010). The basic approach is to analyse wildfire ignition locations (analogous to locations of species occurrence) in relation to environmental variables hypothesised to influence the spatial distribution of ignitions (or species occurrences). Models estimate the response of wildfire ignitions (or species) to these environmental predictor variables. Just as in SDM, there are two types of ignition data that can be analysed: presence (the occurrence of an ignition event in a point in space) or abundance (the number of ignitions per unit area). The data type affects the choice of model type because presence data typically require a binomial response, whereas abundance data require a continuous response. Furthermore, although presence data are often accompanied by absence data, special modelling methods have also been developed for presence-only data, where presence locations are compared against background environmental conditions ('used *v.* available') (Elith *et al.* 2006; Franklin 2010), because it is often impossible to identify places where no ignition can occur.

A diverse set of models has been used to model ignitions (or species occurrences), including statistical methods such as generalised linear models (Nelder and Wedderburn 1972), general additive models (Hastie and Tibshirani 1990), and machine-learning algorithms such as Random Forests (Prasad *et al.* 2006) and Maximum Entropy (Maxent) (Phillips *et al.* 2006). Given the wide range of model types, the question is to what extent model choice affects the results? Are some model types consistently better at explaining or predicting wildfire ignitions? In wildlife studies, model type strongly affects the performance and resulting maps of SDMs (Elith *et al.* 2006; Guisan *et al.* 2007; Elith and Graham 2009), and we therefore hypothesised that ignition models would also be affected by model choice.

Our goal was to compare the performance of three common model types: generalised linear models, Random Forests and Maximum Entropy, for predicting the spatial distributions of wildfire ignitions in a fire-prone region in Michigan, USA. We compared the models in terms of: (1) predictive performance; (2) importance of explanatory variables and (3) spatial patterns of ignition probability.

Methods

Study area

We conducted our analysis across 273 398 ha of the Huron–Manistee National Forest (HMNF) in north-eastern Lower Michigan, USA (Fig. 1). The landscape is mostly flat to rolling with several major streams and lakes. Elevation ranges from 180 m ASL on the shores of Lake Huron in the east, to 420 m in the west. Forests are dominated by jack pine (*Pinus banksiana* Lamb), a species that promotes extreme fire activity. There are also stands of red pine (*P. resinosa* Ait.), white pine (*P. strobus* L.) and hardwood forests. The study area contains many settlements, ranging in sizes from towns (Mio, population of 1972), through subdivisions (e.g. Mack Lake, which burned

Table 1. Environmental predictor variables used in model development

Type	Name	Mean	Median	Minimum	Maximum
Landcover	% agriculture	4.5	0.0	0.0	100.0
	% conifers	35.1	17.2	0.0	100.0
	% grassland	8.0	0.0	0.0	100.0
	% hardwood	28.0	6.8	0.0	100.0
	% mixed	0.8	0.0	0.0	96.5
	% riparian	15.0	0.0	0.0	100.0
Anthropogenic	Distance to nearest structure (m)	2087.0	750.0	0.0	23 200.0
	Distance to nearest road (m)	1413.0	966.0	0.0	16 630.0
	Structure density, 1 km radius (km^{-1})	4.7	0.6	0.0	150.2
Topographic	Elevation (m)	302.9	306	177	474
	Slope (°)	1.6	1	0	27
	South-westness	104.7	103.5	0	201

completely in 1980), to thousands of more isolated structures interspersed throughout the forest. There is a gridded network of roads across the entire area, ranging from two-lane highways to double-track unpaved forest roads, in addition to hundreds of kilometres of smaller recreational trails.

The north-eastern lower peninsula of Michigan is a fire-prone area that experiences a large number of wildfires annually, most due to human activity (Cleland *et al.* 2004). The current fire regime in HMNF is characterised by short return intervals, with multiple wildfire occurrences annually, ranging from small, low-intensity surface fires to severe crown fires that burn thousands of hectares and result in substantial property loss. The historical fire regime was characterised by frequent stand-replacing fires (average fire return intervals as low as 26 years; Cleland *et al.* 2004), as a result of the fast-draining sandy soils and the high flammability of jack pine. Since 1979, there were 550 wildfires in the Huron part of the HMNF, with an average size of 26.79 ha (the largest fire, Mack Lake in 1980, burned 10 015 ha). Three fires burned more than 1000 ha of forest, and six fires burned between 100 and 1000 ha.

Data

We obtained ignitions data for 1994–2009 from the Mio ranger district office. This dataset details the location and cause of the 340 ignitions within the boundaries of the national forest in that period, all of which were anthropogenic. As predictors, we used three types of environmental data with a resolution of 30 m (Table 1). Landcover variables were based on the 2001 National Land Cover Database (NLCD) (Homer *et al.* 2007). We transformed the NLCD classes to continuous variables by calculating the relative proportion of each landcover type within 100 m of each ignition. The 100-m radius reduces the effect of any ignition location inaccuracies. Topographic variables were elevation, slope and south-westness. Human variables were distance to nearest road and structure (any type of house or non-residential building), and structure density within 1 km of each ignition. We calculated Pearson's correlation coefficient among predictive variables and found that correlations were always less than 0.8.

In addition to the 340 ignition points (presence data), we generated a random dataset of 1200 non-ignition points. These 1200 non-ignition points are not true absence points, because it is unknown if ignition could not (a true absence point), or simply did not occur. Instead, these 1200 non-ignition locations are background points, which we used as pseudo-absence data for models requiring absence data, and we identified the values of environmental variables for them in the same way as for the ignition points. Ultimately though, ignition data like ours does not represent presence–absence data, but rather presence-only data, and the best modelling approaches are those that employ a used-*v*.-available strategy.

We used a larger sample of background points because they need to capture the entire distributions of environmental variables across the study area. We withheld 20% of the points for model validation, and developed the models using 272 presence points and 960 background points (except for the Maxent model, where only presence points were used).

Statistical models

Generalised linear models (GLMs hereafter) are extensions of linear regression models that can handle non-normal distributions such as binomial distributions that are used for presence–absence data (Guisan *et al.* 2002; Franklin 2010). For the GLMs, we specified a logit link and binomial response because our response was binary (presence–background data):

$$\log\left(\frac{\mu}{1-\mu}\right) = \hat{\beta}_0 + \sum_{j=1}^{p} X_j \beta_j + \varepsilon$$

where $\mu = \mathrm{E}(Y)$ is the probability of class '1' (ignition occurrence) and $(1 - \mu)$ is the probability of class '0', β_0 is the model intercept, p the number of predictor variables X and ε the error term. To calculate variable importance, we used best subsets (Gavier-Pizarro *et al.* 2010). In best subsets, GLMs for all possible combinations of predictor variables are parameterised (Hosmer *et al.* 1989). Then, models are ranked according to their goodness of fit, and for each predictor variable, the number of times it appears in top ranked models (the 30 best models based

on Akaike's information criterion (AIC)) is counted. We ran the best subsets analysis using the bestglm package (McLeod and Xu 2010) in the R statistical software package (ver. 2.13, R Foundation for Statistical Computing, Vienna, Austria). Finally, we checked for spatial autocorrelation by visual interpretation of the variogram of the GLM residuals. Because we found no evidence for spatial autocorrelation in the GLM residuals, we assumed it did not affect the machine learning approaches either (Dormann *et al.* 2007).

Random Forest modelling (RF hereafter) is an ensemble technique that extends the classification trees modelling approach (Breiman 2001) by averaging the predictions of many individual classification trees, each developed using a subset of the training data. This algorithm overcomes the problem of instability in using single classification trees, resulting in higher prediction accuracy (Franklin 2010; Syphard and Franklin 2010). RF estimates model error and variable importance by comparing the prediction of each tree with a testing dataset consisting of all observations that were held back during its development (out-of-bag samples). Variable importance in RF is computed by calculating the mean squared error (MSE) of each tree by comparing its prediction based on the subset observations to those based on the out-of-bag observations. For each variable, observations from the out-of-bag sample are randomly permuted, and the MSE is recalculated. For any permuted variable, the difference in MSE for each tree (usually an increase), averaged across the forest and normalised by the standard deviation of the differences, is a measure of its importance, with larger differences measured as percent increase in MSE indicating more importance (Cutler *et al.* 2007). We generated the RF models using the randomForest package (Liaw and Wiener 2002) in the R statistical software package (R Foundation for Statistical Computing).

Maximum Entropy is a presence-only machine learning algorithm that iteratively contrasts environmental predictor values at occurrence locations with those of a large background sample taken across the study area (Phillips *et al.* 2006; Elith *et al.* 2011). Maxent has been used to model fire ignition in the US (Parisien and Moritz 2009), and fire occurrence in India's Ghats Mountains (Renard *et al.* 2012). Maxent represents π, the distribution of ignition locations for a set x of sites in the study area (Phillips and Dudík 2008). Each site x is assigned to a non-negative value $\pi(x)$, and these values across all sites sum to one. Maxent then generates a model of π based on environmental constraints obtained from the occurrence data, where each constraint is defined as a function of an environmental variable. The model requires that the average value of each of these functions is within a small margin of error from the empirical average of the corresponding environmental variables across all sample sites. Among the large number of possible probability distributions, Maxent selects the one with the Maximum Entropy (or the most uniform) as the best representative of the data (Phillips *et al.* 2006), and assigns a probability of occurrence to each location. Among presence-only SDMs, Maxent consistently demonstrates higher prediction accuracy, especially with small sample sizes (Elith *et al.* 2006; Pearson *et al.* 2007). Variable importance is quantified based on the increase of the regularised training gain through the iterations of model development, where training gain denotes the increase in the probability of ignition in the training locations (i.e. variables that contribute to increased chance of ignition in the actual ignition locations are considered more important). We used the stand-alone Maxent software, available at: http://www.cs.princeton.edu/~schapire/maxent/ (accessed 29 July 2011).

Model evaluation and comparison

To compare the prediction accuracy among the three models, we calculated the area under the curve (AUC) of the receiver-operating characteristic (ROC) plot (Hanley and McNeil 1982). The ROC plot depicts the relationship between the false-positive error rate (1-specificity; where specificity is the proportion of non-ignitions correctly predicted) and the true-positive rate (sensitivity, or the proportion of ignitions correctly predicted) for each threshold value to the probability of presence predicted by the model. Our three models predict continuous probabilities of occurrence between zero and one, thus a threshold needs to be set to determine whether the probability denotes presence or absence. The AUC is a threshold-independent metric because it evaluates the performance of a model at all possible threshold values (Franklin 2010). AUC ranges from 0.5 to 1, where 0.5 is analogous to a completely random prediction and 1 implies perfect prediction. AUC values between 0.5 and 0.7 denote poor, between 0.7 and 0.9 denote moderate, and larger than 0.9 denote high model performance (McCune and Grace 2002). Although AUC has been criticised (Lobo *et al.* 2008), it is the standard method to assess prediction accuracy because of its threshold independence and the ease of interpreting its results.

Evaluating variable importance

We used two approaches to measure variable importance in the ignition models. First, we calculated variable importance using the native metric of each model: deviance for GLM, percent increase in MSE for RF, and regularised training gain for Maxent. We then calculated the average importance rank of each variable within its model, and created a combined ranking of variable importance (Syphard and Franklin 2009).

In addition, given the different characteristics of the models, we used a jackknife estimator of variable importance based on the change in AUC using the testing data. This yields directly comparable results across the three model types. The approach consists of removing predictor variables from the full model one at a time, training the model and calculating the AUC using the testing data. The difference between the full- and partial-model (without the variable) AUC indicates the contribution of each variable to the model. Thus it represents the information provided by a given variable that is not present in other variables. In addition, we quantified the AUC of the model using one variable at a time, compared the AUC values of single variable models and ranked the variables accordingly.

Comparing prediction maps of ignition probability

For each model, we generated prediction maps of ignition probability based on raster maps of all predictor variables. These maps are equivalent to habitat suitability maps in SDM and depict the suitability of a given pixel for ignition, relative to all other pixels. Pixel values range from 0 (unsuitable)

to 1 (perfectly suitable for ignition). In the case of the GLM, it is also possible to calculate actual ignition probabilities for a 16-year period by adjusting probabilities according to the sampling effort of the background points (Preisler *et al.* 2004). This is done to correct for differences between the ratio of presence and background points in model development (1 : 3.5) *v*. their actual, much smaller ratio. We estimated the ratio between the number of background points used and the potential number of background points in the study area, and subtracted its logarithm from the GLM intercept. We estimated the potential number of background points by converting the ignition locations from points to 30-m pixels in the map of the study area (Syphard *et al.* 2008). All non-ignition pixels (3 037 476) were considered as potential background points, out of which we had used 960. The correction factor was therefore the logarithm of 960/3 037 476 or 8.05. This adjustment reduced ignition probabilities by three orders of magnitude (from 0.01–1 to 0.0003–0.0008) while perfectly retaining the spatial pattern of probabilities.

To assess the differences among models, we compared the three maps visually and calculated Spearman's correlation coefficient between each pair of maps. In addition, we calculated the proportion of area suitable for ignition for each model by applying thresholds to the prediction maps that divide them into 'suitable for ignition' and 'unsuitable for ignition' classes. Because there are multiple criteria for choosing thresholds (Freeman and Moisen 2008), we evaluated threshold values between 0.1 and 0.9 in increments of 0.05, and compared the resulting curves of 'suitable for ignition' area *v*. threshold values among the three models. For the GLM, we used the unadjusted prediction map because the adjustment systematically lowered the ignition probabilities outside the range of our thresholds. The spatial patterns of the adjusted and unadjusted prediction maps generated by GLM were identical, and our focus here was on model comparison and not on absolute prediction values *per se*.

Results

Performance of modelling approaches

Predictive performance based on independent test data was intermediate for all models, and similar among model types. The Maxent model had the best performance with test data AUC of 0.716, followed by RF with 0.694, and the full GLM with 0.664.

Variable importance

In all three models, and based on both native and jackknife measures of importance, predictor variables related to human settlement and infrastructure were the strongest predictors of wildfire ignitions (Fig. 2, Table 2). Using the average rankings from the native variable importance measures, elevation was the best predictor of wildfire ignitions, followed closely by distance to nearest road and structure density within 1 km (Table 2). However, elevation shared the first rank with distance to nearest road in the GLM, and was also the highest ranked in Maxent, but only the fourth highest ranked in RF. Structure density was the highest ranked in RF (3rd in GLM and 4th in Maxent), and distance to nearest house was ranked 3rd, 6th and 3rd in RF, GLM and Maxent. Among the landcover variables, percentage of conifer forest was the best predictor (ranked 8th in GLM, 5th in RF and 5thin Maxent). Native measures of importance within

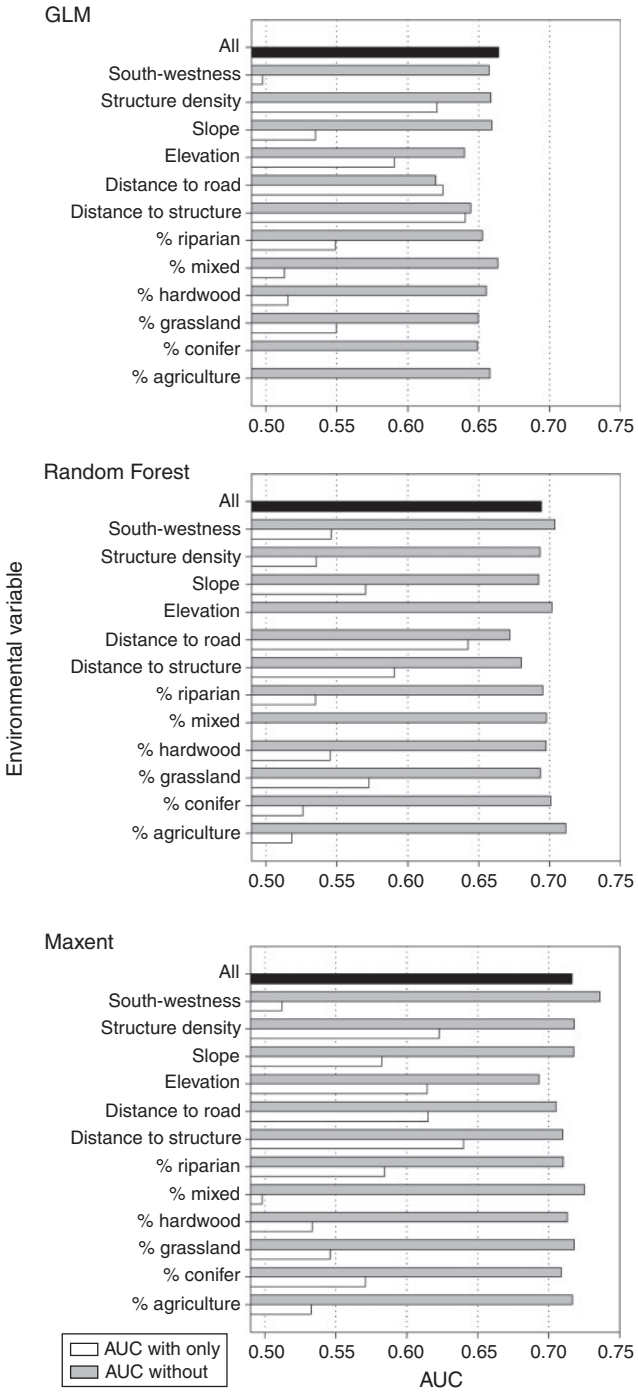

Fig. 2. Jackknife estimations of variable importance for the three types of ignition-distribution models. Bars denote the area under the curve of the receiver operator characteristic curve (AUC of ROC). Black bars depict the full-model AUC, whereas white bars denote the AUC of univariate models for the corresponding variables and grey bars denote the AUC of models with all variables except the corresponding variable.

(rather than across) the models revealed that for RF and Maxent, five variables (distance to nearest road and nearest house, structure density, elevation and percent conifer forest) proved to be the most important (Table 2). Similarly, only five variables in

Table 2. Variable importance measures and their corresponding rankings per model and overall based on the average rank for each model: % presence in 30 best GLM models, % increase in MSE in RF and regularised training gain in Maxent

Variable name	% presence in top 30 models (rank)	% increase in MSE (rank)	Regularised training gain (rank)	Average rank
Elevation	100.00 (1)	0.01380 (2)	18.8 (2)	1.67
Distance to nearest road	100.00 (1)	0.01200 (4)	19.3 (1)	2.00
Structure density	96.67 (3)	0.02090 (1)	13.2 (4)	2.67
Distance to nearest house	30.00 (6)	0.01400 (3)	17.0 (3)	4.00
% conifer	13.33 (8)	0.01140 (5)	12.0 (5)	6.00
South-westness	70.00 (4)	0.00311 (9)	4.2 (6)	6.33
% hardwood	56.67 (5)	0.00891 (6)	2.7 (9)	6.67
% agriculture	20.00 (7)	0.00003 (11)	3.8 (7)	8.33
% riparian	10.00 (10)	0.00345 (7)	2.6 (10)	9.00
% grassland	6.67 (11)	0.00272 (8)	3.4 (8)	9.00
Slope	13.33 (8)	0.00132 (10)	0.8 (11)	10.00
% mixed	6.67 (11)	−0.00054 (12)	2.1 (12)	11.33

the GLM appeared in more than half of the 30 best models but they were slightly different variables (Table 2).

According to the jackknife measures, when predictor variables were analysed independently (Fig. 2), distance to nearest structure was the strongest predictive variable in both GLM and Maxent, and the second strongest predictive variable in RF (Fig. 2). Distance to nearest road was the strongest predictor variable in RF by far, the second strongest in GLM and the third strongest in Maxent. Structure density, was the second best predictor in Maxent and third in GLM, but one of the weakest predictive variables in RF. Among the topographic variables, elevation was the strongest (and the fourth overall) in GLM and Maxent, whereas slope was the strongest variable in RF. Landcover variables were always weaker than human and topographic variables, and in general were not useful predictors of wildfire ignition by themselves.

The jackknife estimate of variable importance (Fig. 2) revealed slightly different results, as the order of the ranking changed, but the general trend of variable types remained. Human variables (distance to road in GLM and RF) remained the strongest predictors of wildfire ignitions, but the effect of omitting single variables on AUC was small. In Maxent, elevation was the strongest predictor, followed by distance to road, distance to nearest structure and several landcover variables: percentage of riparian, percentage of conifer forests and percentage of hardwood forests around the ignition point.

Spatial patterns of ignition probabilities

Results from the three models revealed distinctive spatial patterns of ignition probabilities (Fig. 3). Ignition probabilities tended to be higher near roads and houses, and increased with housing density, in agreement with the relative variable importance. The south-eastern and northern parts of the study area had the highest predicted ignition probabilities, corresponding with areas of higher housing densities. Because the road network is nearly uniformly distributed across the study area, there was an underlying lattice of high ignition probabilities in all three maps.

The GLM and RF models produced similar prediction maps (Spearman's $R = 0.74$) whereas Maxent had a low correlation with the RF model (Spearman's $R = 0.6$) and a moderate correlation with the GLM (Spearman's $R = 0.73$). These differences were also evident from the decay curve of the area suitable for ignition *v*. ignition threshold (Fig. 4). GLM and RF both depict a sharp decrease of the area suitable for ignition with increasing ignition threshold, whereas Maxent retained larger areas suitable for ignitions at higher threshold values. Whereas all three models exhibited higher ignition probabilities near roads, Maxent had a slower decay of probability values with increasing distance to roads, and distance from areas of high housing densities.

Discussion

Determining which model type to use for ignition-distribution modelling is important, because the outcomes of ignition-modelling studies may have direct management implications and inform fire science. Previous findings from SDM for wildlife suggested that machine learning algorithms may be more suitable than statistical models (Elith *et al.* 2006) given that they do not require normally distributed data, which wildfire ignitions data typically are not. However, variations in both model methodology (Elith *et al.* 2006) and species characteristics (Syphard and Franklin 2009) result in different outcomes, suggesting that there may not be one perfect modelling approach.

Unlike some comparisons in the SDM literature (Elith *et al.* 2006; Elith and Graham 2009), the three model types compared here were similar in terms of prediction accuracy. GLM, RF and Maxent all had moderate prediction accuracies, with RF and Maxent slightly outperforming GLM. Compared with a similar study conducted for a broader area (Parisien and Moritz 2009), the AUC for all three model types was low (0.71 compared with 0.8–0.9 in their study). This could be because of the relatively low variability in both topography and distance to roads, which were the main drivers of ignitions. However, in a study in the Santa Monica Mountains in California where there was higher variability in environmental variables (Syphard *et al.* 2008), a GLM-based ignition-distribution model had a similar AUC value to ours (0.71 *v*. 0.66).

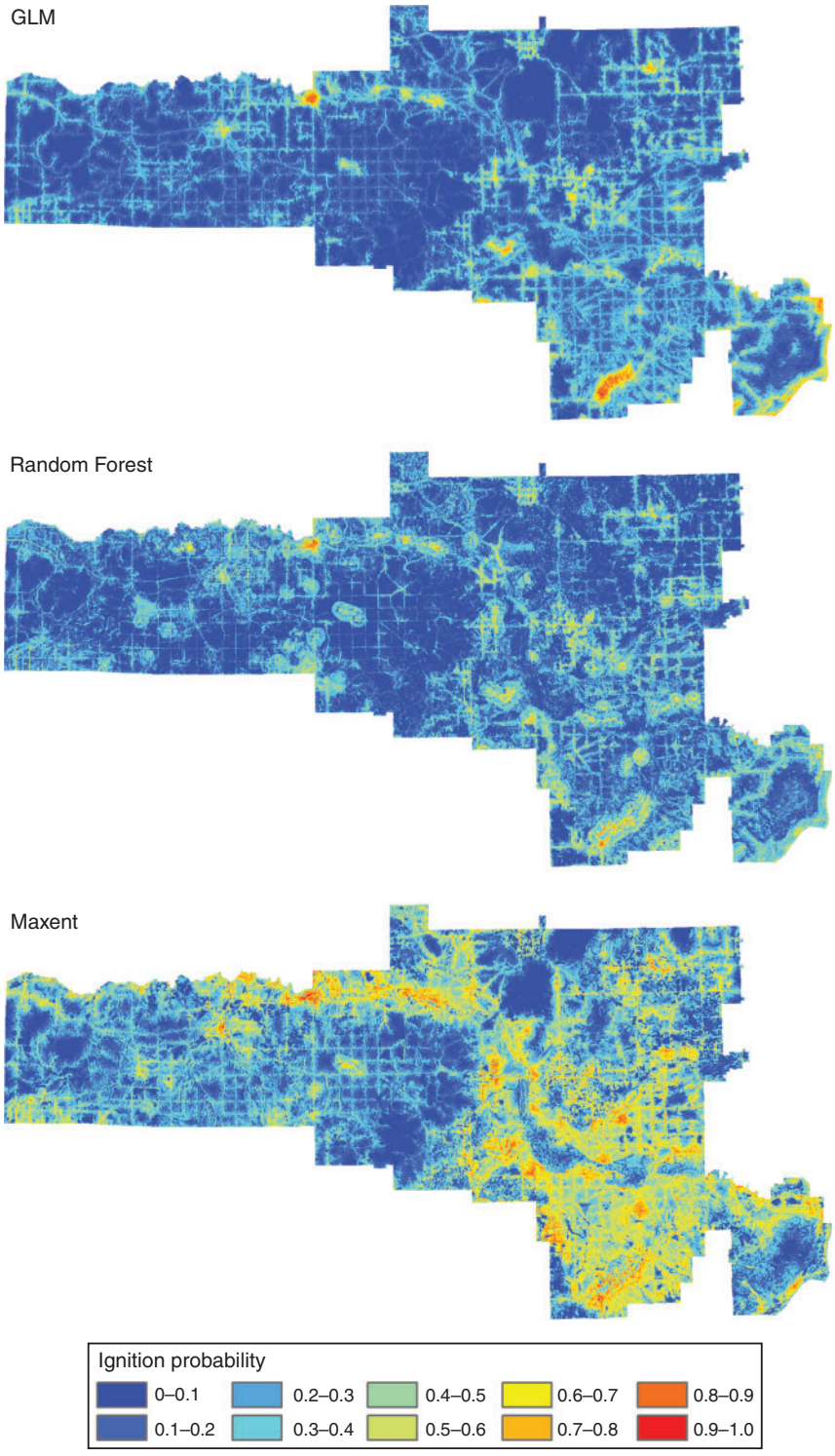

Fig. 3. Predictive maps of ignition probability for the three model types. Values denote the suitability of each pixel for having an ignition occurrence. For the GLM, we show the unadjusted probabilities to enhance comparability. Adjusted probabilities differ in value but not in pattern.

Importance of predictive variables

The most important predictive variables identified in this study were distance to nearest road, distance to nearest house and structure density. These variables are often strong predictors of wildfire ignitions in human-dominated landscapes (Sturtevant and Cleland 2007; Syphard *et al.* 2008). Distance to nearest road is a measure of accessibility, whereas distance to nearest house and structure density are measures of ambient population

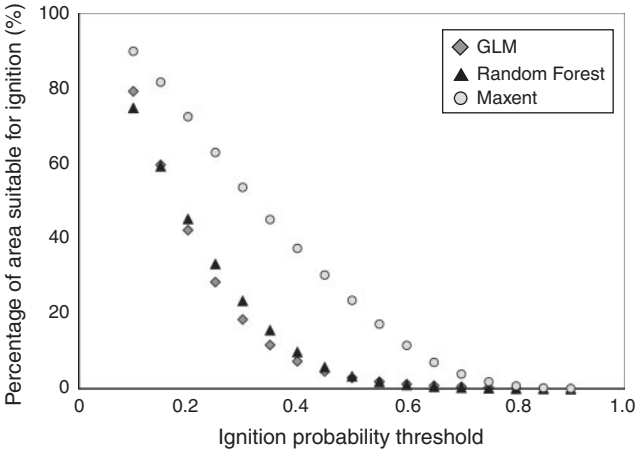

Fig. 4. Relationship between area suitable for ignition in the prediction map of each model type and the threshold to ignition probability (a value between 0 and 1, above which a pixel is considered suitable for ignition; the curve denotes the total percentage of the study area that falls under this ignition-suitability criterion, for a given threshold value). The analysis for the GLM is based on unadjusted probabilities.

density. In most human-dominated landscapes, in which anthropogenic ignitions surpass natural ignitions, both human accessibility and population density are likely to be strong predictors of ignition risk. In less-populated but accessible areas, distance to nearest road is likely to be the best predictor; in remote areas with little human activity and where lightning is a major ignition source, landcover and topography become the strongest predictors of ignitions (Krawchuk *et al.* 2006). Interestingly, structure density had a positive relationship with ignition probability (evident from the shape of Maxent marginal response curves and GLM coefficients). By contrast, other studies (Syphard *et al.* 2007, 2009) have indicated a nonlinear relationship where intermediate population densities promote the highest fire frequencies. The lack of high population densities in the HMNF is most likely the reason for this difference.

We did not expect topography to play a major role in explaining ignitions, because the study area's terrain is flat to rolling. Nevertheless, elevation proved to be an important predictor of wildfire ignition in all models. This may be because the wildland fuels in relatively high areas dry faster than those in lowlands. Finally, land cover variables were not important, indicating that all vegetation types in the study area can support ignitions. Although fire spread rates are markedly higher in coniferous than in deciduous forest in the HMNF, we did not model this.

Prediction maps of ignition probability

Whereas our measure of model accuracy (AUC) revealed only minor differences in predictive performance among model types, the maps of ignition probability showed pronounced differences. GLM and RF produced similar maps, though RF tended to indicate particularly high probability around developed areas. In a similar comparison of SDMs (Syphard and Franklin 2009), the prediction maps of GLM and RF had a fairly low correlation of 0.6, compared with 0.74 in our study. In contrast to GLM and RF, Maxent generated larger areas of high ignition probability, again especially near developed areas.

Despite the moderate correlations among the predicted maps, there were notable differences in patterns and magnitude. For example, although the correlations between RF and GLM, and RF and Maxent prediction maps were nearly the same (0.73 and 0.74), the actual patterns in these maps were markedly different. The GLM–RF curves in Fig. 4 were very similar, whereas the RF–Maxent curves were different. Unfortunately, comparing continuous probability maps is not a straightforward task, as there is no single metric that can quantify the complexity of the spatial patterns they portray. Setting thresholds and then comparing actual values or patterns may be a first step, but comparing binary maps is not much simpler than the approach that we used (Pontius and Millones 2011). In addition, quantifying differences in spatial patterns is hampered by the vast number of available metrics, none of which provides a comprehensive measure of pattern (Li and Wu 2004).

Despite the inherent limitations of map comparisons, even the visual differences in spatial patterns among our models raise an important question: if different models produce similar prediction accuracies, but different spatial patterns of ignition probabilities, which one should be chosen for management purposes? We suggest three possible answers: (1) use the most accurate, regardless of how minor the differences are in accuracy; (2) use all maps, and treat their predictions as a range of possible probabilities and (3) average the map predictions, possibly weighted by their measure of accuracy (such as AUC). The latter is akin to ensemble modelling, which has been explored for SDM, and can have higher predictive power than any one model type (Marmion *et al.* 2009).

Presence–absence data for ignition modelling

A main conceptual difference among the three model types is that GLM and RF are presence–absence models, whereas Maxent is a presence-only model. Both presence–absence and presence-only models are suitable for predicting wildfire ignitions, as long as the meaning of their predictions is properly understood (Parisien and Moritz 2009): presence–absence models are based on the differences in environmental conditions between ignition and non-ignition locations, whereas presence-only models are based on the differences in conditions between ignition locations and the entire landscape. Presence-only models such as Maxent may be superior to presence–absence models in areas that experienced several large fires in a short time (Parisien and Moritz 2009). Such a pattern would suggest that those areas that did not burn have the potential to do so, making it impossible to determine whether unburned areas are true absences. On the other hand, a presence–absence modelling approach may be more justified in areas with long fire records, and where only a small proportion of the landscape can carry a fire. A main conceptual hurdle in understanding the outcomes of ignition distribution models is the meaning of 'absence'. In ignitions data, most or all of the presences (ignition locations) are accounted for, which may make it seem as if all other locations in the landscape are therefore true absences. However, these other locations are really only pseudo absences because, although they did not experience an ignition during the temporal span of the dataset, they could feasibly support an ignition in the future, or may have ignited in the past. We do not have adequate knowledge to confirm these as suitable or unsuitable for

ignition; they are best described as representative of the background environment. Furthermore, the inability of presence-only models to estimate the prevalence of occurrences in the landscape, and the sample-selection bias (Ward *et al.* 2009; Elith *et al.* 2011), are not a problem with ignitions data, because all presences are accounted for.

Having conducted our analysis on a single landscape, one question is how widely applicable our results are. Comparing Maxent and boosted regression trees across different regions and at different scales, Parisien and Moritz (2009) found that model performance was similar, as were rankings of predictive variables. Similarity between their results and ours, despite differences in spatial resolution and ignition type, suggests that in landscapes characterised by a human-driven ignition regime, anthropogenic variables such as housing and road density and distance to roads will be the most important predictors of ignition, regardless of model type. As for model performance, we predict that in landscapes where there is a strong nonlinear relationship between ignitions and environmental variables, machine-learning methods will be more suitable than parametric models.

Conclusion

We analysed the differences in performance, variable importance and prediction maps of ignition probability among three common types of statistical models and found that machine-learning algorithms (RF and Maxent) performed slightly better than the chosen parametric approach (GLM) in terms of prediction accuracy. Despite similar predictive power of the different model types, the resulting prediction maps of ignition probability were very different; with the Maxent map being the most different as it predicted higher ignition probabilities across larger proportions of the landscape. Because there is no single, perfect modelling tool (Elith *et al.* 2006), studies of wildfire ignitions may benefit from using multiple approaches, yielding a range of predictions rather than a single map. The number of issues still to be clarified around ignition distribution modelling argues strongly for continuing improvement in ignition data, and for further research into ignition patterns and processes.

Acknowledgements

We gratefully acknowledge support for this research by the US Forest Service Northern Research Station, and the assistance of S. Goldman, J. Alyea and L. McNichols of the Huron–Manistee National Forest.

References

Bar Massada A, Radeloff VC, Stewart SI, Hawbaker TJ (2009) Wildfire risk in the wildland–urban interface: a simulation study in northwestern Wisconsin. *Forest Ecology and Management* **258**, 1990–1999. doi:10.1016/J.FORECO.2009.07.051

Bar Massada A, Syphard AD, Hawbaker TJ, Stewart SI, Radeloff VC (2011) Effects of ignition location models on the burn patterns of simulated wildfires. *Environmental Modelling & Software* **26**(5), 583–592. doi:10.1016/J.ENVSOFT.2010.11.016

Breiman L (2001) Random forests. *Machine Learning* **45**, 15–32.

Cleland DT, Crow TR, Saunders SC, Dickmann DI, Maclean AL, Jordan JK, Watson RL, Sloan AM, Brosofske KD (2004) Characterizing historical and modern fire regimes in Michigan (USA): a landscape ecosystem approach. *Landscape Ecology* **19**(3), 311–325. doi:10.1023/B:LAND.0000030437.29258.3C

Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007) Random forests for classification in ecology. *Ecology* **88**(11), 2783–2792. doi:10.1890/07-0539.1

Diaz-Avalos C, Peterson D, Alvarado E, Ferguson S, Besag J (2001) Space-time modelling of lightning-caused ignitions in the Blue Mountains, Oregon. *Canadian Journal of Forest Research* **31**, 1579–1593.

Dormann CF, McPherson JM, Araújo MB, Bivand R, Bolliger J, Carl G, Davies RG, Hirzel A, Jetz W, Kissling WD, Kühn I, Ohlemüller R, Peres-Neto PR, Reineking B, Schröder B, Schurr FM, Wilson R (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**(5), 609–628. doi:10.1111/J.2007.0906-7590.05171.X

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, Graham CH, Anderson RP, Dudík 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 JMM, Townsend Peterson A, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberón J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**(2), 129–151. doi:10.1111/J.2006.0906-7590.04596.X

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Diversity & Distributions* **17**(1), 43–57. doi:10.1111/J.1472-4642.2010.00725.X

Finney MA (2005) The challenge of quantitative risk analysis for wildland fire. *Forest Ecology and Management* **211**(1–2), 97–108. doi:10.1016/J.FORECO.2005.02.010

Franklin J (2010) 'Mapping Species Distributions.' (Cambridge University Press: New York)

Freeman EA, Moisen GG (2008) A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling* **217**(1–2), 48–58. doi:10.1016/J.ECOLMODEL.2008.05.015

Gavier-Pizarro GI, Radeloff VC, Stewart SI, Huebner CD, Keuler NS (2010) Housing is positively associated with invasive exotic plant species richness in New England, USA. *Ecological Applications* **20**(7), 1913–1925. doi:10.1890/09-2168.1

Guisan A, Edwards TC, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling* **157**(2–3), 89–100. doi:10.1016/S0304-3800(02)00204-1

Guisan A, Zimmermann NE, Elith J, Graham CH, Phillips S, Peterson AT (2007) What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? *Ecological Monographs* **77**(4), 615–630. doi:10.1890/06-1060.1

Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**(1), 29–36.

Hastie T, Tibshirani R (1990) 'Generalized Additive Models.' (Chapman and Hall: New York)

Homer C, Dewitz J, Fry J, Coan M, Hossain N, Larson C, Herold N, McKerrow A, VanDriel JN, Wickham J (2007) Completion of the 2001 National Land Cover Database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing* **73**(4), 337–341.

Hosmer DW, Jovanovic B, Lemeshow S (1989) Best subsets logistic regression. *Biometrics* **45**(4), 1265–1270. doi:10.2307/2531779

Krawchuk M, Cumming S, Flannigan M, Wein R (2006) Biotic and abiotic regulation of lightning fire initiation in the mixedwood boreal forest. *Ecology* **87**, 458–468. doi:10.1890/05-1021

Latham D, Williams E (2001) Lightning and forest fires. In 'Forest Fires'. (Eds E Johnson, K Miyanishi) pp. 375–418. (Academic Press: San Diego, CA)

Li H, Wu J (2004) Use and misuse of landscape indices. *Landscape Ecology* **19**, 389–399. doi:10.1023/B:LAND.0000030441.15628.D6

Liaw A, Wiener M (2002) Classification and regression by random forest. *R News* **2**(3), 18–22.

Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* **17**(2), 145–151. doi:10.1111/J.1466-8238.2007.00358.X

Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W (2009) Evaluation of consensus methods in predictive species distribution modelling. *Diversity & Distributions* **15**(1), 59–69. doi:10.1111/J.1472-4642.2008.00491.X

McCune B, Grace J (2002) 'Analysis of ecological communities.' (MJM Software Design: Glenden Beach, OR)

McLeod AI, Xu C (2010) bestglm: Best subsets GLM. In 'R package version 0.31'. Available at http://cran.r-project.org/web/packages/bestglm/vignettes/bestglm.pdf [Verified 22 August 2012]

Moritz MA, Hessburg PF, Povak NA (2011). Native fire regimes and landscape resilience. In 'The Landscape Ecology of Fire'. (Eds D McKenzie, C Miller, DA Falk) pp. 51–86. (Springer: Dordrecht, the Netherlands)

Narayanaraj G, Wimberly M (2011) Influences of forest roads on the spatial pattern of wildfire boundaries. *International Journal of Wildland Fire* **20**(6), 792–803. doi:10.1071/WF10032

Nelder JA, Wedderburn RWM (1972) Generalized linear models. *Journal of the Royal Statistical Society. Series A (General)* **135**(3), 370–384. doi:10.2307/2344614

Parisien M-A, Moritz MA (2009) Environmental controls on the distribution of wildfire at multiple spatial scales. *Ecological Monographs* **79**(1), 127–154. doi:10.1890/07-1289.1

Pearson RG, Raxworthy CJ, Nakamura M, Townsend Peterson A (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography* **34**(1), 102–117. doi:10.1111/J.1365-2699.2006.01594.X

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* **31**(2), 161–175. doi:10.1111/J.0906-7590.2008.5203.X

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**(3–4), 231–259. doi:10.1016/J.ECOLMODEL.2005.03.026

Pontius RG, Jr, Millones M (2011) Death to kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing* **32**(15), 4407–4429. doi:10.1080/01431161.2011.552923

Prasad A, Iverson L, Liaw A (2006) Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* **9**(2), 181–199. doi:10.1007/S10021-005-0054-1

Preisler HK, Brillinger DR, Burgan RE, Benoit JW (2004) Probability based models for estimation of wildfire risk. *International Journal of Wildland Fire* **13**, 133–142. doi:10.1071/WF02061

Radeloff VC, Hammer RB, Stewart SI, Fried JS, Holcomb SS, McKeefry AJ (2005) The wildland–urban interface in the United States. *Ecological Applications* **15**(3), 799–805. doi:10.1890/04-1413

Renard Q, Pélissier R, Ramesh BR, Kodandapani N (2012) Environmental susceptibility model for predicting forest fire occurrence in the western Ghats of India. *International Journal of Wildland Fire* **21**, 368–379. doi:10.1071/WF10109

Sturtevant BR, Cleland DT (2007) Human and biophysical factors influencing modern fire disturbance in northern Wisconsin. *International Journal of Wildland Fire* **16**, 398–413. doi:10.1071/WF06023

Syphard AD, Franklin J (2009) Differences in spatial predictions among species distribution modeling methods vary with species traits and environmental predictors. *Ecography* **32**(6), 907–918. doi:10.1111/J.1600-0587.2009.05883.X

Syphard AD, Franklin J (2010) Species traits affect the performance of species distribution models for plants in southern California. *Journal of Vegetation Science* **21**(1), 177–189. doi:10.1111/J.1654-1103.2009.01133.X

Syphard AD, Radeloff VC, Keely JE, Hawbaker TJ, Clayton MK, Stewart SI, Hammer RB (2007) Human influence on California fire regimes. *Ecological Applications* **17**(5), 1388–1402. doi:10.1890/06-1128.1

Syphard AD, Radeloff VC, Keuler NS, Taylor RS, Hawbaker TJ, Stewart SI, Clayton MK (2008) Predicting spatial patterns of fire on a southern California landscape. *International Journal of Wildland Fire* **17**, 602–613. doi:10.1071/WF07087

Syphard AD, Radeloff VC, Hawbaker TJ, Stewart SI (2009) Conservation threats due to human-caused increases in fire frequency in Mediterranean-climate ecosystems. *Conservation Biology* **23**, 758–769. doi:10.1111/J.1523-1739.2009.01223.X

Ward G, Hastie T, Barry SC, Elith J, Leathwick JR (2009) Presence-only data and the EM algorithm. *Biometrics* **65**, 554–563. doi:10.1111/J.1541-0420.2008.01116.X

Yang J, He H, Shifley S, Gustafson E (2007) Spatial patterns of modern period human-caused fire occurrence in the Missouri Ozark Highlands. *Forest Science* **53**, 1–15.