

geospatial technologies

Mapping Percent Tree Mortality Due to Mountain Pine Beetle Damage

John A. Long and Rick L. Lawrence

The mountain pine beetle (*Dendroctonus ponderosae* Hopkins) is a pervasive and particularly destructive species of insect that has killed vast areas of conifers in western North America. To map large areas of infestation satellite imagery is often used because of its ability to cover large areas, but the spatial resolution often precludes fine-scale analysis. Furthermore, maps of insect-caused tree mortality have been attempted on the basis of presence-absence. We present here a two-step method to estimate and map tree mortality as a percentage within a Landsat-sized pixel. The first step delineates presence-absence, and the second estimates percent tree mortality for those pixels with tree mortality. We tested a total of 25 two-step models. The binary presence-absence first step of the models produced overall accuracies between 89.9 and 98.0% and Cohen's kappa values between 0.69 and 0.94. The full two-step models that predicted percent tree mortality produced estimated differences between modeled and observed percent tree mortality that ranged from <0.1 to 2.4% with associated root mean square errors between 10 and 18%. This method has several advantages over most current methodologies, including the use of a single image, nondependence on wetness measures, magnitude of error that is not density-dependent, and no restriction to trees with red-shaded faded crowns.

Keywords: forest disturbance, insect damage, tree mortality, Landsat imagery

Vast areas of North America's western conifer forests have experienced insect damage at epidemic levels, and mapping the location of insect-caused tree mortality is critical to monitoring forest health, informing management strategies, and understanding ecological relationships. The mountain pine beetle (MPB) (*Dendroctonus ponderosae* Hopkins) is a native component of the coniferous forests of western North America (Bentz et al. 2009). It is a pervasive species and can be found from South Dakota's Black Hills to the Pacific coast and from northwestern Mexico to the northern portions of British Columbia and Alberta. There is no shortage of suitable habitat for MPB, as they can successfully attack 22 species of the genus *Pinus*, including 6 species from the eastern United States and 4 species that are not native to North America; however, the principal host species are lodgepole pine (*Pinus contorta*), ponderosa pine (*Pinus ponderosa*), and whitebark pine (*Pinus albicaulis*) (e.g., Safranyik et al. 2010, Meddens et al. 2013).

Many populations of MPB are innocuous, causing little mortality to healthy trees, and periodic outbreaks of increased insect populations are a natural ecosystem process (Creeden et al. 2014); however, epidemic outbreaks can cause widespread tree mortality and affect millions of hectares of forest (Meddens et al. 2012). These

outbreaks produce substantial ecological and economic impacts, because they affect forest health and ecosystem services. Large regions of mortality in mature trees, for example, can reduce both the amount and quality of harvestable timber (Schwab et al. 2009), alter the dynamics of biogeochemical cycling (Kurz et al. 2008, Pfeifer et al. 2010), alter forest composition and structure as well as the distribution of habitat (Klenner and Arsenault 2009), and affect fuels and wildfire behavior (Jenkins et al. 2014).

The symptoms of an attacked tree follow a predictable trajectory of changes in foliage coloration (e.g., Skakun et al. 2003). These changes, known as stages, proceed over the course of 2 or 3 years. The foliage in newly attacked trees changes from a bright green to a duller green (green stage) during the first year due to changes in the cellular structure derived primarily from water loss. The second stage (red stage) is characterized by a substantial loss of chlorophyll and an accompanying change in foliage coloration to red. The final stage (gray stage) results once infested trees have lost the majority of their foliage (e.g., White et al. 2005, Wulder et al. 2006) and may last several years to a decade or longer. Details of the stages of infestation and the ecological relationship between the MPB, its blue-stain fungus symbiont (*Grossmannia clavigera*), and host tree

Manuscript received April 27, 2015; accepted February 25, 2016; published online April 14, 2016.

Affiliations: John A. Long (john.long@northern.edu), Department of Biology, Northern State University, Aberdeen, SD. Rick L. Lawrence (rickl@exchange.montana.edu), Montana State University.

Acknowledgments: This project was supported by the Agriculture and Food Research Initiative Competitive Grant 2013-68005-21298 from the USDA National Institute of Food and Agriculture, Grant/Cooperative Agreement 08HQGR0157 from the US Geological Survey (USGS) via a subaward from AmericaView, and the Montana Agricultural Experiment Station. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the USDA, USGS, or Montana Agricultural Experiment Station.

species have been extensively covered (e.g., see Tsui et al. 2012 or Lundquist and Reich 2014).

Much of the scientific work at the regional scale relies heavily on mapping efforts. The need to locate and map very large areas of infestation and tree mortality has motivated many studies based on multispectral imaging, because the spectral characteristics of foliage from healthy trees and from trees in various stages of MPB attack have been shown to be sufficiently different in several wavelength bands as to be useful for discrimination (Ahern 1988). The majority of these studies exploited differences in the spectral characteristics of healthy and insect-damaged trees (e.g., Wulder et al. 2006, Hicke and Logan 2009, Meddens et al. 2011, 2013), including differences in spectral wetness indices between the green- and red- or gray-stage trees (e.g., Coops et al. 2006, Goodwin et al. 2008, Cheng et al. 2010, Jewett et al. 2011). Many of these studies also relied on multitemporal imagery (e.g., Coops et al. 2006, Jewett et al. 2011, Meddens et al. 2013).

These studies have been generally successful in identifying presence-absence for each stage of attack, although spatial resolution of the sensor has been a limiting factor. Homogeneous regions are easily classified and mapped; however, success in heterogeneous regions with a mix of healthy trees and those in one of the stages of attack has been scale dependent (Bentz and Endreson 2003, Franklin et al. 2003). The pixel size of moderate-resolution satellite imagery, such as Landsat (900 m²), is problematic, because a single pixel can include several whole and partial trees, some attacked and some not. The ability to identify the percentage of tree mortality *within* Landsat-sized pixels is important for several reasons, including better understanding of the early stages of outbreak, when the percentage of tree mortality is low, and better informing of economic analyses when harvest is a viable option.

To our knowledge, there have been only two peer-reviewed published attempts to characterize mortality within a Landsat-sized pixel. Meddens et al. (2013) evaluated the ability of single-date and multitemporal Landsat imagery to estimate the percentage of a pixel that contained trees in the red stage; trees in the gray stage were excluded. They employed four-band aerial imagery (0.3 m resolution) that was ultimately aggregated to 30 m “superpixels” as the reference data. Success was a function of homogeneity in the reference data; pixels with large percentages of trees in the red stage had high classification accuracies, whereas cells with small percentages were not classified with high accuracy. Their method was an important first step, but it was unable to detect pixels with fewer than 25% trees in the red stage with acceptable accuracy (Meddens et al. 2013). In an additional study, Meddens and Hicke (2014) used multitemporal Landsat imagery to develop a continuous measure of red-stage tree mortality as part of a larger study of bark beetle spatial and temporal patterns.

Modeling tree mortality as a percentage is confounded by the fact that the data set is likely to contain many instances of no mortality (i.e., locations with no dead trees). These *zero-inflated* data have more observations equal to zero than expected under the distributional assumptions of standard statistical methods. This condition can lead to overdispersion and is common in field studies based on counts (Barry and Welsh 2002) and in those based on percentages (Vieira et al. 2010). Models that fail to account for zero-inflated data can give misleading results regarding the statistical significance of explanatory variables and, more importantly, can produce predictions that are good in aggregate, but poor for any given location (e.g., Potts and Elith 2006).

The purpose of the research reported in this article was to investigate the ability of several modeling approaches appropriate for zero-rich data to estimate and map the within-pixel percentage, rather than presence-absence, of tree mortality in a section of the Helena National Forest with moderate-resolution imagery. We used single-date Landsat-8 imagery and used four-band aerial imagery at 1-m spatial resolution as the reference data. The result was a map of percent tree mortality, rather than a binary dead/live characterization. The methods presented in this article are general and can be used for any analysis in which the goal is to determine the within-pixel percentage of any identifiable class.

Materials and Methods

Study Area

The study area comprises the south central section of the Helena National Forest in west central Montana (Figure 1). This site is in the Northern Rocky Mountains physiographic province and includes a large portion of the Elkhorn Mountains. The study area is a mountainous region of 57,260 ha (141,491 acres), which consists of approximately 75% coniferous forest and 25% montane steppe. Elevations range between 1,500 and 2,600 m. Local MPB populations began to increase relative to historical levels in 2005 and reached epidemic levels the following year, with substantial increases every year through 2009 (US Department of Agriculture [USDA] Forest Service 2010). Forests are primarily lodgepole pine, with MPB-caused tree mortality within Landsat pixel-sized areas that span the gradient from no mortality to complete mortality. Mortality within the study area overwhelmingly represented trees in the gray stage; we estimated that <2% of the dead trees were in the red stage. We define tree mortality in this study to include trees in the red stage and gray stage, as well as defoliated standing dead trees that may have died from other causes such as disease.

Data

We used both fine-resolution aerial imagery and moderate-resolution satellite imagery. The aerial images consisted of 29 color-infrared digital ortho quarter quad tiles from the National Agricultural Image Program (NAIP), which were mosaicked to provide complete coverage of the study area. Spectral changes across adjacent flight-lines is a potential issue with aerial imagery. This was not the case with the NAIP imagery because post-2006 NAIP images are radiometrically preprocessed to compensate for atmospheric absorption, solar illumination angle, and bidirectional reflectance (Montana State Library 2014). A visual inspection of all NAIP images was conducted with no radiometric issues that would affect visual interpretation of the imagery noted. These images were acquired during July 13–19, 2013, and were obtained from the Montana Geographic Information Clearinghouse.

The satellite imagery consisted of a single Landsat-8 scene (path 39, row 28) from July 18, 2013, which was obtained from the US Geological Survey's Earth Explorer portal. The date of the Landsat image was chosen to coincide as closely as possible to the dates of the NAIP imagery. The Landsat image had one small cumulus cloud and associated shadow, located along the southwestern edge of the study. The geometric accuracy design specification for the Landsat-8 Operational Land Imager (OLI) is 12 m, but the evaluation of test images suggests that Landsat-8 imagery is delivered at a geometric accuracy of ≤ 4.1 m (Storey et al. 2014). NAIP imagery has a geometric accuracy of < 6 m (USDA 2015). Consequently, there were no substantial issues regarding georegistration between

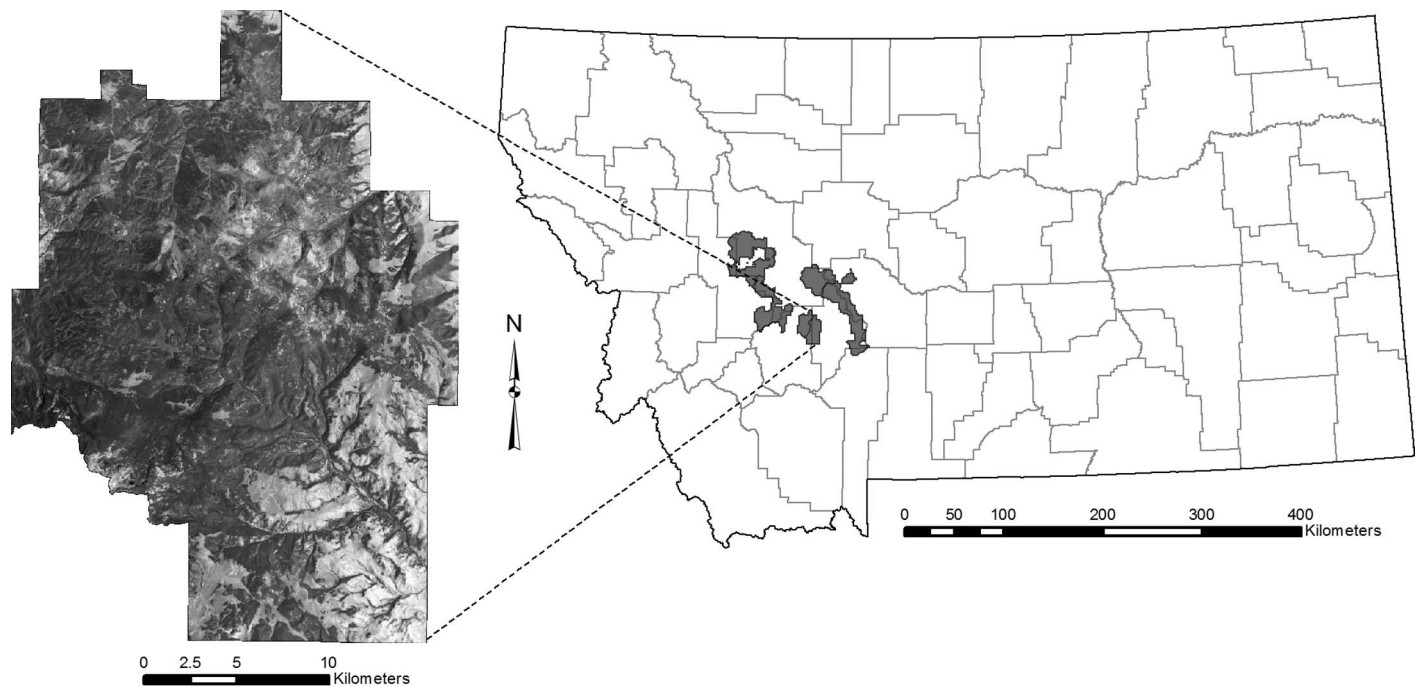


Figure 1. Aerial imagery and locator map of the study area located in the south central portion of the Helena National Forest (shaded area) in western Montana.

the Landsat-8 image and the NAIP images. This conclusion was supported by a close visual inspection of all images.

A total of 397 random locations were selected from within the study area; 296 of these were used to build the models, 99 were used for validation, and 2 were deleted because they were located in a cloud-covered area. The sample size for the training data was based on power calculations comparing two binomial proportions ($\alpha = 0.05$, power = 0.80) to detect a minimum difference of 10% (e.g., Rosner 2010). We extracted tree mortality (response variable), spectral reflectance data, topographic data, and four derived measures for each of these locations. The use of fine-scale imagery has been shown to be an acceptable substitute for ground reference data (e.g., White et al. 2005, Hicke and Logan 2009, Meddens et al. 2013); therefore, tree mortality, measured as the percentage of dead trees within a Landsat-8 pixel (900 m²), was estimated using the point-counting method (Belhouse 1981) with NAIP imagery (1 m²) as the reference. Point-counting used an acetate overlay printed with 100 points arranged in a 10 × 10 grid (Figure 2), which resulted in 3-m horizontal and vertical spacing between points, the approximate diameter of the average lodgepole pine tree crown and similar to the 2.4 m spatial resolution found to produce the highest classification accuracy by others (Meddens et al. 2011). This overlay was the same size as a Landsat pixel and was used on the NAIP imagery to identify the land cover (dead tree, live tree, live vegetation (nontree), bare, or shadow) under each of the 100 points at all 397 locations. A complete absence of tree mortality was observed in approximately 20% of the random locations. The distribution of class percentage was relatively consistent between the locations used to build the models and in the validation set (Figure 3). Shadows were common in the aerial images, particularly in forested areas. Consequently, the percentage of dead trees at each of the random locations was shadow-normalized (Dennison et al. 2010) by dividing by the sum of the nonshadow classes. The shadow-normalized percentage of dead trees per Landsat-sized pixel in the study area averaged 29.2%.

Spectral data (scaled radiance), the normalized difference vegetation index (NDVI), and the first 3 principal components (PC1, PC2, and PC3) were extracted from the Landsat-8 image for each of the 397 locations. Finally, we extracted the topographic variables, elevation, slope, and aspect, which were derived from the National Elevation Dataset 1 arc-second (~30 m) resolution product. Aspect was categorically coded (Flat, N, NE, E, SE, S, SW, W, and NW).

Modeling Approaches and Techniques

There are several approaches to modeling zero-rich data (e.g., see Vieira et al. 2010). We focused on two approaches. The first was a single-step approach using a nonparametric method that does not rely on distributional assumptions; the second was a two-step approach in which presence-absence of tree mortality was modeled initially, and the nonzero percent mortality was subsequently modeled in a second step (e.g., Welsh et al. 1996, Rideout et al. 1998). Using a two-step process effectively eliminates problems caused by zero-rich data because the zeros (absence of dead trees) are removed in the first step and therefore eliminated from further analysis.

We began with the single-step nonparametric approach by modeling percent mortality with random forest regression. Random forest (RF) is a tree-based ensemble classification and regression approach that creates multiple decision trees, each using a different bootstrapped random subset of the training data (Breiman 2001). In RF regression, the result is the mean of all individual tree predictions. RF is useful when the data contain numerous weak explanatory variables (Breiman 2001, Lawrence et al. 2006), and it is well established in the remote sensing literature (Long et al. 2013). We used two single-step RF models: RF1 and RF2. The first, RF1, used the training data as is; in the other, RF2, we reassigned all instances of predicted percent mortality that were <5% to zero. This reassignment was based on one-half of the minimum detectable difference based on our power test, i.e., the sample size precluded the ability to tell the difference reliably between 5% and 0%.

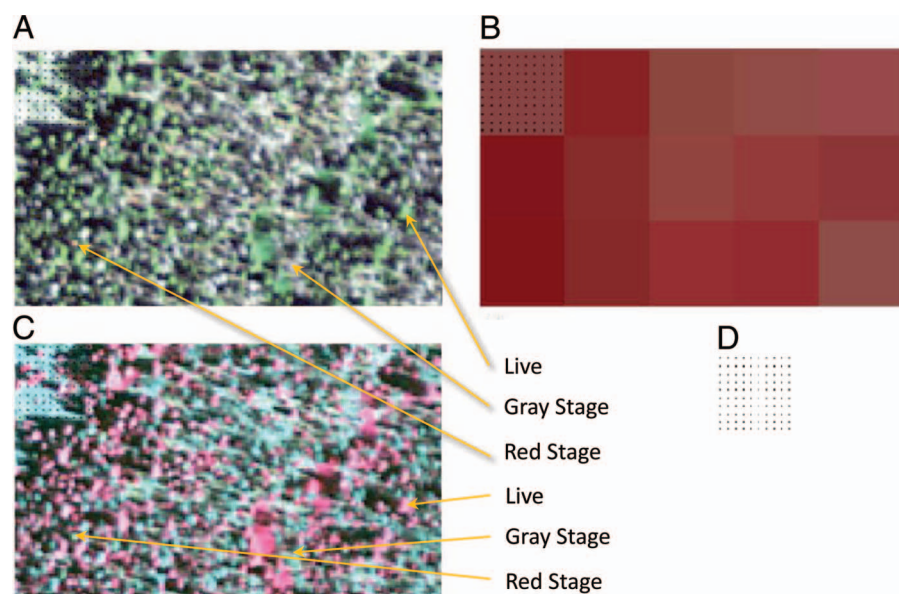


Figure 2. Close-up of a representative portion (150 m $\times$ 75 m) of the study area color infrared NAIP image displayed in true color (A), the same color infrared NAIP image displayed in standard false color (B), Landsat-8 image in standard false color (C), and the 10 $\times$ 10 grid overlay to scale. The point-counting grid has been superimposed on panels A, B, and C in the upper left-hand corners. Note that these colors are brighter and more easily distinguished on computer monitors (RGB color model) than they are in the CMYK color model required for printing. Live trees are green in the true color NAIP image and red in the standard false color NAIP image. Gray-stage trees are gray in the true color image and are greyish-blue to purple in the standard false color image. Red-stage trees are rare, and there are only two in this figure (in the lower left-hand quadrant); they are pink to red in the true color image and a pale greenish-yellow in the standard false color image. Shadows are dark in both images. Note that we are showing a representative scale; the actual classification would involve shifting between several scales before assigning a class.

We also used several combinations of techniques in various two-step analyses. The specific techniques included the following: (1) RF; (2) generalized linear models (GLMs); (3) boosted logistic regression (BLR); (4) beta regression (BR); (5) linear support vector machines (LSVMs); and (6) polynomial support vector machines (PSVMs). Two-step analyses began by modeling presence-absence with one of five techniques (RF, GLMs, BLR, LSVMs, or PSVMs). Then, all nonzero observations were used in the second step with one of five techniques (RF, GLMs, BR, LSVMs, or PSVMs). In total, we evaluated the ability of 27 models (Table 1) to estimate within-pixel tree mortality.

RF has already been discussed. GLMs use standard linear regression techniques and are based on maximum likelihood principles. These are well-known models and have been used in forestry-related remote sensing applications for decades. BLR is standard logistic regression that uses the boosting technique (Freund 1995, Freund and Schapire 1997) to improve prediction accuracy. Boosting, like RF, is an ensemble approach; it sequentially and iteratively reweights the training data and then takes the mean response from the multiple resulting models (e.g., Collins et al. 2002). We could find no studies that directly used the BLR technique in a remote sensing application. In the absence of any a priori assumptions, we separated 0s from 1s at a probability of 0.5 in all logistic regression models. BR (Ferrari and Cribari-Neto 2004) is analogous to multiple linear regression with a continuous response variable (y) that is restricted to the unit interval, $0 < y < 1$. Consequently, BR is well suited to responses measured as a percentage when they are converted to proportions; however, the dependent variable cannot be zero (i.e., there can be no instances of observed 0% mortality; hence, this method was only appropriate for the second step of the two-step approaches). Standard multiple linear regression is not appropriate

for percentages because the models can produce results outside of the unit interval, and it assumes a distribution that is theoretically unbounded. BR has successfully predicted percent forest canopy cover using field measurements (Korhonen et al. 2007) and has been used with satellite imagery (Coulston et al. 2012). Support vector machines (SVMs) are nonparametric techniques, which construct a set of linear hyperplanes (boundaries) in n -dimensional space that separate the training data such that the margins around the hyperplanes are maximized (Burges 1998). LSVMs use standard hyperplanes, whereas the nonlinear SVMs (such as PSVMs) gain linear separation by mapping the training data to a higher-dimensional feature space where the data are linearly separable (e.g., Mangasarian and Wild 2007). SVMs are also well established in remote sensing applications (for an excellent review, see Mountrakis et al. 2011).

Data Analysis

We used the statistical computing program R, including the packages *betareg*, *caret*, *caTools*, *e1071*, *MASS*, *randomForest*, *raster*, *rgdal*, *shapefiles*, and *sp*, for all data analysis efforts. The explanatory variables available for modeling were elevation (EL), slope (SL), the nine categorically coded aspects (AS), scaled radiance values for Landsat-8 bands (B1–B7 and B10–B11), NDVI, and the first three principal components (PC1–PC3). Accordingly, a total of 24 explanatory variables were considered. Each of the single-step models and the first step of the two-step models (presence-absence) used the full suite of variables. In the second step of the two-step models (percent mortality), RF and SVMS (LSVM and PSVM) used the full suite of variables, whereas the final versions of the BR and GLMs were selected by dropping statistically insignificant variables stepwise based on the Akaike information criterion. Models

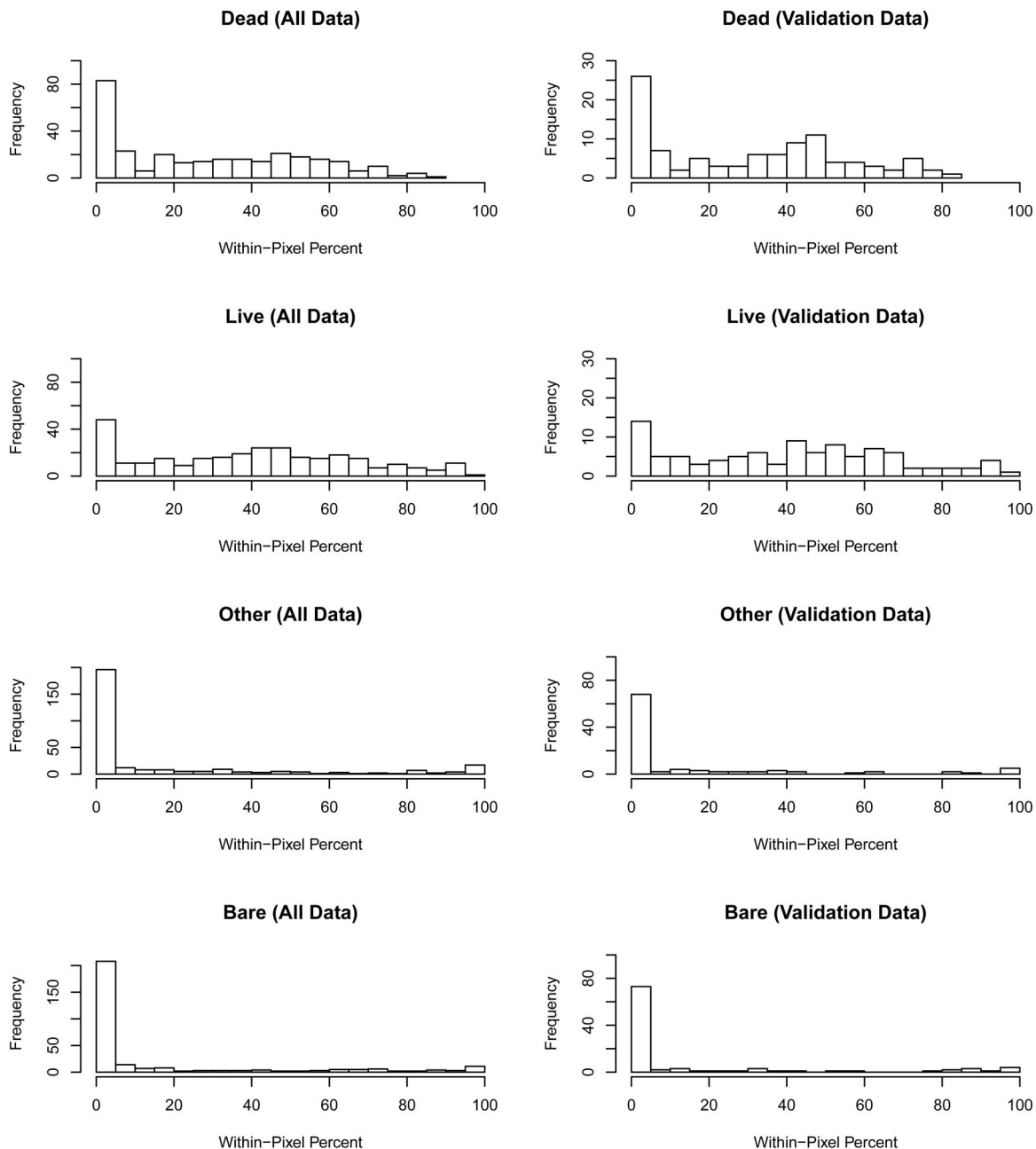


Figure 3. Distribution of within-pixel percentages by class for the training data and the validation data sets. These plots indicate that the classes do not strongly favor any particular within-pixel percentage and that the validation data are representative of the training data.

were constructed with the training data, and then tree mortality was estimated for the validation data using each model.

We assessed the classification performance of the presence-absence step of the two-step models based on overall accuracy and Cohen's kappa coefficients, κ (Cohen 1960) and the associated 95% confidence intervals (CIs), although performance was ultimately evaluated when combined with the second step of the models. The two-step models were evaluated by comparing the predicted percent tree mortality estimated from each model using the validation data with the observed percentage in the validation data by calculating the two-sample pseudomedian of differences between groups, along with the associated 95% CIs. The pseudomedian (Hodges-Lehmann estimator of location shift) of differences is a nonparametric measure of location functionally equivalent to the mean differ-

ence (Hodges and Lehmann 1963) and is the associated location measure for the Wilcoxon signed-rank test. We used the Wilcoxon signed-rank test (Wilcoxon 1945), a nonparametric alternative to the two-sample paired t -test, to test the hypothesis that the distributions of the predicted percentages and the observed percentages were equal (Demšar 2006). In addition, we computed the root mean square error (RMSE) as an additional measure that is robust to differences in error distributions (Chai and Draxler 2014). Finally, we computed error matrices for the presence-absence step of the two-step models.

Results

We begin with the single-step models RF1 and RF2, which gave differences in predicted and observed pseudomedians of 0.62 and

Table 1. Summary of models and explanatory variables included.

Presence-absence step	Percent mortality step	Explanatory variables
GLM	RF1	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3
	RF2	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3
	BR	SL + AS + B2 + B3 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2
	GLM	SL + AS + B1 + B3 + B5 + B6 + B9 + B10 + B11 + NDVI + PC2 + PC3
	RF	SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC3
RF	LSVM	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3
	PSVM	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3
	BR	SL + AS + B2 + B3 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2
	GLM	SL + AS + B1 + B3 + B5 + B6 + B9 + B10 + B11 + NDVI + PC2 + PC3
	RF	SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC3
BLR	LSVM	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3
	PSVM	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3
	BR	SL + AS + B2 + B3 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2
	GLM	SL + AS + B1 + B3 + B5 + B6 + B9 + B10 + B11 + NDVI + PC2 + PC3
	RF	SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC3
LSVM	LSVM	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3
	PSVM	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3
	BR	SL + AS + B2 + B3 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2
	GLM	SL + AS + B1 + B3 + B5 + B6 + B9 + B10 + B11 + NDVI + PC2 + PC3
	RF	SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC3
PSVM	LSVM	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3
	PSVM	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3
	BR	SL + AS + B2 + B3 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2
	GLM	SL + AS + B1 + B3 + B5 + B6 + B9 + B10 + B11 + NDVI + PC2 + PC3
	RF	SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC3
	LSVM	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3
	PSVM	EL + SL + AS + B1 + B2 + B3 + B4 + B5 + B6 + B7 + B10 + B11 + NDVI + PC1 + PC2 + PC3

RF1, single-step random forest; RF2, single step random forest with observations of <5% reassigned to 0%; GLM, generalized linear model; RF, random forest; BLR, boosted logistic regression; BR, beta regression; LSVM, linear support vector machine; PSVM, polynomial linear support vector machine; EL, elevation; SL, slope; AS, aspect; B1–B11, Landsat-8 bands 1–11 (B1, 0.43–0.45 μm , B2, 0.45–0.51 μm , B3, 0.53–0.59 μm , B4, 0.64–0.67 μm , B5, 0.85–0.88 μm , B6, 1.57–1.65 μm , B7, 2.11–2.29 μm , B10, 10.60–11.19 μm , B11, 11.50–12.51 μm); NDVI, normalized difference vegetation index: $(B5 - B4)/(B5 + B4)$; PC1–PC3, principal components 1–3.

Table 2. Performance of the one-step models.

Model ^a	<i>P</i> value ^b	Pseudomedian of differences ^c	CI ($\alpha = 0.05$) ^c	RMSE (%)
RF1	0.4884	0.62	–1.07 to 2.50	10.47
RF2	0.2968	1.21	–1.16 to 3.37	10.44

^a RF1, single-step random forest; RF2, single step random forest with observations of <5% reassigned to 0%.

^b Results are based on Wilcoxon signed-rank tests (note that *large P* values are “better” because they indicate failure to reject the null hypothesis that observed and predicted are the same).

^c Results are based on Wilcoxon signed-rank tests where the difference was formulated as observed – predicted and are given as percent.

1.21%, with associated 95% CI widths of 3.57 and 4.53%, respectively (Table 2). The RMSEs were 10.47% for RF1 and 10.44% for RF2. Both RF1 and RF2 had narrower CIs and smaller RMSEs than many of the two-step models; however, their pseudomedians of differences were intermediate in value, and their *P* values were among the smallest (note that *large P* values indicate no statistically significant difference between predicted and observed and suggest that the model predicts well). Reassigning all observed percent tree mortality of <5% to zero, as we did in RF2, did not provide any additional benefit and, arguably, produced a worse, although not statistically significantly worse, model. The single-step models also gave a positive pseudomedian of differences, suggesting that, on average, these models underestimated percent tree mortality. Concurrently, RF1 tended to predict low levels of presence when tree mortality was absent, which was less of an issue for RF2 as all predictions of <5% were reassigned to zero. The failure to predict absences did not have a substantial effect on any of the metrics but is

generally undesirable in mapping as it results in users of the map (e.g., land managers) being unable to distinguish early infestations (which might be a high priority for management) from lack of infestation (which might be the lowest priority for beetle management).

Overall classification accuracies for the presence-absence step of the two-step models ranged from 90.91% for the GLM to 97.98% for the RF model (Table 3). Kappa coefficients ranged from $\kappa = 0.6903$ (GLM) to $\kappa = 0.9396$ (RF). Cohen’s kappa coefficient accounts for the chance agreement between the observed accuracy and the expected accuracy; consequently, it can be a less misleading metric than overall accuracy for comparing methods. Nonetheless, arranging the models in descending order based on kappa coefficients, RF > BLR > PSVM > LSVM > GLM, produces the same ordering as arranging them based on overall accuracy. Overall, RF and BLR performed the best with respect to correctly predicting presence-absence, regardless of which measure is used to assess performance. More than three-quarters of the errors, regardless of model, resulted from predicting presence when the actual observed class was absence (Table 3; Figure 4). When tree mortality was absent, RF1 and RF2 predicted low levels of mortality, 3.66 and 2.27%, respectively (Table 4).

Each of the two-step models that we considered here performed well with respect to estimating percent tree mortality (Table 5); however, performance differed, depending on which metric was used for assessment. There was insufficient evidence statistically to suggest that model performance of any particular model was “better” than that of any of the other models considered; nonetheless, we make the following observations regarding differences in performance. The pseudomedian of differences, formulated as observed –

Table 3. Performance of the presence-absence step of the two-step models.

Classified data	Reference data		User's accuracy (%)
	Present	Absent	
GLM			
Present	77	9 (77/86)	89.53
Absent	1	12 (12/13)	92.31
	77/78	12/21	
Producer's accuracy (%)	98.72	57.14	
Overall accuracy	89.90%	$\kappa = 0.69$; CI ^a = 0.50 to 0.88	
RF			
Present	77	1 (77/78)	98.72
Absent	1	20 (20/21)	95.24
	77/78	20/21	
Producer's accuracy (%)	98.72	95.24	
Overall accuracy	97.98%	$\kappa = 0.94$; CI ^b = 0.96 to 1.00	
BLR			
Present	76	1 (76/77)	98.70
Absent	2	20 (20/22)	90.91
	76/78	20/21	
Producer's accuracy (%)	97.44	95.24	
Overall accuracy	96.97%	$\kappa = 0.91$; CI = 0.80 to 1.00	
LSVM			
Present	77	6 (77/83)	92.77
Absent	1	15 (15/16)	93.75
	77/78	15/21	
Producer's accuracy (%)	98.72	71.43	
Overall accuracy	92.93%	$\kappa = 0.78$; CI = 0.62 to 0.94	
PSVM			
Present	77	5 (77/82)	93.90
Absent	1	16 (16/17)	94.12
	77/78	16/21	
Producer's accuracy (%)	98.72	76.19	
Overall accuracy	93.94%	$\kappa = 0.84$; CI = 0.70 to 0.98	

^a CI for Cohen's kappa.

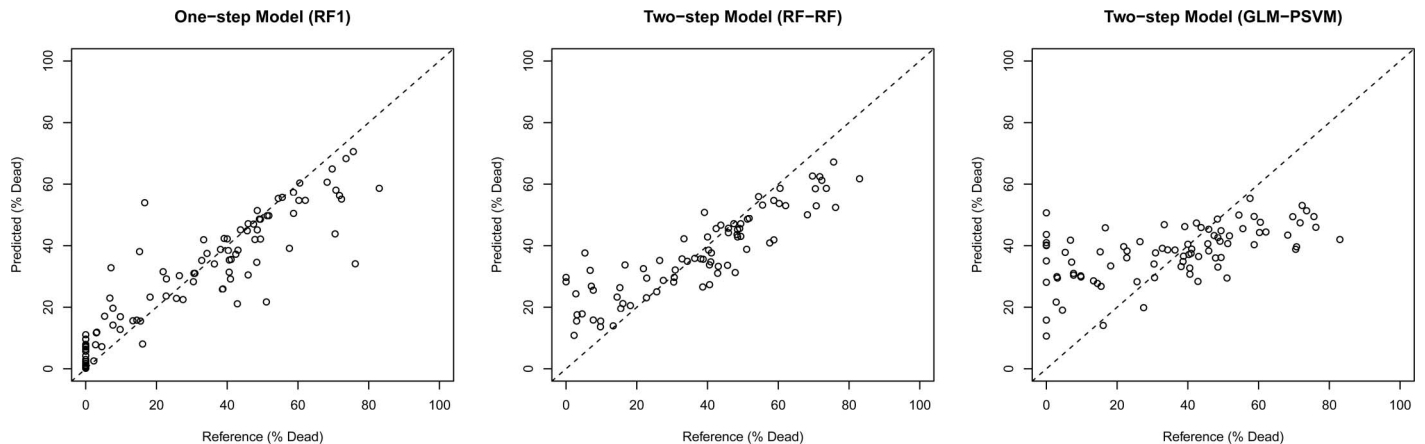


Figure 4. Scatterplot of reference versus predicted within-pixel percentage of tree mortality for the validation dataset using the RF1, RF-RF, and GLM-PSVM models as representatives. All models tend to overestimate low levels of mortality and underestimate high levels of mortality, but the amount is model dependent (Table 4). Note that the range of the errors is approximately equal regardless of the percent mortality.

predicted, ranged between -2.35 and 0.97% , whereas the 95% CI widths were between 3.57 and 8.95% . The smallest pseudomedian of differences (0.002%) resulted from using BLR for the presence-absence model and GLM for the continuous model; the largest (-2.35%) used LSVM for the presence-absence model and PSVM for the continuous model. Hereafter, for the two-step models, we will use a hyphenated nomenclature, e.g., BLR-GLM, to identify the presence-absence model and the continuous model, respectively. The P value for each model was substantially >0.05 (Table 5),

suggesting no statistically significant differences between observed and predicted percentages at the 95% confidence level; consequently, any of the models could be used to estimate percent tree mortality with that level of confidence. Based on pseudomedians of differences and P values, the BLR-GLM model performed best. However, based on RMSE, the RF-RF model was the superior model (RMSE = 10.42%). In general, models using RF, BLR, and BR produced the lowest RMSEs, whereas those using GLM, LSVM, and PSVM had higher RMSEs (Table 5).

Table 4. Overestimation of the absence of tree mortality.

Two-step model	Two-step model	Overprediction	Mean overprediction ^a
			 (%)
Presence-absence model	GLM		
	GLM.BR	10.54	12.45
	GLM.GLM	11.51	
	GLM.RF	13.17	
	GLM.LSVM	12.50	
RF	GLM.PSVM	14.54	
	RF.BR	1.78	2.24
	RF.GLM	1.84	
	RF.RF	2.76	
	RF.LSVM	2.14	
BLR	RF.PSVM	2.67	
	BLR.BR	2.14	2.42
	BLR.GLM	2.47	
	BLR.RF	2.64	
	BLR.LSVM	2.41	
LSVM	BLR.PSVM	2.42	
	LSVM.BR	9.36	10.88
	LSVM.GLM	9.87	
	LSVM.RF	12.42	
	LSVM.LSVM	10.63	
PSVM	LSVM.PSVM	12.13	
	PSVM.BR	4.97	6.29
	PSVM.GLM	5.23	
	PSVM.RF	7.48	
	PSVM.LSVM	6.15	
One-step model	PSVM.PSVM	7.63	
	RF1	3.66	NA
	RF2	2.27	NA

^a Mean overprediction for the five models is based on the presence-absence step. NA, not applicable.

The performance of the presence-absence step drove all subsequent measures. RF, for example, had the highest overall accuracy, the largest kappa, and the smallest pseudomedian of differences, the narrowest CI, the largest *P* value, and the smallest RMSE when

averaging across all two-step models that used RF as the presence-absence step. This ordered relationship held true for RF > BLR > PSVM. The “poorer” performing models, those using LSVM and GLM as the presence-absence model, were nearly indistinguishable in second-step average performance. We note that, within the second step, RF and BR tended to perform very well, regardless of the first-step model, whereas GLM and LSVM were less capable. The relationship between the reference percentage and the predicted percentage was such that the models tended to overestimate when predicting low levels of tree mortality (~20% or less) and underestimate when predicting high levels of tree mortality (~70% or greater). Success in predicting the absence of tree mortality varied by model as indicated by the error matrices (Table 3). When the two-step models incorrectly predicted the absence of tree mortality as presence, RF- and BLR-based models had the smallest average overpredictions, 2.24 and 2.42%, respectively. In contrast, GLM-based models overpredicted by an average of 12.45%, the LSVM-based model by an average of 10.88%, and the PSVM-based model by 6.29% (Table 4; Figure 4).

Discussion

Our purpose was to investigate the ability of several modeling approaches to estimate and map the within-pixel percentage of insect-caused tree mortality with moderate-resolution imagery. All of the models that we evaluated predicted percent tree mortality with what we believe likely to be acceptable accuracy from a management perspective, <3% in pseudomedians of differences and RMSE <18% in all cases, suggesting that it is possible to use moderate-resolution imagery, such as Landsat, in conjunction with fine-scale imagery and various modeling approaches to estimate percent tree mortality over large areas (Figure 5).

The single-step model RF1 produced very accurate predictions, <1% error in the pseudomedian of differences on average (Table 2),

Table 5. Summary of two-step model performance.

Presence-absence step	Percent mortality step	<i>P</i> value ^a	Pseudomedian of differences ^b	CI ($\alpha = 0.05$) ^b	RMSE (%)
GLM	BR	0.6644	−0.76	−4.56 to 2.55	15.47
	GLM	0.3551	−1.64	−5.55 to 1.68	15.79
	RF	0.5464	−0.84	−4.21 to 1.74	13.59
	LSVM	0.1832	−2.27	−5.78 to 1.21	16.18
	PSVM	0.3463	−2.27	−6.67 to 2.22	18.29
RF	BR	0.6061	0.85	−2.67 to 4.01	13.34
	GLM	0.9751	−0.04	−3.65 to 3.05	13.26
	RF	0.6061	0.56	−1.95 to 2.93	10.42
	LSVM	0.6129	−0.88	−4.16 to 2.66	13.75
	PSVM	0.9943	−0.04	−4.56 to 4.07	15.80
BLR	BR	0.6062	0.87	−2.84 to 4.14	14.00
	GLM	0.9981	0.00	−3.35 to 3.30	14.05
	RF	0.4500	0.97	−1.67 to 3.32	11.19
	LSVM	0.7137	−0.63	−4.10 to 3.01	14.50
	PSVM	0.8235	0.56	−4.17 to 4.59	16.24
LSVM	BR	0.6605	−0.76	−4.56 to 2.54	15.37
	GLM	0.3479	−1.67	−5.62 to 1.66	15.57
	RF	0.5466	−0.86	−4.15 to 1.77	13.53
	LSVM	0.1723	−2.31	−5.95 to 1.13	16.29
	PSVM	0.3283	−2.35	−6.77 to 2.18	17.99
PSVM	BR	0.9873	0.06	−3.65 to 3.28	13.93
	GLM	0.6255	−0.89	−4.69 to 2.36	14.00
	RF	0.9475	−0.07	−3.01 to 2.40	11.98
	LSVM	0.3324	−1.67	−5.09 to 1.88	14.79
	PSVM	0.5874	−1.31	−5.74 to 3.10	16.94

^a Results are Based on Wilcoxon signed-rank tests (note that *large P* values are “better” because they indicate failure to reject the null hypothesis that observed and predicted are the same).

^b Results are based on Wilcoxon signed-rank tests where the difference was formulated as observed − predicted and are given as percentages.

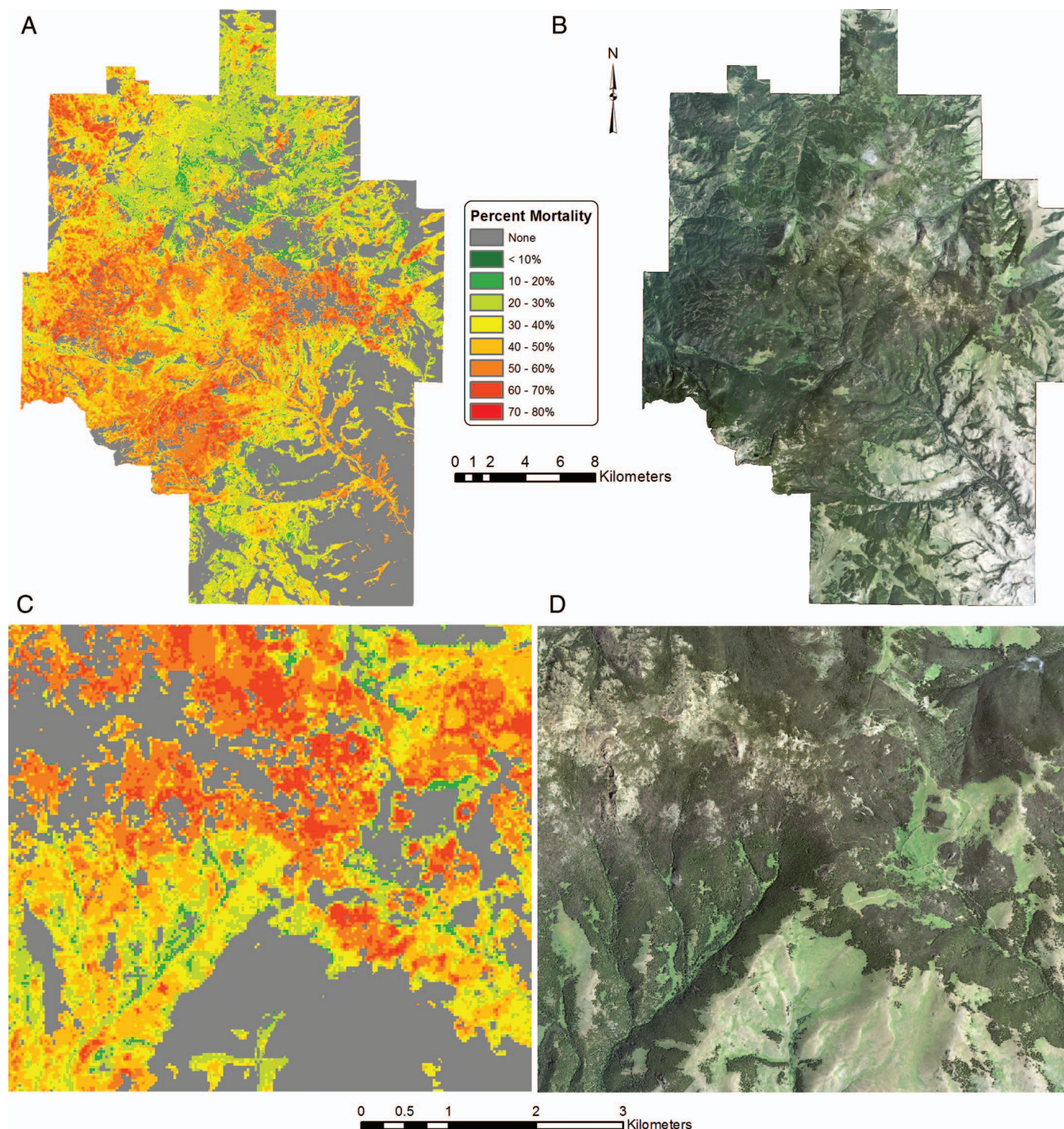


Figure 5. A. Map of the estimated percent mortality made by applying the RF-RF model across the study area. B. NAIP imagery for the study area. C. Close-up view of the estimated percent mortality map. D. NAIP imagery of the same close-up area. The full-region panels (A and B) indicate that the central and west central regions of the study area tend to have the highest levels of mortality, whereas the northern and northeastern portions are characterized by low-density mortality. The map matches the NAIP imagery well, including correctly identifying meadows as “no-mortality,” in the close-up panels (C and D).

but this is potentially misleading because the results were largely a consequence of overestimating the zero values (absence of tree mortality) while simultaneously underestimating the nonzero values (Figure 4). We attempted to account for this situation by reassigning all instances of predicted percent tree mortality of $< 5\%$ to zero in the RF2 model. This reassignment worsened performance with re-

spect to all measures, with the single exception of RMSE, which remained essentially unchanged. The RF2 model also suffered from a tendency to overestimate zero values and underestimate the nonzero values. Both single-step models produced excellent results in terms of overall model performance; nonetheless, they might be less desirable, depending on specific research interests, for mapping

because the models incorrectly identify pixels with no tree mortality as pixels with low levels of mortality, whereas the opposite was rarely the case. Predicting low densities of tree mortality in areas where no actual damage has occurred might obscure important spatial relationships, particularly in regions that are in the early stages of MPB outbreak or generate false conclusions that lead to inappropriate forest management decisions. We do not recommend either single-step model for mapping percent within-pixel tree mortality due to MPB attack.

The two-step models, collectively performed well, but varied depending on the metric used to assess performance. The inclusion of a presence-absence model as a first step substantially improved the ability of the models to correctly predict pixels with no tree mortality. This initial step in each of the two-step models accurately predicted presence-absence with >90% overall accuracy; however, the Cohen's kappa coefficients suggested that the ensemble-based methods, RF and BLR, were substantially better than the other methods, which were not ensemble-based. This finding is consistent with studies in the literature (e.g., Freund and Schapire 1997). Errors with models that used RF or BLR as the presence-absence model were less biased because the misclassifications were approximately equal between the classes, whereas the other methods showed a strong propensity to misclassify absence as presence rather than the other way around.

The focus of the second step of modeling was to estimate the nonzero responses, i.e., accurately predict percent tree mortality in pixels with actual damage. Overall model performance was ultimately driven by the performance of the first step, suggesting that partitioning the observations into presence-absence is a necessary initial step. Two-step models using RF or BR in the second step consistently produced better metrics than the other models, regardless of which technique was used in the first step. These findings suggest robustness across methods for RF and BR. Surprisingly, the more traditional GLM tended to outperform both the modern machine learning methods LSVM and PSVM in the second step, although the differences were not statistically significant.

All combined models were able to estimate percent tree mortality with acceptable accuracy for mapping MPB beetle damage. Nonetheless, we recommend considering either RF or BLR for the presence-absence model and either RF or BR for the second step of the model. The success of the models did vary according to which measure was used to assess performance. We recommend that researchers use the RF-RF model if RMSE is the desired performance measure. Alternatively, if the pseudomedian of differences is preferred, then the BLR-GLM model is recommended.

It is important in the evaluation of our results to clearly understand the context in which the models were evaluated. The objective in this study was to obtain the best predictive map, which in turn led us to evaluate models based on measures of accuracy and precision. Alternative objectives, which might include multiple objectives, could result in alternative measures for evaluation. If understanding the relationship between response and explanatory variables is a priority, for example, boosted and bagged statistical methods, such as RF and BLR, might be less desirable models because such information is limited or nonexistent, whereas GLM expressly shows such relationships. Similarly, if computational efficiency is important, for example, because of large data sizes and time constraints, an analysis would probably find significant differences among the methods.

The methods presented here have broad applicability across a

wide range of ecological applications (and elsewhere) where zero-inflated data are common. Species abundance data, in particular, tend to be zero rich (e.g., Welsh et al. 1996, Smith et al. 2012), and extending methodological approaches beyond traditional regression models to take advantage of modern, machine-learning approaches such as decision trees, random forest, or support vectors could benefit these applications. Other remote-sensing applications, as well as other studies, in which within-observation concentrations add to our understanding of processes of interest will also benefit from these methods, as we see the ability of sensors, when combined with appropriate methods, to accurately model within-observation variability.

Regardless of method, our work advances the science of mapping MPB damage in four notable respects: (1) only a single image is required rather than multitemporal imagery, which is often difficult to acquire as cloud free over large areas; (2) success is not based on wetness measures, which have low classification accuracy relative to our results (e.g., Skakun et al. 2003); (3) the magnitude of the errors and therefore accuracy is not density dependent—low densities of tree mortality can be estimated within approximately the same range of error as pixels with higher densities (although the direction of error, i.e., tendency to over- or underestimate, is density dependent); and (4) we are not restricted to mapping trees in the red stage and include trees in the gray stage in our definition of “mortality.”

Literature Cited

- AHERN, F.J. 1988. The effects of bark beetle stress on the foliar spectral reflectance of lodgepole pine. *Int. J. Remote Sens.* 9(9):1451–1468.
- BARRY, S.C., AND A.H. WELSH. 2002. Generalized additive modelling and zero inflated count data. *Ecol. Model.* 157:179–188.
- BELHOUSE, D.R. 1981. Area estimation by point-counting techniques. *Biometrics* 37:303–312.
- BENTZ, B., J. LOGAN, J. MACMAHON, C.D. ALLEN, M. AYRES, E. BERG, A. CARROLL, ET AL. 2009. *Bark beetle outbreaks in western North America: Causes and consequences. Bark beetle symposium, 2005 November, Snowbird, Utah.* Univ. of Utah Press, Salt Lake City, UT. 42 p.
- BENTZ, B.J., AND D. ENDRESON. 2003. *Evaluating satellite imagery for estimating mountain pine beetle-caused lodgepole pine mortality: Current status. Mountain pine beetle symposium: Challenges and solutions, 2003 October 30–31, Kelowna, British Columbia.* Inform. Rep. BC-X-399, Canadian Forest Service, Pacific Forestry Centre, Victoria, BC, Canada. 298 p.
- BREIMAN, L. 2001. Random forests. *Mach. Learn.* 45:5–32.
- BURGES, C.J.C. 1998. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Disc.* 2:121–167.
- CHAI, T., AND R.R. DRAXLER. 2014. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* 7:1247–1250.
- CHENG, T., B. RIVARD, G.A. SÁNCHEZ-AZOFEIFA, J. FENG, AND M. CALVO-POLANCO. 2010. Continuous wavelet analysis for the detection of green attack damage due to mountain pine beetle infestation. *Remote Sens. Environ.* 114:899–910.
- COHEN, J. 1960. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20(1):37–46.
- COLLINS, M., R.E. SCHAPIRE, AND Y. SINGER. 2002. Logistic regression, AdaBoost, and Bregman distances. *Mach. Learn.* 48:253–285.
- COOPS, N.C., M. JOHNSON, M.A. WULDER, AND J.C. WHITE. 2006. Assessment of QuickBird high spatial resolution imagery to detect red attack damage due to mountain pine beetle infestation. *Remote Sens. Environ.* 103:67–80.
- CREEDEN, E.P., J.A. HICKE, AND P.C. BUOTTE. 2014. Climate, weather, and recent mountain pine beetle outbreaks in the western United States.

- For. Ecol. Manage.* 312:239–251.
- DENNISON, P.E., A.R. BRUNELLE, AND V.A. CARTER. 2010. Assessing canopy mortality during a mountain pine beetle outbreak using GeoEye-1 high spatial resolution satellite data. *Remote Sens. Environ.* 114:2431–2435.
- FERRARI, S.L.P., AND F. CRIBARI-NETO. 2004. Beta regression for modeling rates and proportions. *J. Appl. Stat.* 31(7):799–815.
- FRANKLIN, S.E., M.A. WULDER, R.S. SKAKUN, AND A.L. CARROLL. 2003. Mountain pine beetle red-attack forest damage classification using stratified Landsat TM data in British Columbia, Canada. *Photogramm. Eng. Remote Sens.* 69(3):283–288.
- FREUNDE, Y. 1995. Boosting a weak learning algorithm by majority. *Inform. Comput.* 121:256–285.
- FREUNDE, Y., AND R.E. SCHAPIRE. 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* 55:119–139.
- GOODWIN, N.R., N.C. COOPS, M.A. WULDER, S. GILLANDERS, T.A. SCHROEDER, AND T. NELSON. 2008. Estimation of insect infestation dynamics using a temporal sequence of Landsat data. *Remote Sens. Environ.* 112:3680–3689.
- HICKE, J.A., AND J. LOGAN. 2009. Mapping whitebark pine mortality caused by a mountain pine beetle outbreak with high spatial resolution satellite imagery. *Int. J. Remote Sens.* 30(17):4427–4441.
- HODGES, J.L., AND E.L. LEHMANN. 1963. Estimates of location based on rank tests. *Ann. Math. Stat.* 34(2):598–611.
- JENKINS, M.J., J.B. RUNYON, C.J. FETTING, W.G. PAGE, AND B.J. BENTZ. 2014. Interactions among the mountain pine beetle, fires, and fuels. *For. Sci.* 60:489–501.
- JEWETT, J.T., R.L. LAWRENCE, L.A. MARSHALL, P.E. GESSLER, S.L. POWELL, AND S.L. SAVAGE. 2011. Spatiotemporal relationships between climate and whitebark pine mortality in the Greater Yellowstone Ecosystem. *For. Sci.* 57(4):320–335.
- KLENNER, W., AND A. ARSENAULT. 2009. Ponderosa pine mortality during a severe bark beetle (Coleoptera: Curculionidae, Scolytinae) outbreak in southern British Columbia and implications for wildlife habitat management. *For. Ecol. Manage.* 258(Suppl.):S5–S14.
- KORHONEN, L., K.T. KORHONEN, P. STENBERG, M. MALTAMO, AND M. RAUTIAINEN. 2007. Local models for forest canopy cover with beta regression. *Silva Fenn.* 41(4):671–685.
- KURZ, W.A., C.C. DYMOND, G. STINSON, G.J. RAMPLEY, E.T. NEILSON, A.L. CARROLL, T. EBATA, AND L. SAFRANYIK. 2008. Mountain pine beetle and forest carbon feedback to climate change. *Nature* 452:987–990.
- LAWRENCE, R.L., S.D. WOOD, AND R.L. SHELEY. 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). *Remote Sens. Environ.* 100:356–362.
- LONG, J.A., R.L. LAWRENCE, M.C. GREENWOOD, L. MARSHALL, AND P.R. MILLER. 2013. Object-oriented crop classification using multitemporal ETM+ SLC-off imagery and Random Forest. *GISci. Remote Sens.* 50(4):418–436.
- LUNDQUIST, J.E., AND R.M. REICH. 2014. Landscape dynamics of mountain pine beetles. *For. Sci.* 60(3):464–475.
- MANGASARIAN, O.L., AND E.W. WILD. 2007. Feature selection for non-linear kernel support vector machines. P. 231–236 in *Proc. of the seventh IEEE international conference on data mining: ICDM workshops 2007*, Tung, A.K.H., Q. Zhu, N. Ramakrishnan, O.R. Zaiane, Y. Shi, C.W. Clifton, and X. Wu (eds.). IEEE Computer Society, Los Alamitos, CA.
- MEDDENS, A.J.H., AND J.A. HICKE. 2014. Spatial and temporal patterns of Landsat-based detection of tree mortality caused by a mountain pine beetle outbreak in Colorado, USA. *For. Ecol. Manage.* 322:78–88.
- MEDDENS, A.J.H., J.A. HICKE, AND C.A. FERGUSON. 2012. Spatiotemporal patterns of observed bark beetle-caused tree mortality in British Columbia and the western United States. *Ecol. Applic.* 22(7):1976–1891.
- MEDDENS, A.J.H., J.A. HICKE, AND L.A. VIERLING. 2011. Evaluating the potential of multispectral imagery to map multiple stages of tree mortality. *Remote Sens. Environ.* 115:1632–1642.
- MEDDENS, A.J.H., J.A. HICKE, L.A. VIERLING, AND A.T. HUDAK. 2013. Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date Landsat imagery. *Remote Sens. Environ.* 312:49–58.
- MONTANA STATE LIBRARY. 2014. *Metadata for Montana 2013 color NAIP orthophotos*. Montana State Library, Helena, MT.
- PFEIFER, E.M., J.A. HICKE, AND A.J.H. MEDDENS. 2010. Observations and modeling of aboveground tree carbon stocks and fluxes following a bark beetle outbreak in the western United States. *Global Change Biol.* 17:339–350.
- POTTS, J.M., AND J. ELITH. 2006. Comparing species abundance models. *Ecol. Model.* 199:153–163.
- RIDEOUT, M., C.G.B. DEMÉTRIO, AND J. HINDE. 1998. Models for count data with many zeros. P. 179–190 in *Proc. of the 19th international biometrics Conference*, Cape Town, South Africa.
- ROSNER, B. 2010. *Fundamentals of Biostatistics*, 7th ed. Cengage Learning, Boston, MA. 888 p.
- SAFRANYIK, L., A.L. CARROLL, J. RÉGNIÈRE, D.W. LANGOR, W.G. RIEL, T.L. SHORE, B. PETER, B.J. COOKE, V.G. NEALIS, AND S.W. TAYLOR. 2010. Potential for range expansion of mountain pine beetle into the boreal forests of North America. *Can. Entomol.* 142:415–442.
- SCHWAB, O., T. MANESS, G. BULL, AND D. ROBERTS. 2009. Modeling the effect of changing market conditions on mountain pine beetle salvage harvesting in the British Columbia forest products industry. *Can. J. For. Res.* 39:1806–1820.
- SKAKUN, R.S., M.A. WULDER, AND S.E. FRANKLIN. 2003. Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage. *Remote Sens. Environ.* 86:433–443.
- SMITH, A.N.H., M.J. ANDERSON, AND R.B. MILLAR. 2012. Incorporating the intraspecific occupancy-abundance relationship into zero-inflated models. *Ecology* 93(12):2526–2532.
- STOREY, J., M. CHOATE, AND K. LEE. 2014. Landsat 8 Operational Land Imager on-orbit geometric calibration and performance. *Remote Sens.* 6:11127–11152.
- TSUI, C.K.M., A.D. ROE, Y.A. EL-KASSABY, A.V. RICE, S.M. ALAMOUTH, F.A.H. SPERLING, J.E.K. COOKE, J. BOHLMANN, AND R.C. HAMELIN. 2012. Population structure and migration pattern of a conifer pathogen, *Grosmannia clavigera*, as influenced by its symbiot, the mountain pine beetle. *Mol. Ecol.* 21:71–86.
- US DEPARTMENT OF AGRICULTURE. 2015. *National Agriculture Imagery Program (NAIP) information sheet*. USDA, Washington, DC. 2 p.
- USDA FOREST SERVICE. 2010. *State of the bugs—HNF summary 01/15/2010*. USDA For. Serv., Helena, MT. 1 p.
- VIEIRA, A.M.C., J.P. HINDE, AND C.G.B. DEMETRIO. 2000. Zero-inflated proportion data models applied to a biological control assay. *J. Appl. Stat.* 27(3):373–389.
- WELSH, A.H., R.B. CUNNINGHAM, C.F. DONNELLY, AND D.B. LINDEN-MAYER. 1996. Modelling the abundance of rare species: Statistical models for counts with extra zeros. *Ecol. Model.* 88:297–308.
- WHITE, J.C., M.A. WULDER, D. BROOKS, R. REICH, AND R.D. WHEATE. 2005. Detection of red attack stage mountain pine beetle infestation with high spatial resolution satellite imagery. *Remote Sens. Environ.* 96:340–351.
- WILCOXON, F. 1945. Individual comparisons by ranking methods. *Biometrics Bull.* 1(6):80–83.
- WULDER, M.A., J.C. WHITE, B. BENTZ, M.F. ALVAREZ, AND N.C. COOPS. 2006. Estimating the probability of mountain pine beetle red-attack damage. *Remote Sens. Environ.* 101:150–166.

Copyright of Forest Science is the property of Society of American Foresters and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.