Spatially Explicit Prediction of Residual Vegetation Patch Occurrence within Boreal Wildfires

Araya, Y. H.,[1]* Remmel, T. K.[1] and Perera, A. H.[2]
[1]York University, Department of Geography, 4700 Keele Street, Toronto, ON, Canada, M3J 1P3
E-mail: yikalo@yorku.ca, remmelt@yorku.ca
[2]Ontario Forest Research Institute, Ontario Ministry of Natural Resources, 1235 Queen Street East
Sault Ste. Marie, ON, Canada, P6A 2E5, E-mail: ajith.perera@ontario.ca
**Corresponding author*

Abstract

Even after landscapes are disturbed by boreal wildfires, considerable quantities of residual vegetation remain. A method for developing spatially explicit predictive probability maps is presented to identify the presence of residual vegetation within burned boreal landscapes in North-western Ontario, where we learned residual presence expectations from a suite of wildfires that burned from 2002-2003. This approach relies on easily measured variables due to the lack of detailed local information in this remote region. We cross-validate predictions within our training data, a suite of 11 wildfires, using a bootstrapping approach (internal validation) and then test the model on an independent event that burned in 2011 (external validation). The predictive model is based on the Random Forest algorithm and is implemented at 5 separate spatial resolutions (4 to 64 m). The model has a reasonably high predictive power as determined by internal validation. The external validation yielded accuracy better than random prediction and we conclude that the existence of residual vegetation is clearly related to the presence of firebreak features and proximity to wet regions. Our repeatable approach is spatially explicit, implemented in an open software environment, and provides acceptable results where local and detailed data availability is substantially limited.

1. Introduction

Wildfires, the most prevalent natural disturbances in boreal forest ecosystems (Johnson, 1995), are often intense and frequent but rarely consume everything in their path (Rowe and Scotter, 1973). Owing to variations in fire intensity, together with differences in weather, fuel, species composition, and other environmental conditions (Johnson et al., 1998), wildfires contribute to the creation of a complex mosaic of stands of varying age, composition, and structure (Van Wagtendonk, 2004 and Cullinane-Anthony et al., 2014). This mosaic includes remnants of pre-fire forest ecosystems that partially or entirely escape fire (Ouarmim al., 2015), referred to as residual vegetation patches (which we shorten in this paper to *residuals*). Residuals are conceptually defined as anything from an individual tree to an entire stand that escaped burning, but in practice may include areas that were burned by very low intensity fires (Perera and Buse, 2014).

In Canada, mapping of residuals (Remmel and Perera, 2009) and characterizing their patterns and spatial characteristics (Araya et al., 2015, 2016) have become increasingly important for understanding natural disturbance regimes. Undisturbed residuals continue to provide vital ecological services after disturbances, including the sources of seeds for regenerating local species, acting as refuges for disturbance-sensitive species, establishing habitat, cover, and food sources that are not likely to be available in the surrounding disturbed areas (Burton et al., 2008 and Cullinane-Anthony et al., 2014). Although residual mapping (Remmel and Perera, 2009) provides information after a wildfire or harvesting disturbance event has occurred, additional benefits could be attained if the most likely residual patch locations resulting from a fire disturbance could be predicted using a modelling framework. If only pre-fire landscape conditions were known, predictions of likely residual patches within a larger wildfire disturbance landscape could be used to guide harvest layout planning within the context of emulating natural disturbances (END) or to assess biomass and carbon dynamics. We demonstrate and evaluate a framework for predicting the spatially explicit likelihood of residual occurrence from pre-burn data for the fire-dominated boreal forest landscapes of North-western Ontario and incorporating the complex and interactive effects of physical variables using a machine-learning technique (Munoz and Felicisimo, 2004). While several methods for mapping residuals exist, this has not been

ISSN 1686-6576 / © Geoinformatics International

extensively handled in the fire literature (Perera and Buse, 2014). Studies of boreal wildfires have rather focused primarily on their patterns, mechanistic processes, and their ecological effects (Van Wagtendonk, 2004). For example, the natural wildfire patterns and processes across the forested areas of Alberta and central Saskatchewan (Andison, 2013) and western Quebec (Bergeron et al., 2007) have been studied broadly (Andison, 2013). A detailed set of data have also been collected and analysed to map historical fires and characterize residual vegetation in British Columbia and Alberta (Andison, 2004). The majority of these and other studies have been within areas of active fire suppression and hence exist in non-natural settings (Perera and Buse, 2014). Our study examines natural (lightning ignited and non-suppressed) boreal wildfire events that occurred in North-western Ontario.

The occurrence of residuals is a complex ecological process and factors such as vegetation characteristics, geo-environmental factors, and local weather conditions could influence their formation and resulting patterns. Since their interactions are not linear and may involve multiple processes, a robust model should be designed to examine the predictability of the occurrence of residuals within a fire disturbed landscape. We were determined to develop a spatially explicit model to identify areas where residuals are likely to occur due to physical site characteristics and geo-environmental features. The model is a non-parametric method that addresses the complex spatial dependence and non-linear interactions commonly found in nature. Since the abundance and patterns of residuals could vary as a function of both spatial and temporal scales, the model was also designed to generate the spatial probability of residual occurrence as a gradient of scales. However, the remoteness of the fire event sites hindered the possibility of obtaining local weather data for fire scale analysis, especially at the fine pixel-level of this study. Therefore, the likelihood of residual occurrence was limited to obtainable environmental variables related to topography and land cover variables, particularly those forming firebreak conditions.

There are a limited number of studies that have undertaken to understand the factors that govern residual presence (e.g., Cuesta et al., 2009 and Araya et al., 2015). These suggest that the occurrence of residuals is ascribed to interactive effects of various factors, particularly moist sites, but this work is far from complete. Further, regional controls on the pattern of residuals could be explained by a function integrating vegetation characteristics, geo-environmental factors, and the likelihood of macro- (climate and seasonality) and meso-scale (e.g., rainfall, wind, and temperature) fire weather conditions (Perera and Buse, 2014). Variations in measurement and analytical scale (e.g., spatial resolution) among studies further limit comparisons among them. To address these limitations and to improve our understanding of physical site characteristics that influence the presence of residuals, we propose an approach based on Random Forests (RF) that can be applied broadly to predict residual presence likelihood. Our spatially explicit model prediction results are readily interpreted in cartographic format and comparable with other studies across multiple spatial resolutions.

Machine-learning methods, such as RF, provide a framework for identifying the important underlying variables for building robust predictions and for exploring mechanistic relationships within models (Breiman, 2001 and Evans et al., 2011). Unlike traditional statistical methods, machine-learning methods are non-parametric models and do not assume a predefined distribution or hypotheses; they handle the complex, non-linear, and multi-dimensional interactions among variables that are common in ecological data (Evans et al., 2011 and Waljee al., 2013). In this context, our goal is to internally validate a spatially explicit and readily repeatable predictive method implemented in RF using a bootstrapping approach. We then externally validate the performance of our model by producing spatially explicit likelihood maps of residual occurrence for an independent wildfire event using pre-fire conditions and then overlay those results with the map of actual residuals that remained following the fire to assess the accuracy of those predictions.

2. Methods

2.1. Study Area

We studied twelve natural wildfires situated in two different ecoregions in North-western Ontario. All twelve wildfires burned within the Ontario boreal shield ecozone (Hills, 1961), which is characterized by relatively long, cold and dry winters, and short (and warm), and moist summers (Thompson, 2000). Eleven of the wildfires, having footprint areas ranging from approximately 58 to 4225 ha, are identical to the ones examined by Remmel and Perera (2009) and Araya et al., (2015, 2016) and are located within the boreal region of Ontario (Figure 1). These 11 wildfires were all ignited by lightning and were not suppressed; they form the data domain for training the predictive model and for performing the internal validations.

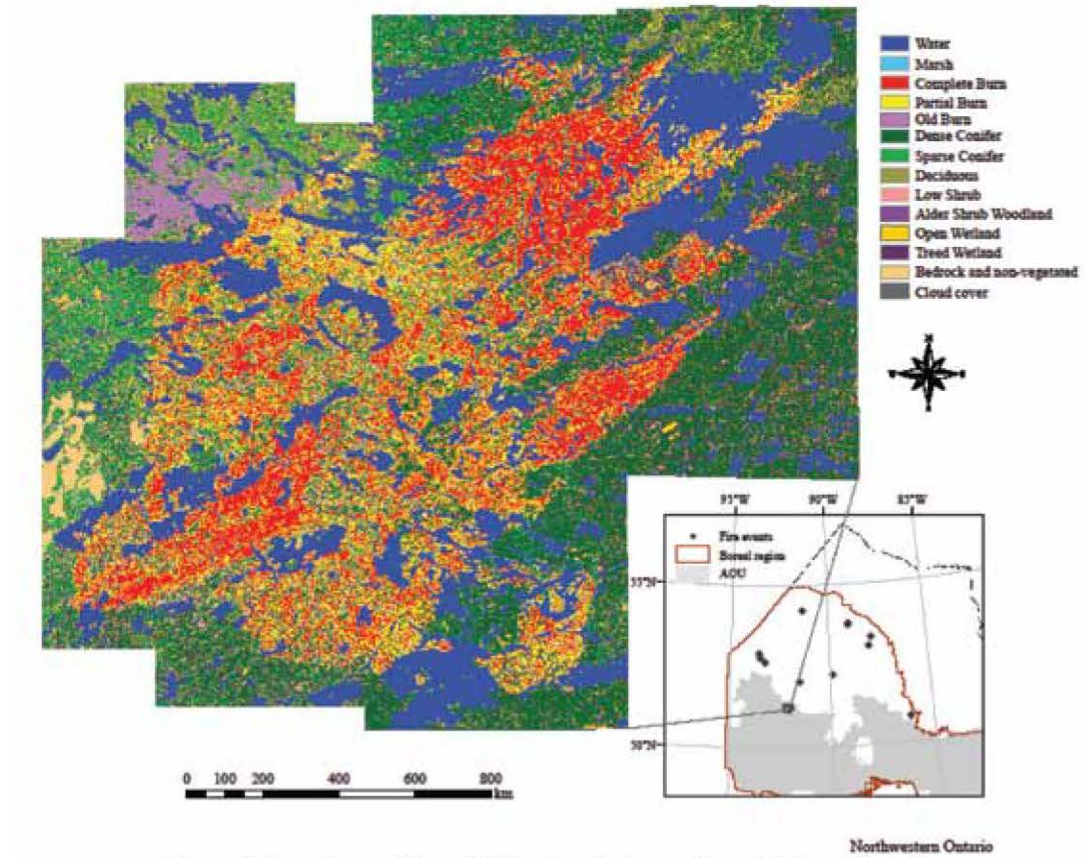

Figure 1: Locations of the wildfires in relation to Ontario's boreal region

A twelfth wildfire event (RED084) with a footprint of 54,828 ha (containing 14,591 ha of residual vegetation) was used for externally validating the model.

The eleven wildfires used for training occurred among the 2001 to 2003 fire seasons and occurred within the Ontario Ministry of Natural Resources (OMNR) extensive fire management zone where fires are monitored and recorded, but not actively suppressed (Perera et al., 2009). Unlike the training wildfires, RED084 burned in the Red Lake area in July 2011 and is situated within the area of undertaking (AOU) where forest management practices occur. Minor areas of this fire event had previously burned or been harvested and fire suppression efforts were minimal relative to the extent of the footprint. RED084 was selected to independently validate the performance of our model because: 1) the fire was ignited naturally by lightning, 2) despite suppression efforts, the fire burned until extinguishing naturally, 3) the event is contained within a similar ecological setting as the training fires, and 4) imagery from the same satellite was available for both training and validation sites, which facilitated our assessment.

2.2 Data

Our model was based on presence–absence data, with residuals forming the presence data. Presence data (residuals) were determined from classified Ikonos imagery as described in Remmel and Perera (2009). The mapping of residuals follows the Ontario Ministry of Natural Resources' guidelines, with residuals defined based on their size and location in relation to the perimeter of a disturbed landscape (OMNR, 2014). Residuals are clusters of unburned pixels, with an area greater than 0.25 ha of a treed land cover class at a distance greater than 1 pixel inset from the fire perimeter (OMNR, 2014); note that the spatial resolution of the data alters the definition for identifying residuals. Training data for our model was originally classified from Ikonos imagery acquired between 1 June and 7 August 2005; the residuals used as the validation data were derived from Ikonos imagery captured between 30 October 2011 and 12 July 2012. In each case, a supervised maximum-likelihood classifier was used to produce post-fire vegetation maps with different feature classes such that residuals would be extracted and spatially aggregated at 4, 8, 16, 32, and 64 m spatial resolutions (R_4, R_8, R_{16}, R_{32}, and R_{64} respectively) (Remmel and Perera, 2009 and

Araya et al., 2016). Our study examined residuals from eleven wildfires in northern-western Ontario, and we follow the works of Araya et al., (2016) to categorize each of the eleven training wildfires into one of three groups based on footprint area: small fires (≤ 260 ha), medium fires extend to < 3000 ha, and large fires ≥ 3000 ha (identified by a natural breakpoints in the data).

Since information about absence data was not readily available, a spatial sampling algorithm (Araya et al., 2016) was used to randomly identify areas that could be considered absence data (we describe them in this paper as *null-residuals*), mimicking the residuals in size, shape, orientation, and frequency of occurrence. Internal validation was based on data from the eleven wildfires used for training and external validation used the RED084 wildfire that was kept separate from the training of the model and from internal validation, such that it could be used afterward for independent external validation.

The model tested the interactive effects of variables selected based on *a priori* ecological assumptions (i.e., that residual patch occurrence is explained by vegetation characteristics and geo-environmental conditions). While published literature indicates that the occurrence of residuals results from local wildfire behaviour in response to environmental conditions, limited availability of such data for the study area limited the types of variables that could be utilized in the predictive model. Thus, we rely on the combined and interactive effects of the topographic variables (ruggedness index – RI, slope – SL, and elevation – EL), natural firebreak features (distance to wetlands – WL, water – WA, and bedrock and non-vegetated areas – BV), and land cover class – LC to predict the likely occurrence of residuals. RI quantifies the relative elevation difference between a central cell and its eight nearest neighbouring cells as presented on a digital elevation grid as shown in Equation 1 (Riley et al., 1999).

$$RI = Y\left[\sum (x_{ij} - x_{oo})^2\right]^{1/2}$$

Equation 1

x_{oo} is the elevation of the centre cell; x_{ij} are the elevations of each of the neighbouring cells whereas Y indicates the cell size.

Distance to wetland – WL and distance to water – WA variables are computed as Euclidean distances to the nearest wetland or water body cell respectively. Our land cover class was based on pre-burn conditions extracted from the Ontario Land Cover Data Base that is derived from Landsat imagery acquired between 1999 and 2002, mostly from 2000 onward. Weather variables (wind, temperature, and precipitation) that may contribute in the ignition and spread of wildfire were not incorporated in our study since they were not available at the required spatial scales for this remote region.

2.3 Model Construction and Calibration

Spatial datasets were converted into tables with 8 columns; seven predictor variables and a binomial response variable (i.e., residuals – null residuals) and rows representing individual pixels. This data was fed into a RF predictive algorithm implemented in R (R Core Team, 2015). The RF technique is an ensemble learning technique developed out of CART and boosting and bagging approaches that generates numerous classification trees that are subsequently aggregated to produce a single classification. The RF algorithm begins by generating multiple bootstrapped samples and then building a set number of un-pruned classification trees for each bootstrapped sample, classifying both presence and absence based on a suite of geo-environmental factors (Breiman, 2001). While two-thirds of the bootstrapped samples are used for constructing a classification tree, the remaining one-third, referred to as the out-of-bag sample (OOB), are used for computing accuracies and error rates. Each classification tree assigns a class (i.e., residual or null-residual) to each unique combination of predictor variable decision sequences and a majority vote across all randomized trees defines the resulting response class for that pixel.

RF implementation in R requires two parameters to be set: the number of trees in the forest (n_{tree}) and the number of variables (m_{try}) in the random subset at each node used to split the nodes (Breiman, 2001). The sensitivity of RF algorithm to the parameters (n_{tree} and m_{try}) based on the same dataset (training data), given the physical variables, was assessed in Araya et al., (2016). The study suggested that the error estimates showed stabilization of the overall error across a suite of n_{tree} and m_{try} values; adding more trees to the model did not substantially change the error rate (mean: 9.71%, standard deviation: 0.10). Thus, the model was constructed and calibrated with parsimonious options n_{tree} = 100 and m_{try} = 3. The model was constructed and calibrated (and internally validated) using the training data by splitting the data into *training* and *test* groups using a hold-out approach. The procedure undertaken to partition the data, and construct and calibrate the model is graphically summarized in Figure 2

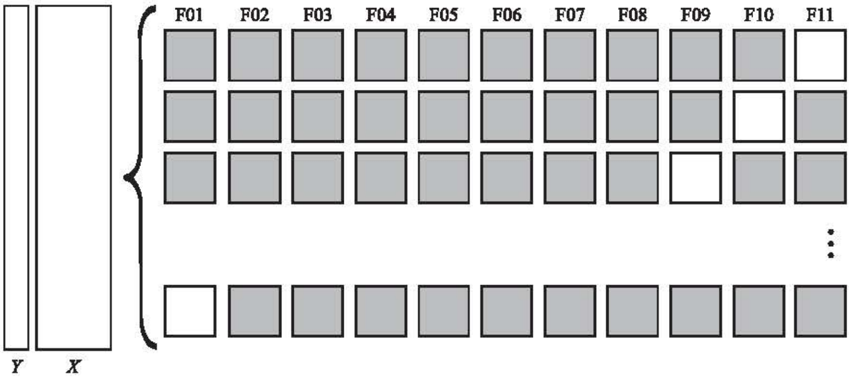

Figure 2: Overview of the data partitioning (hold-out approach for individual fire events) that was implemented for calibrating the model. Allocation of events into *training* (grey) and *test* (white) sets. *Y* refers to the response variable (the likelihood of a pixel being a residual) while *X* indicates the physical variables from the training data

The prediction by RF produces either deterministic response classes or probabilities of class membership (i.e., matrix of class probabilities). Since most of the environmental attributes are inherently continuous, classifying spatial elements into binary classes yields an overly simplistic representation of the landscape (Evans and Cushman, 2009).

We presented the presence of residuals as a separate probability surface, determined at the individual pixel level, rather than a mosaic of discrete patches that are implicitly assumed to be categorically discrete. The result from the model present predicted probability values scaled from 0 to 1 for each grid cell, with predictions closer to 1.0 indicating greater chance of residual vegetation occurrence. Contiguous cells with high probability of residual occurrence can be grouped and considered as a residual; where the probability threshold needs to be decided by practitioners and will be context sensitive. In order to examine the threshold values that determine the probability of residual occurrence, and compare the observed with the expected values, arbitrary threshold values were selected to convert the predicted probability values into a binary surface.

2.4 Model Validation

The inherent uncertainty of a predictive model needs to be validated to determine its suitability for a specific application and to compare different modelling techniques (Pearce and Ferrier, 2000 and Beauvais et al., 2006). Several methods have been proposed, but model validation is generally performed using internal or external validation. Our first approach, internal validation, was done by splitting the study data set into training and test groups that are respectively used to build and validate a model; this was a leave-one-out method, where a dataset from a single wildfire (e.g., F01 from the training data) is used for testing while the remaining datasets (from all other fires) were used for training the model. Each iteration of testing proceeds with a different event being left out for the internal testing. Next, we tested model validity against an external data set that has never been used in building the model. Since internal validation may overestimate model performance, our approach was also validated externally to provide an independent assessment of model performance. To achieve external validation, our model was built with the training data (eleven wildfires) and validation was performed with data from the RED084 wildfire event.

There are two approaches for evaluating model performance: (1) use a fixed-probability threshold for deciding residual versus null-residuals, or (2) use a threshold-independent measure, that can vary based on the data (Jepsen, 2004 and Beauvais et al., 2006). Both methods translate the predicted probability into a binary (0 or 1) class; the fixed-probability threshold is based on a single value (often 0.5 is considered as a default) while the latter method adjusts to the data. The selection of an appropriate threshold value is difficult, often arbitrary, and will affect measures of a model's performance. A threshold-independent measure, such as the Receiver Operating Characteristics (ROC), based on a broad spectrum of threshold values (Fielding and Bell, 1997, Pearce et al., 2001 and Peters et al., 2007) was implemented in our study to assess the predictive performance of the model. The ROC curve provides a graphical depiction of a model's discrimination ability over a range of threshold values (Zweig and Campbell, 1993 and Pearce and Ferrier, 2000).

Predicted probability values derived from the RF model were used to generate a 2×2 contingency table, one for each threshold value. A threshold value represents a probability above which a cell is considered to contain residual vegetation (Peters et al., 2007). A threshold interval of 0.01 is used to convert the predicted values into a binary category (0 and1), over 100 threshold values evenly spaced across the range of available predicted values (from 0.0 to 1.0). For each threshold of the continuous probability scale, a 2×2 contingency table was created from which sensitivity and specificity were computed (Figure 3).

		Actual	
		Present	Absent
Predicted	Present	a	b
	Absent	c	d

Sensitivity	Specificity
a / (b + c)	d / (b + d)

Figure 3: Metrics of residual validation: a hypothetical 2×2 contingency table used as the base for estimating sensitivity, specificity, ROC, and AUC

Sensitivity refers to the proportion of presence (i.e., residuals) correctly predicted while specificity is the proportion of absence (i.e., null-residuals) correctly classified. A ROC curve, a plot of sensitivity on the *y-axis* against (1 – specificity) on the *x-axis* for varying threshold values is produced. Since comparing ROC curves directly is not easy and can be highly subjective (Eunsik and Wenbao, 2011), a single index, the Area Under the Curve (AUC), is computed as it describes the discrimination ability of a model (Zweig and Campbell, 1993 and Pearce and Ferrier, 2000). We produced ROC curves for each fire event and spatial resolution combination within our study along with the AUC for each curve using the *AUC* package in R. The AUC values serve as indicators of the model's ability to distinguish between residuals and null-residuals, independent of a specific threshold value.

2.5 Spatial Prediction

A model output can be presented as an abstract formula or in the form of qualitative description, but the ability to generate predicted probability maps enable us to examine the spatial distribution of elements and the continuous nature of residual occurrence in a disturbed landscape (Beauvais et al., 2006). Our predicted probability values, ranging from 0 to 1, were converted into digital or GIS readable format maps using the *RSAGA* package in R. The output maps identify the probability of residual occurrence at each grid cell of a fire event footprint.

3. Results

A pair-wise correlation analysis to assess the relationship among pairs of physical variables (topographic, firebreak features, and land cover) was undertaken and identified no redundancies in the variables (Araya et al., 2016). A similar study has examined the relative importance and marginal effects of each of the variables that has been used for spatial prediction in the present study (Araya et al., 2016). Prior to computing the ROC curves and their respective AUC values for the test data, the discriminatory power of the model in relation to the variables was also examined visually by comparing the variability of probability values across five spatial resolutions for the training data (Figure 4). The results show that the median probability values for residual cells are relatively higher than those for null-residual cells; this difference is most dramatic at R_4. The model's inability to distinguish between residual and null-residuals was also evident in some of the wildfires (e.g., F06 and F07) where the median values for residual and null-residuals are very similar.

Our model's predictive performance was initially assessed using selected fixed probability thresholds, cut-off values used to translate the predicted probabilities into a binary class (0 and 1) where the default threshold is often set as 0.5 (to isolate cases performing better than random chance). This involves overlaying deterministic probability map with existing residuals that remained following the fire. The choice of an appropriate threshold values is difficult, often arbitrary, and affects measures of a model's performance. We present fixed-accuracy measures using arbitrary threshold values (0.2, 0.4, 0.5, 0.6, and 0.8), one for each of the three class sizes (large, medium and small fire events), and their respective overall accuracy at R_4.

The sequence of threshold values were selected on an interval to flank the 0.5 mid-point, to help identify a reasonable threshold. The overall accuracy for our internal validation at R_4 always exceeds 75% for the selected fire events across different cut-off values while low accuracy (less than 35%) was attained for our external validation (RED084) at 0.2 and 0.4 threshold values (Table 1). Yet, an overall accuracy of greater than 50% was observed for Red Lake fire event for the remaining threshold values.

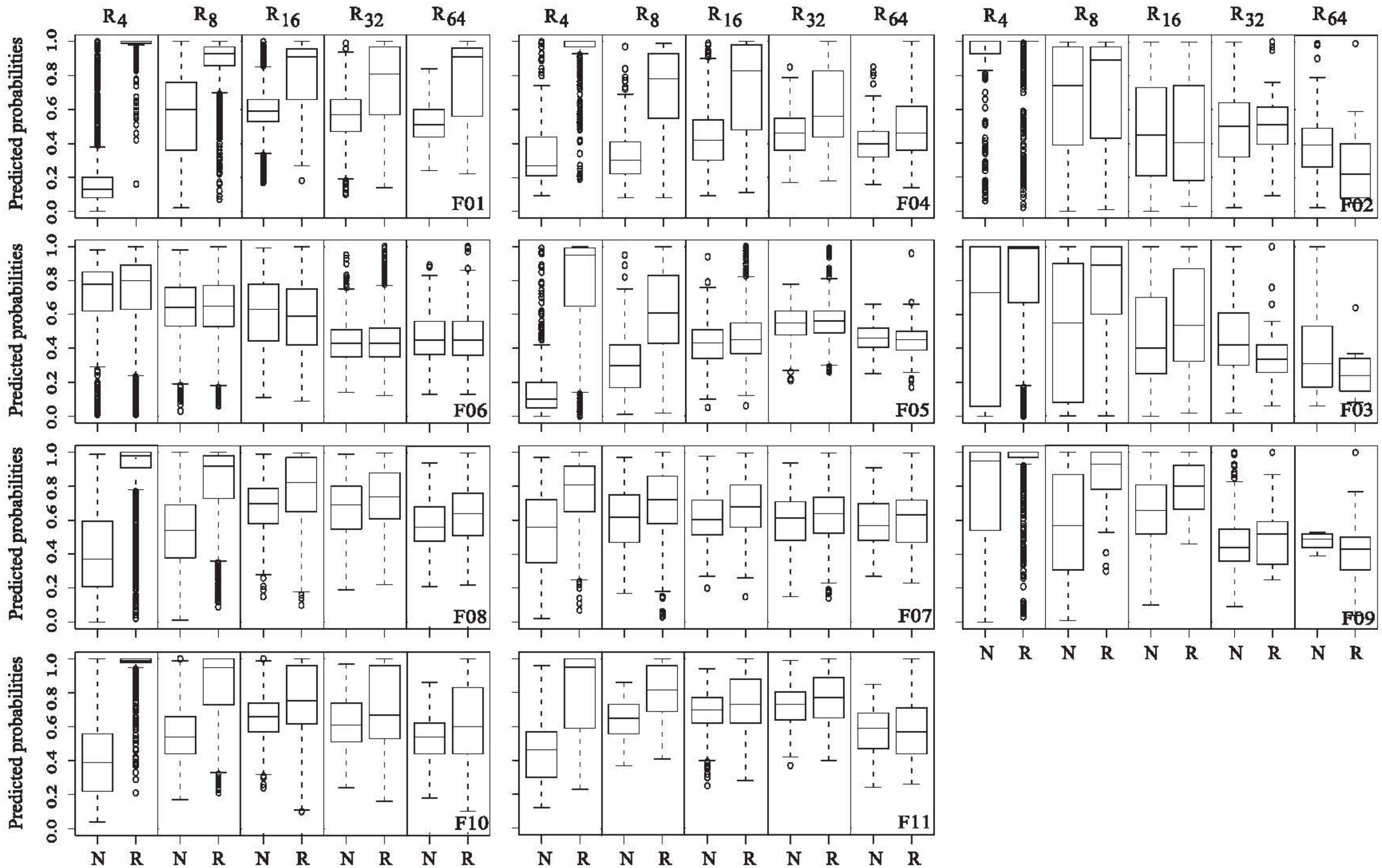

Figure 4: Model's discrimination ability to distinguish between residual and null-residuals: variability of predicted probability values associated with residual (R) and null-residual (N), for wildfires from the training data

Table 1: Predictive performance and overall accuracy of the RF model using fixed-thresholds (OA – Overall accuracy), fixed accuracy measures were set using arbitrary threshold values (0.2, 0.4, 0.5, 0.6, and 0.8)

			Actual							
			F01		F04		F02		RED084	
Threshold			Present	Absent	Present	Absent	Present	Absent	Present	Absent
0.20	Predicted	Present	2.66	0.52	1.01	0.28	0.03	0.00	1562.44	9136.97
		Absent	0.00	1.15	0.00	0.07	0.00	0.00	19.93	91.06
		OA (%)	88.0		79.1		90.5		15.3	
0.40		Present	2.66	0.19	1.00	0.10	0.03	0.00	1172.68	6783.96
		Absent	0.00	1.48	0.02	0.25	0.00	0.00	409.69	2444.08
		OA (%)	95.6		91.4		90.5		33.5	
0.50		Present	2.66	0.15	0.99	0.06	0.03	0.00	872.12	4000.11
		Absent	0.00	1.52	0.03	0.29	0.00	0.00	710.26	5227.92
		OA (%)	96.4		93.2		90.5		56.4	
0.60		Present	2.65	0.12	0.97	0.04	0.03	0.00	580.71	1816.27
		Absent	0.01	1.55	0.05	0.31	0.00	0.00	1001.66	7411.76
		OA (%)	96.9		93.9		90.5		73.9	
0.80		Present	2.64	0.06	0.92	0.02	0.03	0.00	332.30	435.98
		Absent	0.03	1.61	0.09	0.33	0.00	0.03	1250.07	8792.05
		OA (%)	98.1		92.0		100.0		84.4	

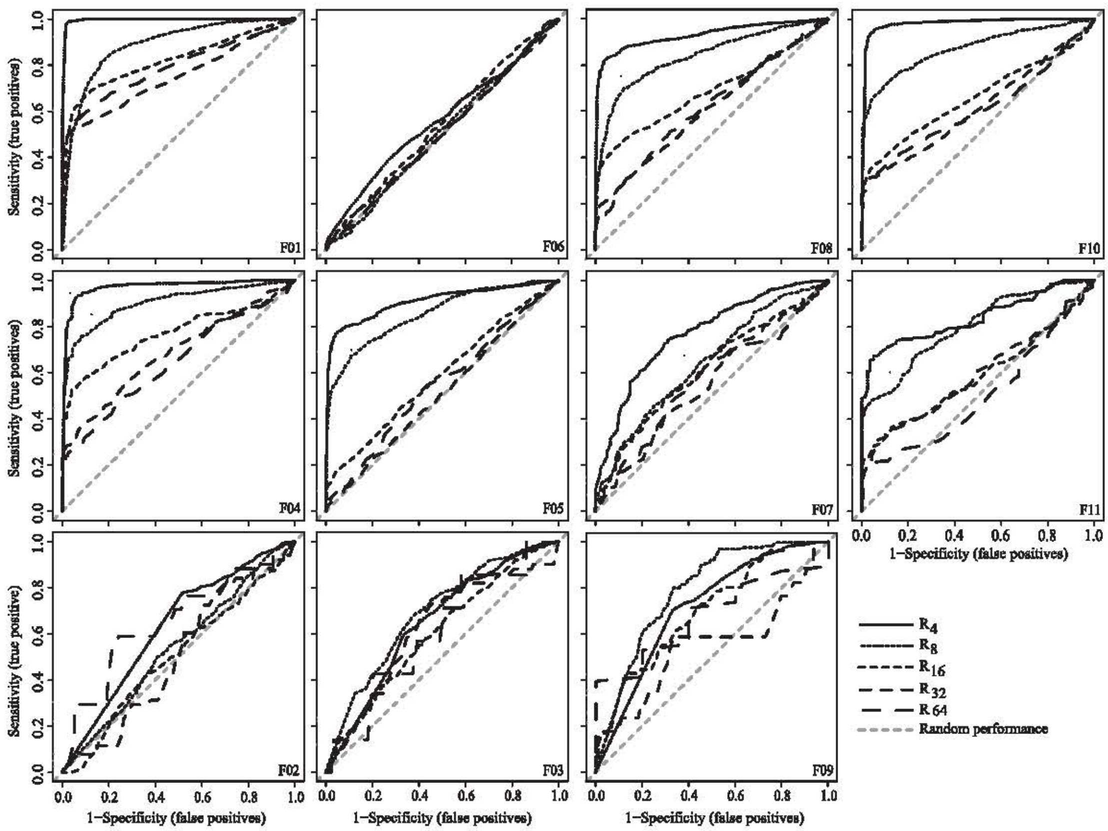

Figure 5: Graphical representation of the predictive performance of RF model for the eleven wildfires from training data; each ROC curve in the plot depicts model's performance at specific spatial resolution

To further examine model performance, the ability of our model to discriminate residuals was assessed using an internal validation and respective ROC curves; the results are graphically summarized in Figure 5. A model that perfectly predicts residuals generates an ROC curve that follows up the left axis and right across the top axis of the plot, while a model with random predictions produces a ROC curve that follows a 45° diagonal from the lower left corner to the upper right corner (Zweig and Campbell, 1993, Fielding and Bell, 1997 and Pearce and Ferrier, 2000).

Plots generally higher in the upper triangle have greater observed accuracy (Zweig and Campbell, 1993); such a trend is evident in Figure 5 with changing spatial resolution where curves for some events (e.g., F01, F04, F05, F08, and F10) approached near perfect discrimination at R_4. The AUC provides a summary measure of a model's predictive accuracy, since the area under the curve will increase to 1.0 when the ROC for a perfect prediction occurs, but more generally a ROC curve characterized by a large AUC will be more accurate than if the AUC is small (Pearce and Ferrier, 2000). As a general rule, the AUC values range from a random guess (AUC ≤ 0.5), low accuracy (0.5 ≥ AUC < 0.7), reasonable accuracy (0.7 ≥ AUC < 0.9), and high accuracy (AUC ≥ 0.9) (Swets, 1988).

The results of our internal validation, based on the training data, showed that the model had the highest accuracy for F01, F04, and F05 with an AUC of 0.995, 0.970, and 0.910 respectively at R_4 (Table 2), suggesting the model could discriminate between residual and null-residuals more than 90% of the time. At R_4, the model also produced AUC > 0.7 for all fire events except F02, F03, and F06. Based on the general rule by Swets (1988), the ability to discriminate at this scale was evaluated as being between reasonable and high accuracy. The lowest accuracy was observed for F02 and F06, with AUC values of 0.629 and 0.563 respectively, but still better than random chance. The statistical significance of AUC values, compared with a random prediction, was computed using a Wilcox test as implemented in R. Based on this test, our predictive model had significantly higher discrimination ability ($p < 0.05$) than random prediction (i.e., AUC = 0.5) for all the wildfires at R_4.

Table 2: AUC values for each predictive model treatment, representing the 12 wildfires at 5 spatial resolutions, reflecting Swets (1988) rules: light grey (high accuracy) to darker grey (random prediction)

		Spatial resolutions				
Size Class	**Fire ID***	**R_4**	**R_8**	**R_{16}**	**R_{32}**	**R_{64}**
Large	F01	0.995	0.886	0.816	0.749	0.793
	F06	0.563	0.506	0.544	0.503	0.500
	F08	0.933	0.844	0.685	0.605	0.613
	F10	0.981	0.874	0.659	0.616	0.611
Medium	F04	0.970	0.902	0.771	0.688	0.643
	F05	0.910	0.854	0.584	0.512	0.546
	F07	0.770	0.642	0.611	0.555	0.567
	F11	0.837	0.799	0.588	0.590	0.503
Small	F02	0.629	0.537	0.507	0.507	0.648
	F03	0.647	0.688	0.601	0.622	0.611
	F09	0.710	0.786	0.699	0.551	0.677
Independent	RED084	0.571	0.583	0.572	0.615	0.617

*** The training wildfires are categorized into three groups (large, medium, and small) while RED084 was kept separate for independently validating the model**

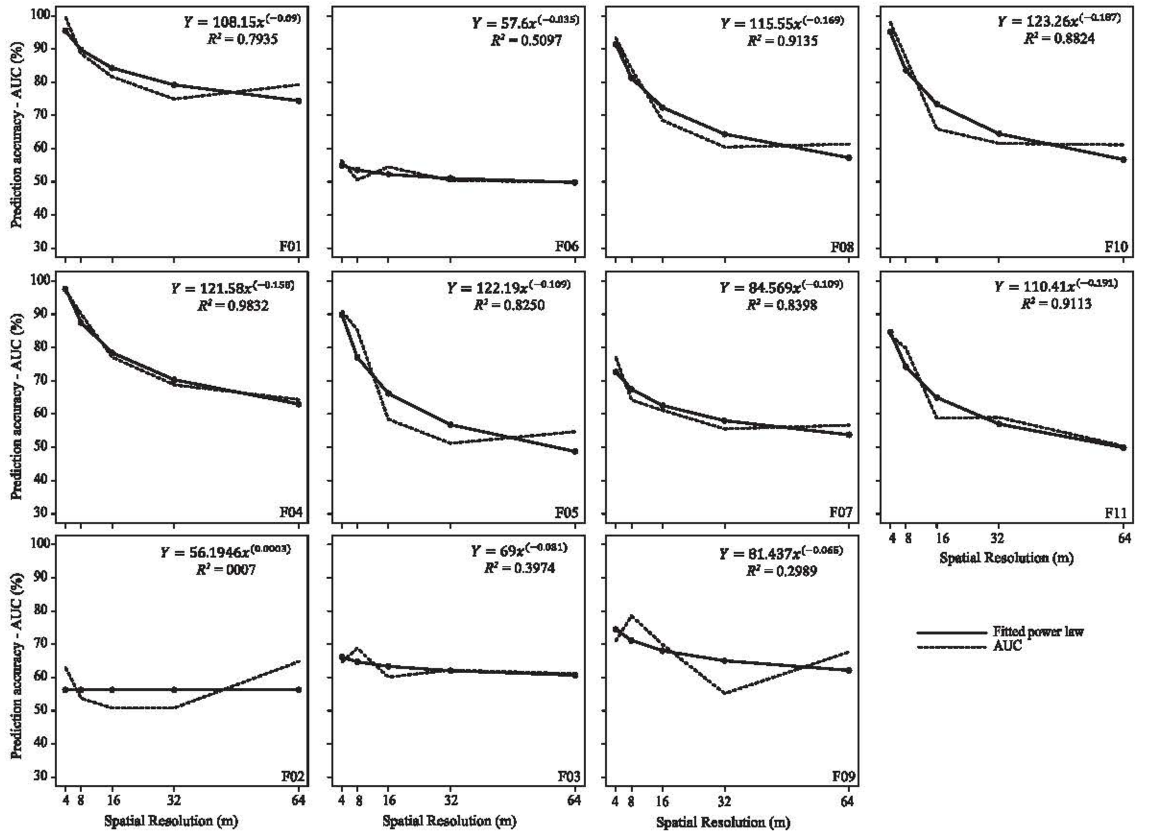

Figure 6: The predictive performance of the model as a function of scale: the dashed line in each plot shows the AUC values computed at different spatial resolutions while the solid line is the best fit model

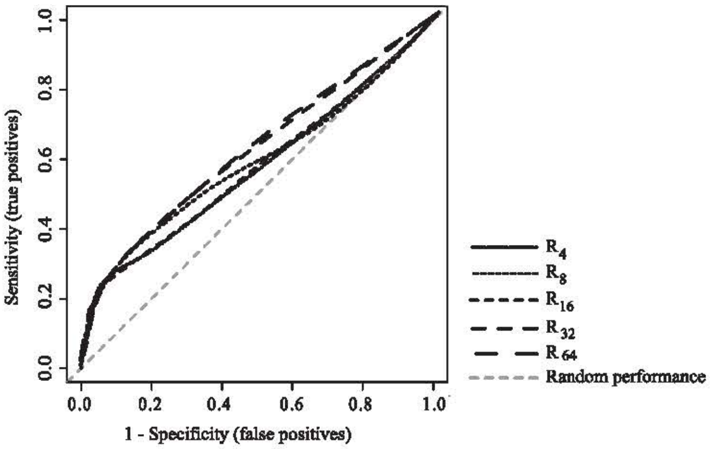

Figure 7: Graphical representation of the predictive performance of RF models for the validation wildfire (RED084); each ROC curve in the plot depicts model performance at specific spatial resolution

At R_8, the model also produced a higher accuracy measure with an index of 0.902 for only one fire event (i.e., F04) while a reasonable accuracy was obtained for the remaining fire events except for F02, F03, F06, and F07. At coarser spatial resolutions such as the one at R_{16}, the model produced low accuracy measure for most of the wildfires with an index value ranging between 0.507 and 0.669; yet a reasonable accuracy was attained for F01 and F04, with AUC index values of 0.816 and 0.771 respectively. The model's ability to discriminate residuals from non-residuals at R_{32} and R_{64} was generally low with AUC values ranging from 0.500 (equal to random prediction) for F06 at R_{64} to 0.688 for F04 at R_{32}.

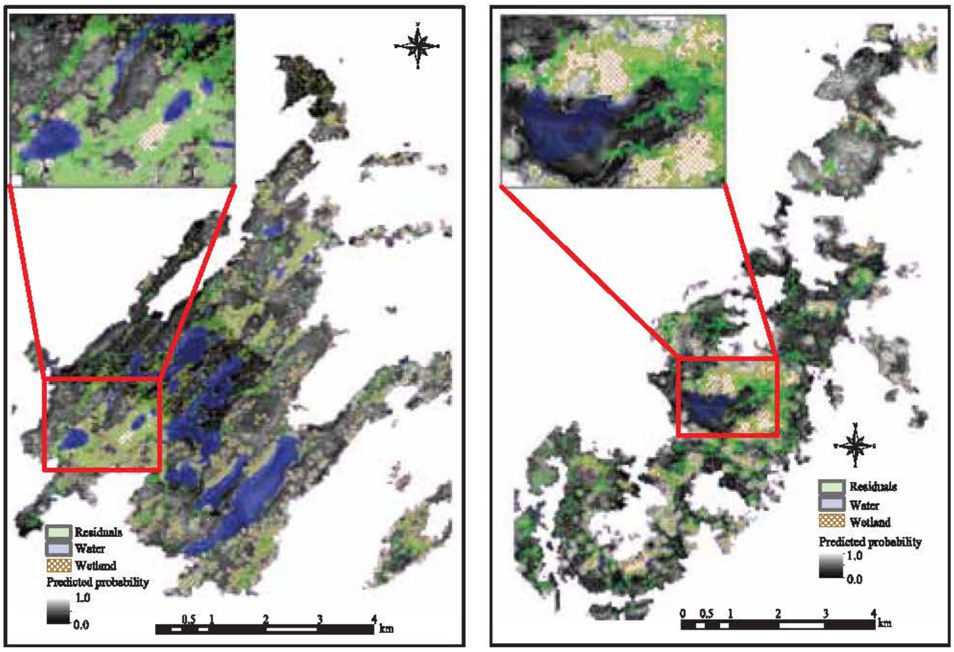

Figure 8: Maps of residual vegetation occurrence probability for selected wildfires from training fire events F01 and F10 at R_{32}. Lighter shading indicates a greater probability of residual occurrence, cross-hatched areas display wetlands, and light blue represents open water land cover

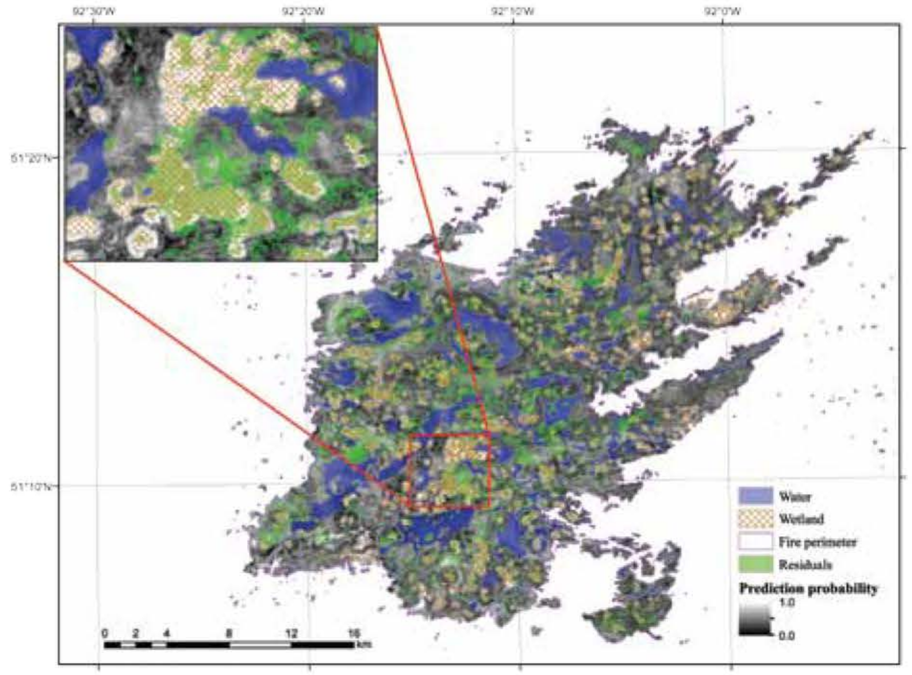

Figure 9. Maps of residual vegetation occurrence probability for the independent external validation (RED084) fire event at R_{32}. Lighter shading indicates a greater probability of residual occurrence, cross-hatched areas display wetlands, and light blue represents open water land cover.

The predictive performance of the model was assessed at different spatial resolutions to determine whether residual cells could be modelled more successfully at specific scales. A comparison of the RF model across scale gradient revealed that some distinctive patterns in model performance were observed as a function of spatial resolution. At the finest spatial resolution, R_4, the model provided better discrimination, but the predictive accuracy decreases with increasing spatial resolution size in a consistent power-law relationship for large and medium size wildfires (except for F06), with R^2> 80% (Figure 6). Such a monotonic decreasing trend with a power-law relationship was not observed across the three small wildfires: F02, F03, and F09. We also implemented an external validation using an independent validation dataset. The ROC curves representing the predictive performance of our models at R_4, R_8, R_{16}, R_{32}, and R_{64} are graphically presented in Figure 7. The curves lie between the two extremes of perfect discrimination and random prediction, but overall the model had low accuracy with an index value that lies within the range of 0.5 and 0.7 across all the spatial resolutions (Table 2). Yet, the AUC values found to be significantly better than that expected from a random model ($p < 0.05$). While the statistical validity and accuracy of a model is important, graphical representation of the model output (i.e., spatially explicit maps of residual occurrence prediction) facilitate examination of the spatial distribution of spatial elements (Fielding and Bell, 1997). The maps of predicted probable residual vegetation occurrence were constructed for all wildfires (training and independent events) at each of the five spatial resolutions. However, due to space restrictions, we only present the probability maps for selected wildfire events representing the training data (Figure 8) and validation data (Figure 9) at R_{32} as this is the closest spatial resolution to how Landsat-type remote sensing sensors would represent landscapes. The model output is a probability value scaled from 0.0 to 1.0 for each grid cell, with predictions closer to 1.0 indicating greater chance of residual vegetation occurrence.

4. Discussion

The interactive effects of seven physical variables (ruggedness index, slope, elevation, distance to wetlands, distance to water, distance to bedrock and non-vegetated areas, and land cover class) were related with a binomial response variable (i.e., presence–absence of residual) using predictive models implemented in RF. Weather variables such as wind, temperature, and precipitation that may affect the patterns of wildfire and post-fire forest structure were not considered in this study as they were not readily available at the required scale for the remote region of north-western Ontario. Our modelling framework provides local error estimates, which are indicative of model fit, but not necessarily of predictive performance of the model. Given the importance of predictive models, continuous and progressive evaluation of models with external validation is necessary, as effective and correct model assessment has real significance to ecological studies (Pearce and Ferrier, 2000, Manel et al., 2001 and Austin, 2007). We chose to perform both internal and external validation to provide statistically independent measures of model performance. The results of our fixed-probability threshold, specifically overall accuracy, indicated that the model did a reasonable job to discriminate residuals from null-residuals at R_4. The overall accuracy from fixed-probability threshold has been widespread in determining classification accuracy, but Manel et al., (2001) argued that it is inherently misleading. We then considered threshold independent approach to further assess the predictive power of our model. We found that internally validated estimates of model performance was generally well supported and significant at $p<$ 0.05 for the finer spatial resolutions (R_4, R_8, and F_{16}), across the majority of the training wildfires. The results suggest that the model has a reasonable to high predictive accuracy at finer spatial resolutions and that the accuracy decreases generally with coarsening spatial resolutions.

Our internal validation showed that the model had low predictive performance for certain wildfires within the training data (i.e., F02, F03, F06, and F09), however three of those observations (F02, F03, and F09) are for the smallest wildfires, and one plausible reason for low prediction accuracy could be attributed to small sample size. The extent of the small wildfires was less than 100 ha; with less than 10 residuals. Likewise, Pearce et al., (2001) found that the predictive performance of rare cases, with less than nine records, was poorer than those with large numbers of records. Similarly, Schwartz et al., (2006) claims that too few observations results in low statistical power for prediction. Our finding supports the argument that predictive models usually attain more accurate predictions with increased sample size (Edwards et al., 2007). Despite the low accuracy, the performance of our model (for F02, F03, and F09) was statistically significant at R_4 and F_8; the model performed significantly better than random prediction. Furthermore, our internal model validation had low predictive accuracy for one of the largest wildfires (F06) within training data. One potential reason could be the differences within the landscape as a function of the environmental attributes and burn severity. There could be inter-regional or inter-landscape differences among the disturbed landscapes (Burton et al., 2008). Another reason for low predictive accuracy at F06 could result from the distribution of un-burnable cover types that determine residual occurrence within a fire footprint. A variable importance assessment conducted by (Araya et al., 2016), on the same dataset, indicated that wetlands are the most important predictors to explain residuals for all large wildfires except for F06; at R_4, wetlands form less than 0.1% of the F06 fire footprint. Therefore, for a model to predict spatial elements with an excellent ability to discriminate residuals from non-residuals (e.g., a strong model), the parameters that determine the response variable (residual vegetation occurrence) should occupy a sizable portion of the fire footprint.

The variation in predictive performance of the model with the dataset from the training data, over multiple-scales, suggested that a scaling rule might be determined for wildfires (i.e., F01, F04, F05, F07, F08, F10, and F11) yielding reasonable to high accuracy measures. In this instance, the variability observed in AUC values as a function of spatial resolutions produced an R^2 ranging from 80% to 95% (Figure 6). In situations such as F02, F03, F06, and F09 where low predictive accuracy was yielded across the spatial resolutions, a robust scaling rule could not be established given the predictor variables, with R^2 less than 40%. External validation is a more rigorous procedure for evaluating a model's performance (Taylor et al., 2008); hence we stressed the need for the external data to be completely independent of building the model. Given the environmental variables considered, external validation yielded estimates of model performance that were lower than those initially reported using internal validation, with AUC values less than 0.6 across all of the spatial resolutions. Given the interregional and inter-landscape differences within the boreal forests as a function of climate and topographic parameters, the results are not surprising. This reflects the idea that modelling approaches developed and applied even in relatively data-rich regions may not necessarily work effectively elsewhere (Ferrier et al., 2002). It was also argued that every wildfire represents a unique combination of fire behaviours that affect forest species and habitat features in the canopy differently (Burton et al., 2008). Regardless, our model does improve the prediction of residual vegetation more accurately than by simple random assignment, and it does so with coarse data for a remote region for which detailed local fire-weather data are not available (i.e., at the spatial resolution of the analysis).

Predictive models, based on presence and absence data, are increasingly used in ecology, particularly for conservation planning (Manel et al., 2001) and effective monitoring programs (Magness et al., 2009), Residuals are generally treated as categorical presence–absence classes in a discrete mosaic (Evans and Cushman, 2009), rather than a continuous probability surface of presence.

5. Conclusions

We developed a spatially explicit predictive model in which the presence of residual vegetation can be determined. Although predictive models can provide objective estimates of the presence of residuals, the challenges of validating a RF-based predictive model are considerable, and attention must be given to the validation approach in the construction of the model. Our study was also aimed at validating the predictive performance of the model in relation to the combined effects of seven physical variables using both internal and external validation. Internal validation (bootstrapping) is more convenient and perhaps yields an overestimated accuracy of prediction, but for a predictive model to be valid and effective, external validation using a dataset from an independent site was considered for further assessment of performance. Our approach provides a consistent and replicable method for learning complex and non-linear ecological relationships and subsequently using those models to predict residual vegetation occurrence in remote locations where detailed weather data burning data do not exist.

A model with an ability to discriminate residuals is the one that correctly distinguishes between residual and null-residuals in the evaluation dataset, irrespective of the reliability of the predicted probabilities. The overall accuracy from fixed-probability threshold yielded a reasonable accuracy to discriminate residuals from null-residuals for selected threshold values at R_4, while low accuracy was attained for our external validation. Since fixed-probability measure has been criticized for providing inherently misleading output, we opted to assess the model's accuracy with threshold independent measures. We found that internally validated estimates of model performance were generally well supported; with AUC values greater than 0.7 at R_4 for the majority of wildfire events while the predictive performance of the model decreased with coarsening spatial resolution in a consistent power-law relationship, with a coefficient of determination >80%.

The ability to generate predicted maps was also one of the goals of our study; hence we produced spatially explicit probability maps that identify locations where residual vegetation is likely to occur. We found that the physical variables considered in our study did a reasonable job to explain residual occurrence. While the model was able to identify areas of potential residual occurrence with an overall accuracy ranging from 68% (F06) to 97% (F01) at R_4, high prediction probability is specifically associated with wetlands. We also inferred that the performance of the predictive model, even with internal validation, that results vary as a function of sample size (number of training residual records) and the spatial resolution. Larger sample sizes and the finest spatial resolutions (e.g., R_4 and R_8) produced highest predictive accuracy. When applied to an independent dataset, the performance of our model was generally lower than the performance observed in the internal validation across the gradient of scales, with AUC values ranging from 0.5 (random prediction) to 0.7. Yet, the results were better than random prediction and the model was able to identify areas of potential residual occurrence in the independent dataset; the predicted residuals coincide with areas where residuals actually formed with an overall accuracy of 55%. This suggests the potential of the variables (and the model itself) to explain residual vegetation occurrence on coarse and sparse data at a rate better than random chance alone.

While our model produces reasonable prediction accuracy across a gradient of scales, the likelihood of residual occurrence was limited to obtainable environmental variables (topography, firebreak features, and land cover variables). Unfortunately, local weather data (wind, temperature, and precipitation) that are known to contribute significantly to wildfire ignition success and spread do not exist at high spatial resolution for this remote location. The performance of our model would likely benefit with the incorporation of finer spatial resolution and local fire weather data.

References

Andison, D. W., 2004, Island Remnants on Foothills and Mountain Landscapes of Alberta: Part II on Residuals. *Alberta Foothills Disturbance Ecology Research Series Report No. 6*, 1-38.

Andison, D. W., 2013, Wildfires Patterns in Western Boreal Canada. *Healthy Landscape Research Series Report No. 8*, 1-65.

Araya, Y. H., Remmel, T. K. and Perera, A. H., 2015, Residual Vegetation Patches within Natural Boreal Wildfires: Characterizing by Pattern Metrics, Land cover Expectations, and Proximity to Firebreak features. *Geomatica,* 69(4): 327-338.

Araya, Y. H., Remmel, T. K. and Perera, A. H., 2016, What Governs the Presence of Residual Vegetation in Boreal Wildfires? *Journal of Geographical Systems,* 18(2): 159-181.

Austin, M., 2007, Species Distribution Models and Ecological Theory: A Critical Assessment and some Possible New Approaches. *Ecological Modeling*, 200: 1-19.

Beauvais, G. P., Keinath, D. A., Hernandez, P., Master, L. and Thurston, R., 2006, Element Distribution Modeling: a primer. *Nature Serve, Arlington Virginia*, 1-42.

Bergeron, Y., Drapeau, P., Gautheir, S. and Lecomte, N., 2007, Using Knowledge of Natural Disturbances to Support Sustainable Forest Management in the Northern Clay Belt. *The Forestry Chronicle*, 83(3): 326-337.

Breiman, L., 2001, Random Forests. *Journal of Machine Learning*, 45(1): 5-31.

Burton, P. J., Parisien, M., Hicke, J., Hall, R. and Freeburn, J., 2008, Large Fires as Agents of Ecological Diversity in the North American Boreal Forest. *International Journal of Wildfire*, 17: 754-767.

Cuesta, R. M., Garcia, M. and Retana, J., 2009, Factors Influencing the Formation of Unburned Forest Islands within the Perimeter of a Large Forest Fire. *Forest Ecology and Management*, 258: 71-80.

Cullinane-Anthony, B. L., Seefelt, N. E., Corace, R. G., Kashian, D. M. and Gehring, T. M., 2014, Influence of Residual Forest Patches on Post-Fire Bird Diversity Patterns in Jack Pine-dominated Ecosystems of Northern Lower Michigan. *Forest Ecology and Management*, 331: 93-103.

Edwards, T. C., Cutler, D. R., Beard, K. H. and Gibson, J., 2007, Predicting Invasive Plant Species Occurrences in National Parks: A Process for Prioritizing Prevention. USGS Utah cooperative Fish and Wildlife Research Unit, Utah State University, Logan, UT 84322-5290 USA, Report no. 2007-1.

Evans, J. E. and Cushman, S. A., 2009, Gradient Modeling of Conifer Species using Random Forests. *Landscape Ecology*, 24: 673-683.

Evans, J. S., Murphy, M. A. and Holden, Z. A., 2011, Modeling Species Distribution and Change using Random Forest. In: Drew, C.A. and Huettman, F., *Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications*, Springer, New York, 139-159.

Eunsik, P. and Wenbao, Y., 2011, Marker Selection by a Threshold Independent Measure Associated with Partial AUC. *Proceeding of the 58th World Statistical Congress*, 6163-6168.

Ferrier, S., Watson, G., Pearce, J. and Drielsma, M., 2002, Extended Statistical Approaches to Modelling Spatial Pattern in Biodiversity in North-East New South Wales. I. Species-level modeling. *Biodiversity and Conservation*, 11:2275-2307.

Fielding, A. H. and Bell, J. F., 1997, A Review of Methods for the Assessment of Prediction Errors in Conservation Presence/Absence Models. *Environmental Conservation*, 24(1): 38-49.

Hills, G.A., 1961, The Ecological Basis for Land-use Planning. *Ontario Department of Lands and Forests, Research Branch, Research Report No. 46*, 1-240.

Jepsen, J. U., 2004, Spatially Explicit Models in Landscape and Species Management. PhD Dissertation, National Environmental Research Institute, University of Copenhagen.

Johnson, E. A., 1995, Fire and Vegetation Dynamics: Studies from the North American Boreal Forest. Cambridge University Press, Cambridge.

Johnson, E. A., Miyanishi, K. and Weir, J. H., 1998, Wildfires in the Western Canadian Boreal Forest: Landscape Patterns and Ecosystem Management. *Journal of Vegetation Science*, 9: 603-610.

Magness, D. R., Huettmann, F. and Morton, J., 2009, Using Random Forest to Provide Predicted Species Distribution Maps as a Metric for Ecological Inventory and Monitoring Programs. *Studies in Computational Intelligence*, 122: 209-229.

Manel, S., Williams, H. C. and Ormerod, S. J., 2001, Evaluating Presence-Absence Models in Ecology: the Need to Account for Prevalence. *Journal of Applied Ecology*, 38: 921-931.

Munoz, J. and Felicisimo, A. M., 2004, Comparison of Statistical Methods Commonly used in Predictive Modeling. *Journal of Vegetation Science*, 15: 285-292.

OMNR, (Ontario Ministry of Natural Resources). 2014. Forest Management Guide for Boreal Landscapes. *Toronto: Queen's printer for Ontario, Technical Report*, 1-104.

Ouarmim, S., Ali, A. A., Asselin, H., Hely, C. and Bergeron, Y., 2015, Evaluating the Persistence of Post-fire Residual Patches in the Eastern Canadian Boreal Mixed Forest. *Boreas*, 44: 230-239.

Pearce, J. and Ferrier, S., 2000, Evaluating the Predictive Performance of Habitat Models Developed using Logistic Regression. *Ecological Modeling*, 133: 225-245.

Pearce, J., Ferrier, S. and Scotts, D., 2001, An Evaluation of the Predictive Performance of Distributional Models for Flora and Fauna in North-East New South Wales. *Journal of Environmental Management*, 62: 171-184.

Perera, A. H., Dalziel, B. D., Buse, L. J. and Routledge, R.G., 2009, Spatial Variability of Stand-Scale Residuals in Ontario's Boreal Forest Fires. *Canadian Journal of Forest Research*, 39: 945-961.

Perera, A. H. and Buse, L. J., 2014, Ecology of Wildfire Residuals in Boreal Forests. Wiley Blackwell, Oxford.

Peters, J., Baets, B. D., Verhoest, N. E., Samson, R., Degroeve, S., Becker, P. D. and Huybrechts, W., 2007, Random Forests as a Tool for Ecohydrological Distribution Modeling. *Ecological Modeling*, 207: 304-318.

R Core Team, 2015, R: a Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. Available at http://www.r-project.org.

Remmel, T. K. and Perera, A. H., 2009, Mapping Natural Phenomena: Boreal Forest Fires with Non-discrete Boundaries. *Cartographica*, 44(4): 274-288.

Riley, S., DeGloria, S. D. and Elliot, R., 1999, A Terrain Ruggedness Index that Quantifies Topographic Heterogeneity. *International Journal of Science*, 1(4): 23-27.

Rowe, J. S. and Scotter, G. W., 1973, Fire in the Boreal Forest. *Quaternary Research*, 3: 444-464.

Schwartz, M. W., Iverson, L. R., Prasad, A. M., Matthews, S. N. and O'Connor, R. J., 2006, Predicting Extinctions as a Result of Climate Change. *Ecology*, 87: 1611-1615.

Swets, J., 1988, Measuring the Accuracy of Diagnostic Systems. *Science*, 240:1285-1293.

Taylor, J. M., Ankerst, D. P. and Andridge, R. R., 2008, Validation of Biomarker-based Risk Prediction Models. *Clinical Cancer Research*, 14(19): 5977-5983.

Thompson, I. D., 2000, Forest Vegetation of Ontario: Factors Influencing Landscape Change. In: Perera, A.H., Euler, D.E. and Thompson, I.D (Eds), *Ecology of a Managed Terrestrial Landscape: Patterns and Processes of Forest Landscapes in Ontario*, University of British Columbia Press, Vancouver, 30-51.

Van Wagtendonk, J. W., 2004, Fire and Landscapes: Patterns and Processes. *USDA Forest Service General Technical Report. PSW-GTR-193*, 1-10.

Waljee, A. K., Higgins, P. D. and Singal, A. G., 2013, A Primer on Predictive Models. *Clinical and Translational Gastroenterolgy*, 5(1): e44.

Zweig, M. H. and Campbell, G., 1993, Receiver-Operating Characteristic (ROC) Plots: a Fundamental Evaluation Tool in Clinical Medicine. *Clinical Chemistry*, 39(4): 561-577.