12-9-2019

Detecting Tree Mortality with Landsat-Derived Spectral Indices: Improving Ecological Accuracy by Examining Uncertainty

Tucker J. Furniss
Utah State University

Van R. Kane
University of Washington

Andrew J. Larson
University of Montana

James A. Lutz
Utah State University

Recommended Citation

Utah State University
MERRILL-CAZIER LIBRARY

Detecting tree mortality with Landsat-derived spectral indices: Improving ecological accuracy by examining uncertainty

Tucker J. Furniss[1,†], Van R. Kane[2], Andrew J. Larson[3], James A. Lutz[1]

[1]*Wildland Resources Department and Ecology Center, Utah State University, Logan, UT 84322 USA*
[2]*School of Environmental and Forest Sciences, University of Washington, Seattle, WA 98195 USA*
[3]*Wilderness Institute and Department of Forest Management, University of Montana, Missoula, MT 59812 USA*

[†]Corresponding author e-mail: tucker.furniss@usu.edu

Abstract

Satellite-derived fire severity metrics are a foundational tool used to estimate fire effects at the landscape scale. Changes in surface characteristics permit reasonably accurate delineation between burned and unburned areas, but variability in severity is much more challenging to detect. Previous studies have relied primarily on categorical data to calibrate severity indices in terms of classification accuracy, but this approach does not readily translate into an expected amount of error in terms of actual tree mortality. We addressed this issue by examining a dataset of 40,370 geolocated trees that burned in the 2013 California Rim Fire using 36 Landsat-derived burn severity indices.

The differenced Normalized Burn Ratio (dNBR) performed reliably well, but the differenced SWIR:NIR ratio most accurately predicted percent basal area mortality and the differenced normalized vegetation index (dNDVI) most accurately predicted percent mortality of stems ≥10 cm diameter at breast height. Relativized versions of dNBR did not consistently improve accuracy; the relativized burn ratio (RBR) was generally equivalent to dNBR while RdNBR had consistently lower accuracy.

There was a high degree of variability in observed tree mortality, especially at intermediate spectral index values. This translated into a considerable amount of uncertainty at the landscape scale, with an expected range in estimated percent basal area mortality greater than 37% for half of the area burned (>50,000 ha). In other words, a 37% range in predicted mortality rate was insufficient to capture the observed mortality rate for half of the area burned. Uncertainty was even greater for percent stem mortality, with half of the area burned exceeding a 46% range in predicted mortality rate. The high degree of uncertainty in tree mortality that we observed challenges the confidence with which Landsat-derived spectral indices have been used to measure fire effects, and this has broad implications for research and management related to post-fire landscape complexity, distribution of seed sources, or persistence of fire refugia. We suggest ways to account for uncertainty that will facilitate a more nuanced and ecologically-accurate interpretation of fire effects.

This study makes three key contributions to the field of remote sensing of fire effects. First, we conducted the most comprehensive comparison to date of all previously published severity indices using the largest contiguous set of georeferenced tree mortality field data, revealing that the accuracy of both absolute and relative spectral indices depends on the tree mortality metric of interest. Second, we conducted this study in a single, large fire that enabled us to isolate variability due to intrinsic, within-landscape factors without the additional variance due to extrinsic factors associated with different biogeographies or climatic conditions. We show that uncertainty is related to fire severity and is greatest at intermediate severity levels. Finally, we identified the range in tree mortality that may be indistinguishable based on spectral indices derived from Landsat satellites, and we demonstrated how this variability translates into a considerable amount of uncertainty in fire effects at the landscape scale.

Introduction

The ecological effects of fire (hereafter "burn severity" *sensu* Morgan et al. 2014) are highly variable in both space and time, posing a challenge to managers and scientists who seek to quantify patterns in fire effects across large landscapes. For all but the smallest fires, satellite imagery is the only feasible way to accomplish this. Satellite-derived spectral indices can provide reasonably accurate estimates of burn severity over large landscapes (Miller et al. 2009), and they are relied upon heavily by land managers and fire scientists alike (Eidenshink et al. 2007). Since the formalized development of dNBR as a tool to detect burn severity in the early 2000's (Miller and Yool 2002, Key and Benson 2006) researchers have continued to develop and refine alternative indices (e.g., RdNBR, RBR, MIRBI), striving to improve accuracy and consistency among fires and across regions (Trigg and Flasse 2001, Roy et al. 2006, Miller and Thode 2007, Parks et al. 2014, McCarley et al. 2018). As numerous indices have been developed, there is a pressing need to compare the growing collection of spectral indices with a consistent dataset to determine which estimate tree mortality with the greatest accuracy, and to characterize the uncertainty that should be expected across the range of satellite-derived spectral index values.

A few studies have conducted such comparisons, but the emphasis has been on classification accuracy regarding categorical severity classes (burned vs. unburned areas [Brewer et al. 2005, Meddens et al. 2016]; or low, medium, and high severity [Epting et al. 2005, Miller and Thode 2007, Cansler and McKenzie 2012]). These approaches can determine overall classification accuracy, but they do not permit error to be quantified and assessed as a continuous variable. Recent studies have begun to address this research gap by calibrating spectral indices using continuous, field-based measures of burn severity (Miller et al. 2009, Whitman et al. 2018, Harvey et al. 2019), but these studies have utilized model performance metrics such as R^2, AUC, AIC, and Kappa – strong in a statistical sense but quite limited in their ability to describe uncertainty in tangible, ecological terms (e.g., range of fire effects that should be expected). To our knowledge no study to date has quantified *the range in tree mortality that should be expected at any given value of a spectral index* (but see Harvey et al. 2019 for a closely related example).

The 30-m resolution of Landsat pixels poses a challenge to ecological interpretation of reflectance values. There is a fundamental mis-match between the size of 30-m pixels and the crown spread of individual trees; pixel values therefore represent an "average" spectral response based on the wide range of spectral changes occurring among individual trees at the sub-pixel scale. The problem this creates is that the mortality of many small trees may elicit the same spectral response as the mortality of one large tree (Fig. 1), yet the ecological implications of these two scenarios are vastly different. Presently, one may refer to the literature to determine which spectral index has the highest classification accuracy, but until recently there was no way to translate spectral index values into an estimate of actual tree mortality (see Whitman et al. 2018 and Harvey et al. 2019 for two recent examples). A patch of Landsat pixels may appear to have burned with homogeneous severity based on satellite-derived indices, but the actual variability in fire effects (e.g., Fig. 1) remains uncharacterized. What is the range in tree mortality that may appear equivalent using satellite-derived data, and with what resolution can different fire severities reliably be delineated?

Understanding the relationship between Landsat-derived spectral indices and actual tree mortality will better enable both researchers and managers to accurately assess fire effects and their concomitant ramifications for changes to forest structure (Morgan et al. 2014). We evaluated 36 spectral indices using spatially-explicit, georeferenced forest inventory data of 40,370 trees that were alive pre-fire. The purpose of this study was not to identify a single

optimal index or to advance a new remote sensing technique. Rather, this study was designed to serve as a guide to enhanced ecological interpretation of Landsat-derived spectral indices. Our objectives were to: 1) quantify the variability in field-measured tree mortality (in terms of number of stems and basal area) that should be expected across different satellite-derived spectral index values; 2) determine which spectral indices have the highest accuracy for those field-based tree mortality metrics (i.e., basal area of mortality, percent change in density, and mortality of large-diameter trees); and 3) examine how variability in satellite-derived estimates of observed mortality may be scaled up to assess uncertainty at the landscape scale (>1,000 km^2).

Fig. 1. Images depicting the range in variability in actual fire effects compared to Landsat-derived spectral indices. Plots were selected to demonstrate how variability in biophysical conditions can elicit a wide range of spectral responses, and this is a source of considerable uncertainty in satellite-derived estimates of fire severity. The images represent two pairs of plots with approximately equivalent dNBR (320 – 322 for the top two images, 150 – 158 for the bottom two images). Images were taken in 2017, four years after the plots burned in the Rim Fire. The images are from four different sampling plots (of 313 total) and are intended to be representative of the overall fire effects. The plots include pixel 85 with the Yosemite Forest Dynamics Plot (top left) and three Joint Fire Science Program plots (G7-P3, top right; G5-P7, bottom left; G7-P4, bottom right). Plot names correspond to plot maps in Fig. S1.

Methods

Study sites

We conducted this study in the lower-montane mixed-coniferous forest zone of the Sierra Nevada Mountains, California, USA (Fig. 2). We combined two datasets: a long-term 25.6-ha permanent monitoring plot established in 2010 (the Yosemite Forest Dynamics Plot [YFDP]; Lutz et al. 2012, Lutz 2015), and a network of 53 0.25-ha plots established in 2017 (hereafter Joint Fire Science Program [JFSP] plots). All study sites were within Yosemite National Park (Yosemite) and were within the footprint of the 2013 California Rim Fire (excluding four unburned JFSP plots).

The YFDP has dimensions of 320 m × 800 m and contains 260 contiguous 30 m × 30 m Landsat pixels (Lutz et al. 2018b; Fig. S1). It is located at a mean elevation of 1,843 m in an old-growth *Abies concolor – Pinus lambertiana* (white fir – sugar pine) forest on predominantly northerly aspects. In 2009 – 2010, all live trees ≥1 cm and all snags ≥10 cm diameter at breast height (DBH; 1.37 m) were tagged, identified, and mapped (34,458 live trees; 2,697 snags). The plot has been censused every year since establishment; newly recruiting trees have been tagged and newly dead trees have been characterized by the factors contributing to their mortality. In August of 2013, the YFDP was burned for the first time since 1900 (Barth et al. 2015). A full post-fire remeasurement was conducted in May 2014.
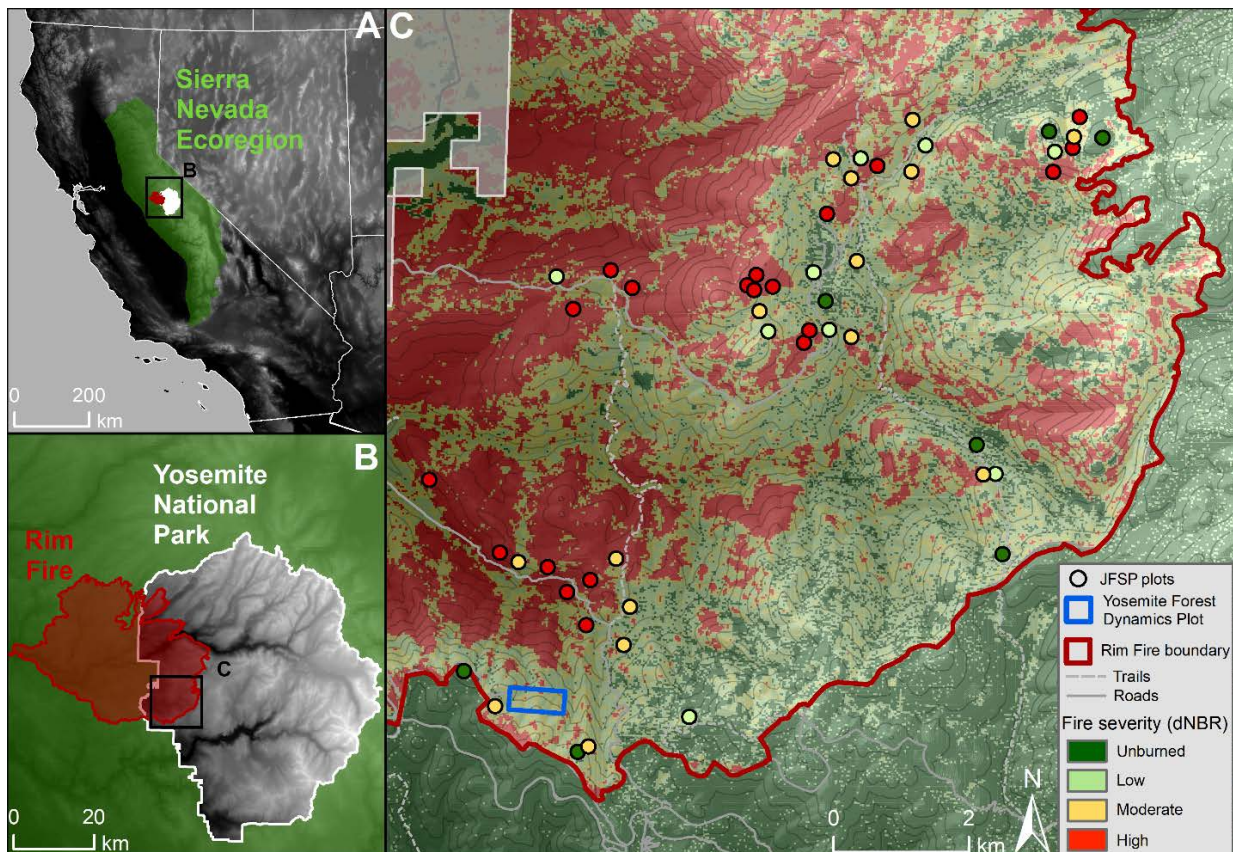

Fig. 2. Location of study sites on the western slope of the Sierra Nevada (A) within the lower-montane mixed-conifer zone of Yosemite National Park (B,C). Study sites were burned in the 2013 California Rim Fire and were selected to represent a broad range in fire severity (B). Burn severity classifications based on differenced Normalized Burn Ratio (dNBR) derived using Landsat 8 images (pre-fire scene July 14, 2013 and post-fire scene July 1, 2014). The study sites include 53 0.25-ha stem-mapped plots (colored according to dNBR severity classification) and the 25.6-ha Yosemite Forest Dynamics Plot (blue rectangle).

The JFSP plots were selected to capture the range of forest types, fire histories, and burn severities within the Yosemite Rim Fire footprint (Figs. 2, S1&S2). The JFSP plots were square 50 m × 50 m plots established in *Pinus ponderosa – Calocedrus decurrens*, *Pinus jeffreyi, Abies magnifica – Abies concolor*, and *Abies concolor – Pinus lambertiana* forest types between 1,431 m and 2,250 m elevation. Plots were installed post-fire based on a randomly chosen locations stratified by severity category (Fig. S2). Satellite-derived severity was estimated before plot installation using dNBR derived from Landsat 8 images (pre-fire scene July 14, 2013 and post-fire scene July 1, 2014). Plots were placed within patches of uniform severity class at least 3×3 pixels (90 m square; 0.81 ha) to buffer against positional error (Fig. S1). The distance to roads and trails was >0.1 km and <1.5 km to minimize edge effects while maintaining accessibility. Plots were located in the field using a handheld GPS, after which we collected precise positional data using a survey-grade GNSS receiver (Topcon HiPer SR). We post-processed plot coordinates to sub-meter accuracy with Natural Resources Canada Precise Point Positioning tool (https://webapp.geod.nrcan.gc.ca/geod/tools-outils/ppp.php). Within each JFSP plot, all trees ≥10 cm DBH were identified, measured, and mapped.

Trees were mapped from permanent grid markers using transect tapes and lasers and are within ±0.25 m to the datum (details in Lutz et al. 2012). Status was recorded as either live, fire-killed, or dead but not due to fire; no trees ≥10 cm DBH were fully consumed. Field personnel used a variety of cues to successfully determine status (see Jeronimo et al. in Press for details).

Rim fire

The Rim Fire began in August 2013 and was active for over two months, burning 104,131 ha of the Stanislaus National Forest and Yosemite National Park. After it finished burning it was the largest fire on record in the Sierra Nevada and the third largest in California, though it has since been surpassed by two similarly large fires in other parts of California (CAL FIRE 2018). The sites used in the present study burned with mixed severity (Figs. 2 & S2; Larson et al. 2016, Cansler et al. 2019, Kane et al. 2015, Furniss et al. 2019), with weather conditions more moderate compared to the low relative humidity and plume-dominated fire that resulted in large, high-severity patches in the Stanislaus NF and some parts of Yosemite (Lydersen et al. 2014, Lutz et al. 2017a, Collins et al. 2019).

Within Yosemite National Park, the Rim Fire burned with a higher proportion of moderate-severity (25% to 75% mortality by basal area) and high-severity (≥75% mortality by basal area) than was characteristic of the pre-suppression fire regime in lower-montane mixed conifer forests of the Sierra Nevada (based on Landsat-derived severity maps; Lydersen et al. 2014, Harris and Taylor 2015). However, since the reintroduction of fire to the park in the 1970s, there has been a shift to larger fires with greater areas of moderate- and high-severity (van Wagtendonk and Lutz 2007, Lutz et al. 2011, Collins et al. 2019), and the severity distribution for the Rim Fire within Yosemite was representative of this contemporary regime (Fig. S2). Fires in the Sierra Nevada are projected to either sustain or increase in size, frequency, and severity in the coming decades (Lutz et al. 2009), making the Rim Fire an optimal case study generalizable to other large, mixed-severity fires in the region.

Landsat data

Landsat 8 scenes were chosen to maximize scene clarity while matching pre- and post-fire phenology (*sensu* Key and Benson 2006). We conducted both initial (post-fire imagery taken immediately following fire) and extended assessments (post-fire imagery taken approximately one year post-fire). While extended assessments are generally considered to more accurately capture fire effects in forests (Key and Benson 2006, Miller and Thode 2007), initial assessment

may be important to inform management action within the year following the fire.

Extended assessments were conducted using data from July 14, 2013 (pre-fire) and July 1, 2014 (post-fire). Initial assessments were conducted using data from August 15, 2013 (pre-fire) and September 16, 2013 (post-fire). We were not able to select a more phenologically matched scene pair for the initial assessment because adequate Landsat imagery was not available the year prior to the fire (2012) – the Landsat 8 satellite was launched in 2013, the striping in Landsat 7 images covers the study area, and the Landsat 5 TM sensor failed in 2011. Though both scenes used in the initial assessment were from the same season, we conducted a phenological offset calibration to minimize the difference in phenology between the two scenes (details below).

Images were terrain corrected and co-registered by the USGS Earth Resources Observation and Science Center (EROS). These orthorectified top-of-atmosphere reflectance images (Level 1T products) were further processed by USGS EROS to remove atmospheric effects (Vermote et al. 2016), resulting in Level-2 surface reflectance (SR) image products that were accessed through the EROS Science Processing Architecture bulk ordering service (https://espa.cr.usgs.gov).

For the bi-temporal indices (i.e., indices based on differencing the pre- and post-fire scenes), we conducted a phenological calibration to minimize phenological mismatch between the two scenes. This was done by selecting a calibration zone adjacent to the fire (and in the same forest type) that did not burn, calculating the mean of the index values from the calibration area of the post-fire image, and subtracting this value from the entire scene (*sensu* Parks et al. 2014, Meddens et al. 2016). This reduces phenological bias and more clearly isolates fire-induced changes to vegetation.

Spectral index calculations

We compiled a list of 36 spectral indices that are sensitive changes in vegetation, with an emphasis on indices that have been used to detect fire effects (Table 1). We included both "snapshot" indices (based only on post-fire scene; 14 indices) and bi-temporal indices (difference between pre- and post-fire scenes; 22 indices). Indices were calculated according to published formulas included in Table 1. We also included slope, aspect, and landscape position as these topographic variables have been shown to influence fire behavior (Kane et al. 2015, Meddens et al. 2016). Topographic variables were calculated based on a 10-m digital elevation model (USGS 2007) that we resampled to match the resolution and extent of the Landsat images. Slope, aspect, and landscape position were calculated based on eight neighboring cells, and aspect was cosine-transformed.

Spectral index validation

We treated each of the 260 Landsat pixels within the YFDP as individual sampling units, and we generated separate tree lists based on the trees that fell within each pixel. We calculated all spectral indices for the entire Rim Fire, then we extracted the index values for each individual pixel. We also assigned an area-weighted average index value to each pixel based on adjacent pixels intersected by a 6-m buffer to account for the positional error of the Landsat scenes and canopy overlap of trees rooted in adjacent pixels. For the JFSP plots, we generated individual tree lists associated with each of the 53 plots. Pixel values for the JFSP plots were assigned using an area-weighted average because each of the 0.25-ha plots contained 4 to 6 partial pixels (Fig. S1). For each tree list, we calculated mortality rate by density and basal area (BA) as the percentage of pre-fire live trees $\geq$10 cm DBH (% stems ha^{-1} for density; % m^2 ha^{-1} for BA) that were killed within one year of the fire.

Table 1. Satellite-derived spectral indices and topographic variables. Subscripts 'pre' and 'post' indicate pre-fire and post-fire; L8 stands for Landsat 8; 'R' stands for spectral wavelengths within the red band, 'G' for the green band, 'B' for the blue band, 'NIR' for the near-infrared band, 'SWIR1' for the shortwave infrared band centered at 1.6 μm, and 'SWIR2' for the shortwave infrared band centered at 2.2 μm. Wavelength thresholds for each band may be found in the Landsat 8 handbook (USGS 2016). Coefficients for the tassled-cap transformation for L8 from in Baig et al. (2014).

	Index	Code	Formula (L8)	Citation
Snapshot	Normalized differenced vegetation index	NDVI	(NIR - Red) / (NIR + Red)	Rouse et al. 1974; Tucker 1979
	Mid-IR bispectral index	MIRBI	10*SWIR2 - 9.8*SWIR1 + 2	Trigg and Flasse 2001
	Normalized burn index	NBR	(NIR - SWIR2) / (NIR + SWIR2)*1000	Key and Benson 2006
	Char soil index	CSI	NIR / SWIR1	Smith et al. 2007
	Soil-adjusted vegetation index	SAVI	$((NIR - Red) * (1.0 + L)) / (NIR + Red + L)$	Huete 1988; Barbosa et al. 1999
	Normalized differenced moisture index	NDMI	(NIR - SWIR1) / (NIR + SWIR1)	Wilson & Sader 2002; Gao 1996
	SWIR1 to NIR ratio	SWIR1:NIR	SWIR1 / NIR	Vogelmann 1990
	SWIR2 to NIR ratio	SWIR2:NIR	SWIR2 / NIR	Kushla and Ripple 1998
	SWIR2 to SWIR1 ratio	SWIR2:SWIR1	SWIR2 / SWIR1	Epting et al. 2005
	NIR to G ratio	NIR:G	NIR / G	Landsat 8 handbook
	NIR to R ratio	NIR:R	NIR / R	Landsat 8 handbook
	Tassled-cap brightness	TC.BRI	$\sum$(coefficients * L8 bands 2 to 7)	Baig et al. 2014
	Tassled-cap greenness	TC.GRE	$\sum$(coefficients * L8 bands 2 to 7)	Baig et al. 2014
	Tassled-cap wetness	TC.WET	$\sum$(coefficients * L8 bands 2 to 7)	Baig et al. 2014
Bi-temporal	Differenced individual bands	dB	$B_{pre} - B_{post}$	McCarley et al. 2017
		dG	$G_{pre} - G_{post}$	McCarley et al. 2017
		dR	$R_{pre} - R_{post}$	McCarley et al. 2017
		dNIR	$NIR_{pre} - NIR_{post}$	McCarley et al. 2017
		dSWIR1	$SWIR1_{pre} - SWIR1_{post}$	McCarley et al. 2017
		dSWIR2	$SWIR2_{pre} - SWIR2_{post}$	McCarley et al. 2017
	Differenced indices	dNDVI	$NDVI_{pre} - NDVI_{post}$	Meddens et al. 2016, McCarley et al. 2017
		dMIRBI	$MIRBI_{pre} - MIRBI_{post}$	McCarley et al. 2017, 2018
		dNBR	$NBR_{pre} - NBR_{post}$	Key and Benson 2006
		RdNBR	$dNBR / (|NBR_{pre}| / 1000)^{0.5}$	Miller and Thode 2007
		RBR	$dNBR / ((|NBR_{pre}| / 1000) + 1.001)$	Parks et al. 2014
		dCSI	$CSI_{pre} - CSI_{post}$	Smith et al. 2007
		dSAVI	$SAVI_{pre} - SAVI_{post}$	McCarley et al. 2017
		dNDMI	$NDMI_{pre} - NDMI_{post}$	Meddens et al. 2016, McCarley et al. 2017
	Differenced band ratios	dSWIR1:NIR	$SWIR1.NIR_{pre} - SWIR1.NIR_{post}$	Meddens et al. 2016
		dSWIR2:NIR	$SWIR2:NIR_{pre} - SWIR2:NIR_{post}$	McCarley et al. 2017
		dSWIR2:SWIR1	$SWIR2:SWIR1_{pre} - SWIR2:SWIR1_{post}$	McCarley et al. 2017
		dNIR:G	$NIR:G_{pre} - NIR:G_{post}$	This study
		dNIR:R	$NIR:R_{pre} - NIR:R_{post}$	This study
	Differenced tassled-cap	dTC.BRI	$TC.BRI_{pre} - TC.BRI_{post}$	Meddens et al. 2016, McCarley et al. 2017
		dTC.GRE	$TC.GRE_{pre} - TC.GRE_{post}$	Meddens et al. 2016, McCarley et al. 2017
		dTC.WET	$TC.WET_{pre} - TC.WET_{post}$	Meddens et al. 2016, McCarley et al. 2017
Topo	Slope angle	Slope	Slope in degrees	Meddens et al. 2016
	Slope aspect	Aspect	cos(aspect)	Meddens et al. 2016
	Topographic position	TPI	elevation - $\sum$(elevation$_{neighbors}$)	Kane et al. 2015
	Solar irradiance	SRI	Relative index	Kane et al. 2015

Quantifying error

To quantify the accuracy of each spectral index, we created individual random forest models for each index with the index value as the predictor variable and observed mortality (% stems ha^{-1} for density and % m^2 ha^{-1} for BA) as the response. Correlations were summarized using percent variance explained (%VE), an independent out-of-bag estimate of error generated by the random forest algorithm. We used random forest analysis because it is a non-parametric statistical method that does not require the assumption of normality (Cutler et al. 2007), making it ideal to model fire severity data that may assume different non-linear response curves and contain non-normal residual distributions (Meddens et al. 2016, Whitman et al. 2018, Parks et al. 2018). We created separate models for each spectral index to avoid biased estimates of model performance due to collinearity between indices.

Because the ecological effects of tree death depend on the metric used (Lutz and Halpern 2006), we partitioned observed mortality in different ways to assess how each index detected the following observed mortality categories: percent basal area mortality (hereafter BA mortality), percent small tree mortality (1 cm ≤ DBH < 10 cm; this was only calculated for pixels within the YFDP), percent stem mortality for trees ≥10 cm DBH (hereafter stem mortality), percent stem mortality for trees ≥50 cm DBH (hereafter stem ≥50 cm DBH mortality), and percent stem mortality for trees ≥100 cm DBH (hereafter stem ≥100 cm DBH mortality). The subset of trees ≥10 cm DBH was used to permit direct comparison with other datasets that do not contain trees smaller than 10 cm DBH (as with many forest inventory datasets). We assessed small-diameter tree mortality separately for the distinct ecological role they play as a regenerating cohort. Our assessment of mortality of medium- and large-diameter (≥50 and ≥100 cm DBH, respectively) trees enabled us to test whether mortality of these less numerous but disproportionately important (Lutz et al. 2018a) trees may be accurately estimated with Landsat-derived indices.

In addition to the random forest analysis, we assessed accuracy using locally-weighted scatterplot smoothing (LOESS) models, a non-parametric statistical method which fits a regression line based on localized subsets of data points. For this analysis, we summarized observed mortality in two ways: percent BA and percent stem ≥10 cm DBH mortality. We avoided using generalized linear models because there is a weak theoretical basis for the shape of the relationship between spectral indices and tree mortality (e.g., van Wagtendonk et al. 2004, McCarley et al. 2017, Whitman et al. 2018), and this would have biased our results towards indices that happen to conform most closely to an arbitrarily-chosen model form. The LOESS technique performs regression based on the observed shape of the data rather than an *a priori* functional form, and may be used with non-linear data and non-normal distributions of residuals. We used the loess.sd() function from the msir package version 1.3.1 (Scrucca 2011) with the "span" parameter set to 0.8 and all other parameters set to their defaults.

We generated LOESS models for each spectral index with observed mortality as the independent variable and normalized index values as the dependent variable, and we extracted the standard deviation of the model as a continuous function of observed mortality values. We multiplied the standard deviation values by ±1.96 to create an error envelope that contained (approximately) 95% of the data points. This error envelope may be interpreted as the amount of *variability in the spectral response of each index as a function of observed burn severity*. This is intentionally different from a confidence interval based on standard error which would reflect the certainty with which one could estimate the "true" mean value. To characterize the variability in observed mortality across the range of spectral index values (as opposed to the variability in spectral indices across the range in observe mortality), we created additional LOESS models

with raw (non-normalized) index values as the independent variable and observed mortality as the dependent variable (same two variables, but axes were reversed; *sensu* Miller and Thode 2007, Miller et al. 2009, Parks et al. 2014, Whitman et al. 2018, Harvey et al. 2019). These models may be interpreted as the expected amount of *variability in observed tree mortality as a function of the values of each spectral index*.

We tested whether variability in satellite-derived spectral index values were related to pre-fire forest characteristics by extracting the residuals (predicted minus observed) from the LOESS models. Pre-fire structure for the JFSP plots was reconstructed by considering all fire-killed trees as live pre-fire and adding them to the populations of trees that were alive when the plots were established. We tested for correlations using linear regression between LOESS residuals and pre-fire stand attributes including pre-fire stand structure (density and basal area), LiDAR-derived topographic metrics (topographic position, slope, aspect, solar irradiance), and LiDAR-derived percentage canopy cover for four canopy height strata (>2 m, 2 m – 8 m, 8 m – 16 m, and 16 m – 32 m; details may be found in Kane et al. 2015). We examined spatial auto-correlation with the YFDP by creating semivariograms for basal area mortality and residuals of the LOESS model for dNBR.

Rarefied spectral index validation

Considering the uneven distribution of severities in the study sites (most pixels were low-severity ($\leq 25\%$ mortality by basal area) to moderate-severity; Fig. S1&S2), we conducted an additional validation based on a subset of the full dataset which was stratified by observed mortality class. The validation based on the full dataset will be most relevant for fires that burn with a similar severity distribution as the data used in this study (primarily low- to moderate-severity, as with many mixed-severity fires in the Sierra Nevada), the rarefied validation dataset will provide a more robust estimate of spectral index accuracy even at severity levels that are under-represented by our dataset (primarily high severity). This validation may therefore be more applicable to areas which burn with larger proportions of high-severity.

To create a rarified dataset, we classified the full dataset into 10% bins based on observed mortality, and we randomly selected two observations from each of bin to create a subset of 20 data points. At first we selected observations without replacement, then we selected with replacement once all observations in a given bin had been used. We continued to select unique subsets until each observation was used at least once, then we generated a random forest model for each spectral index based on each rarefied sample. Percent variance explained was averaged among all subsets to create a final rarefied accuracy estimate for each index. We performed this analysis on the five best performing indices as determined by the random forest models created with the full dataset. We calculated standard error and used Tukey's honest significant difference to identify significant differences between indices at $\alpha = 0.05$.

Estimating uncertainty at the landscape scale

We scaled the satellite-derived mortality predictions and associated variability to the entire Rim Fire footprint to explore how variability in a satellite-derived spectral index may translate to uncertainty in fire effects at the landscape scale ($>1,000$ km^2). We performed this analysis with dNBR, as it is perhaps the most ubiquitous severity index (Eidenshink et al. 2007, Meddens et al. 2016, McCarley et al. 2017). We used LOESS models that were created for dNBR to generate a predicted mortality rate (percent of BA and percent of stems ≥ 10 cm DBH) and associated standard deviation across the full range of dNBR values observed within the Rim Fire. We then assigned the predicted mortality rate and a 95% confidence envelop based on the dNBR value for every Landsat pixel within the fire footprint. We calculated high and low estimates of mortality

based on the predicted mortality rate ± 1.96 standard deviations, and generated an uncertainty map based on the difference between the high and low mortality estimates. The range of severities captured by our study sites represented 92% of the area burned by the Rim Fire (dNBR was lower in 5% and higher in 3% of the fire footprint), and we used conservative estimates of uncertainty for those extremities to avoid over-estimating uncertainty (we estimated 20% uncertainty for the lowest severities and 1% for the highest severities; both values were less than the observed level of uncertainty for at similar severity levels). The resulting map of uncertainty may be interpreted as the range in predicted mortality values that is necessary to bracket the true level of mortality 95% of the time. We conducted this analysis for both percent BA and percent stem mortality, and with both 95% and 68% confidence envelopes (±1.96 SD and ±1 SD, respectively).

All analyses were performed in R version 3.5.2 (R Core Team 2018) using version 1.3.1 of msir (Scrucca 2011), version 4.6.14 of randomForest (Liaw and Wiener 2002), and version 2.6.7 of raster (Hijmans 2016).

Results

Stem mortality ≥10 cm DBH one year post-fire ranged from 0% to 100% within the 53 JFSP plots and 0% to 88% within the 260 contiguous Landsat pixels within the YFDP. Basal area mortality ≥10 cm DBH was 0% to 100% within the JFSP plots and 0% to 74% within the YFDP (Fig. 3). Average mortality among the 313 plots was 43% of trees ≥10 cm DBH and 21% of BA. Crown torching was minimal within the YFDP, and most fire-killed trees retained red needles >1 yr post-fire. Satellite-derived severity ranged from unburned (no change detected) to high severity (minimum dNBR -11, maximum dNBR 870; Table S1), spanning a range of dNBR values representative of 92% of the area within the Rim fire footprint (5% of the fire footprint had dNBR values <-11 and 3% had dNBR values >870). Most of the study area was burned at low and moderate severities (Figs. 2, S1, S2; Table S1).

Detecting mortality with spectral indices

The extended assessment was more accurate than the initial assessment, especially for stem mortality (Table 2). The extended assessment increased percent variance explained (%VE) from 28% to 49% for percent stem mortality and from 58% to 63% for percent BA mortality. The most accurate indices for the initial assessment were dTC.WET (percent stem mortality only), dNBR (percent BA mortality only), and RBR (both stem and BA mortality). The most accurate indices for the extended assessment were dNDVI (percent stem mortality), dSWIR1.NIR (percent BA mortality), and dNBR (both stem and BA mortality). RBR and dNBR detected basal area mortality with similar accuracy (57.3%VE and 56.7%VE, respectively), but dNBR was more accurate for percent stem mortality (Table 2). RBR was more accurate for percent mortality of stems ≥50 cm DBH, and both indices had low accuracy for percent stem mortality of large-diameter stems (≥100 cm DBH). RdNBR was inferior to both dNBR and RBR in all mortality categories, with the exception of the marginal increase in accuracy for large-diameter stem mortality. The results for dSAVI were functionally equivalent to those of dNDVI for all analyses (their formulae are nearly identical); all discussion of dNDVI henceforth applies to dSAVI as well.

Accuracy was generally higher for percent BA mortality compared to percent stem mortality (stems ≥10 cm DBH), with an increase in %VE of 22% for RBR and 15% for dNBR. dNDVI, in contrast, was more accurate for percent stem mortality (Table 2). All spectral indices had very low accuracy in estimating percent mortality of saplings (1 cm ≤ DBH < 10 cm; maximum 8%

for NDVI), but there was a steep increase in %VE for small- and intermediate-sized stems (≥10 cm DBH; 49% for dNDVI). Percent stem mortality was estimated most accurately for stems ≥50 cm DBH (57% for dR), then declined for stems ≥100 cm DBH (33% for dSWIR2). Accuracy was higher for the JFSP plots compared to the YFDP, with maximum %VE for BA mortality of 83% and 38%, respectively (dR for JFSP and dNBR for YFDP; Table 2).

Table 2. Correlations between spectral indices and observed mortality. Values represent the percent variance explained determined with individual random forest models. Spectral index values and observed mortality were calculated for each Landsat pixel without a buffer (results based on buffered pixels may be found in Table S2). Observed mortality was quantified as percent of pre-fire live stems (or basal area [BA]) that was killed by fire. "Initial" columns represent correlations based on a post-fire Landsat scene from immediately after the fire (September 16, 2013), while "extended" columns use a post-fire scene taken one year following the fire (July 1, 2014). Column titles indicate which structural subsets were used to calculate observed mortality. The data were also subset by plot to compare accuracy between the two datasets. Bold indicates the best three spectral indices for each category of observed mortality. This table includes the best five indices for each category of observed mortality; see Table S2 for results for all indices.

	Percent variance explained (%)								
	Initial		**Extended**					**% BA**	
Index	**Stems ≥10 cm**	**Basal area**	**Basal area**	**Stems 1-10 cm**	**Stems ≥10 cm**	**Stems ≥50 cm**	**Stems ≥100 cm**	**YFDP**	**JFSP**
dNBR	**26.5**	**57.6**	**56.7**	-	**41.5**	42.6	3.8	**38.3**	**76.8**
RdNBR	18.7	40.7	49.9	-	30.9	40.3	15.2	**36.9**	44.3
RBR	**26.9**	**56.3**	**57.3**	-	34.9	**51.8**	10.3	28.3	68.0
dSWIR1.NIR	14.8	49.5	**63.1**	-	**42.8**	51.5	9.1	**30.1**	73.2
dNDVI	25.4	**54.8**	42.8	-	**49.4**	31.9	6.4	20.9	76.5
dSWIR2.NIR	23.8	51.6	54.9	-	33.1	46.0	**24.6**	28.7	59.9
dTC.WET	**28.1**	48.3	52.1	-	20.9	48.6	5.6	16.0	73.6
dR	2.3	45.6	54.8	-	40.3	**57.4**	**20.7**	4.5	**83.3**
dB	0.0	45.1	49.3	-	29.3	**56.1**	19.1	0.0	67.7
dSWIR2	3.8	47.3	56.4	-	30.1	48.1	**33.0**	24.6	56.4
dTC.GRE	12.0	33.5	41.7	-	22.8	23.3	14.8	26.0	**79.2**
dNDMI	24.0	46.6	37.0	-	41.0	23.8	8.9	25.9	72.8
NDVI	3.7	33.2	29.1	**8.0**	33.1	34.9	10.8	28.1	28.6

Spectral index accuracy at different levels of observed mortality

Visual interpretation of the LOESS regression error envelopes corroborated the random forest results (Fig. 3&S3&S4). The indices with the highest %VE according to random forest models had the tightest error envelopes, and envelopes grew as %VE decreased. The range in observed mortality (both percent stem and BA morality) was greatest at intermediate spectral index values, but this contained pixels that were burned at both moderate- and high-severity according to the categorical severity classes for dNBR established by Miller and Thode (2007). The range in percent BA mortality was 18% for unburned (dNBR < 41), 31% for low severity (41 ≤ dNBR < 176), 47% for moderate severity (177 ≤ dNBR < 366), and 61% for high severity (dNBR ≥367). The range in percent stem mortality was 34% for unburned, 61% for low severity, 65% for moderate severity, and 37% for high severity (Fig. 3).

The correspondence between observed mortality and spectral indices assumed a different relationship when the axes were swapped (Fig. 4). The relative accuracy of each index remained

the same, but the distribution of uncertainty was different. The previous analysis demonstrated that uncertainty in observed mortality was greatest at intermediate spectral index values, while this analysis demonstrated that the range in spectral index values was greatest at the highest levels of observed mortality. In other words, a strong spectral response was associated with a very low range in observed mortality (because mortality was 100%; Fig. 3), but a high level of observed mortality (>95%) was associated with a wide range of spectral responses (Fig. 4). For example, the range in dNBR values for plots that had 100% mortality was 452 to 870 (Figs. 3&4); a near doubling of spectral response values without any difference in tree mortality. The range in normalized index values was positively related to both percent stem and percent BA mortality; variability was approximately two standard deviations at low (<25%) mortality levels, while the range often exceeded 3-4 standard deviations at high (>75%) mortality levels (Fig. 4).

Variability in observed mortality (residuals of the LOESS models) was not correlated with any of the pre-fire metrics examined: pre-fire basal area, pre-fire density, LiDAR-derived topographic variables, or LiDAR-derived canopy metrics. Accuracy for the YFDP pixels was generally reduced by using an area-weighted average pixel value based on a 6 m buffer around each pixel, though some snapshot indices (derived from the post-fire scene alone) were marginally improved with the buffer (Table S2).

Rarefied results

Random forest models based on a rarefied dataset yielded subtly different results compared to the models based on the full dataset (Table S3). The rarefication procedure enabled us to calculate the standard error of %VE for the random forest models based on different subsets of data and to test for significant differences. As with the full dataset models, dSWIR1.NIR had the highest accuracy for percent BA mortality, while dNDVI had the highest accuracy for percent stem mortality (Table S3). RdNBR, RBR, and dSWIR1.NIR were most accurate for stem mortality of medium-diameter stems, while RdNBR and RBR were most accurate for large-diameter stems. dNBR was not best in any category, but it was never the least accurate.

Uncertainty at the landscape scale

Scaling the relationship between dNBR and observed mortality to the entire Rim fire, estimated mortality within the year following the fire was 45% of basal area and 60% of stems ≥10 cm DBH (Fig. 5&S5). At a 95% confidence level, uncertainty in percent BA mortality was less than 10% for only 5% of the fire footprint, <20% for 8% of the footprint, and <40% for 53% of the footprint (Fig. 5). Uncertainty was generally higher for percent stem mortality, with less than 10% uncertainty for 11% of the fire footprint, <20% for 20% of the footprint, and <40% for 43% of the footprint (Fig. S5). Half of the Rim Fire footprint (median uncertainty value) had a range in estimated mortality of >37% of pre-fire basal area and >46% of pre-fire stems ≥10 cm DBH. Mortality estimates were most accurate in the high-severity and unburned patches (based on dNBR) where uncertainty was within 10%, while uncertainty in moderate-severity areas was as high as 80% for percent BA and 70% for percent stem mortality (Fig. 5&S5).

We observed similar results at the 68% confidence level (Fig. S6&S7). Uncertainty was greatest overall for percent stem mortality, but the greatest uncertainty for an individual dNBR severity class was for percent BA mortality in moderate-severity areas where the range in estimated percent BA mortality was 40% (Fig. S6). In other words, a 40% range in predicted BA mortality rates (e.g., BA mortality predicted to be between 50% and 90%) would only contain the true level of mortality approximately two-thirds of the time. The range in the 68% confidence envelope exceeded 22% for percent BA mortality and 29% for percent stem mortality in half of the Rim Fire footprint (Fig. S6&S7).
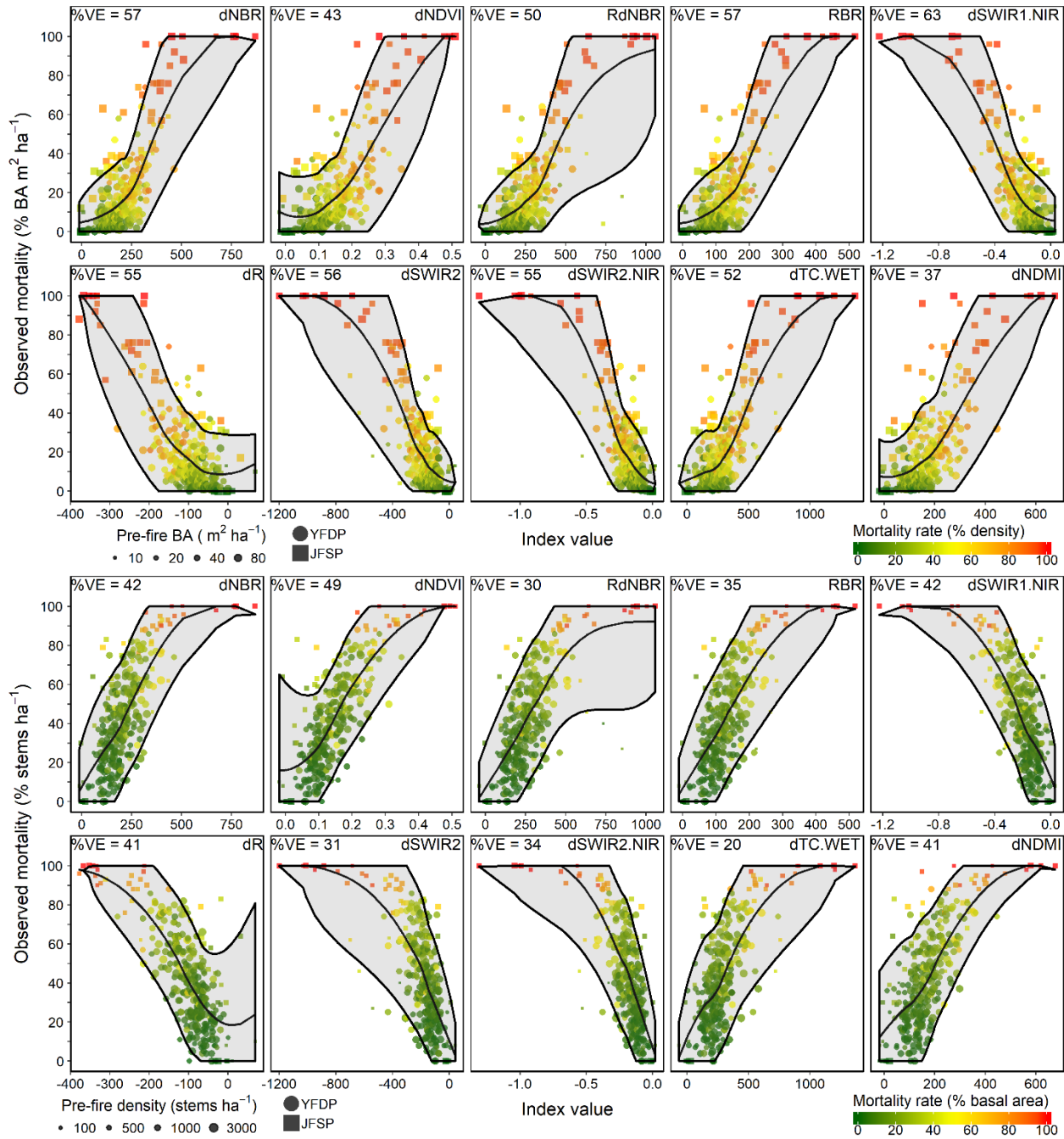
Fig. 3. Relationship between satellite-derived spectral indices and observed mortality of trees ≥10 cm diameter at breast height (1.37 m). Points represent 53 individual plots (JFSP) and 260 Landsat pixels with in the Yosemite Forest Dynamics Plot (YFDP). Shaded envelopes represent the variability in observed mortality as a function of each spectral index. Envelopes were derived from continuous estimates of the mean and standard deviation of the data generated with locally-weighted scatterplot smoothing (LOESS) regression models and scaled to capture 95% of the variability in observations. The %VE indicates the percent variance explained using random forest models (Table 2, S2). See Figs. S3&S4 for results for all 36 indices.
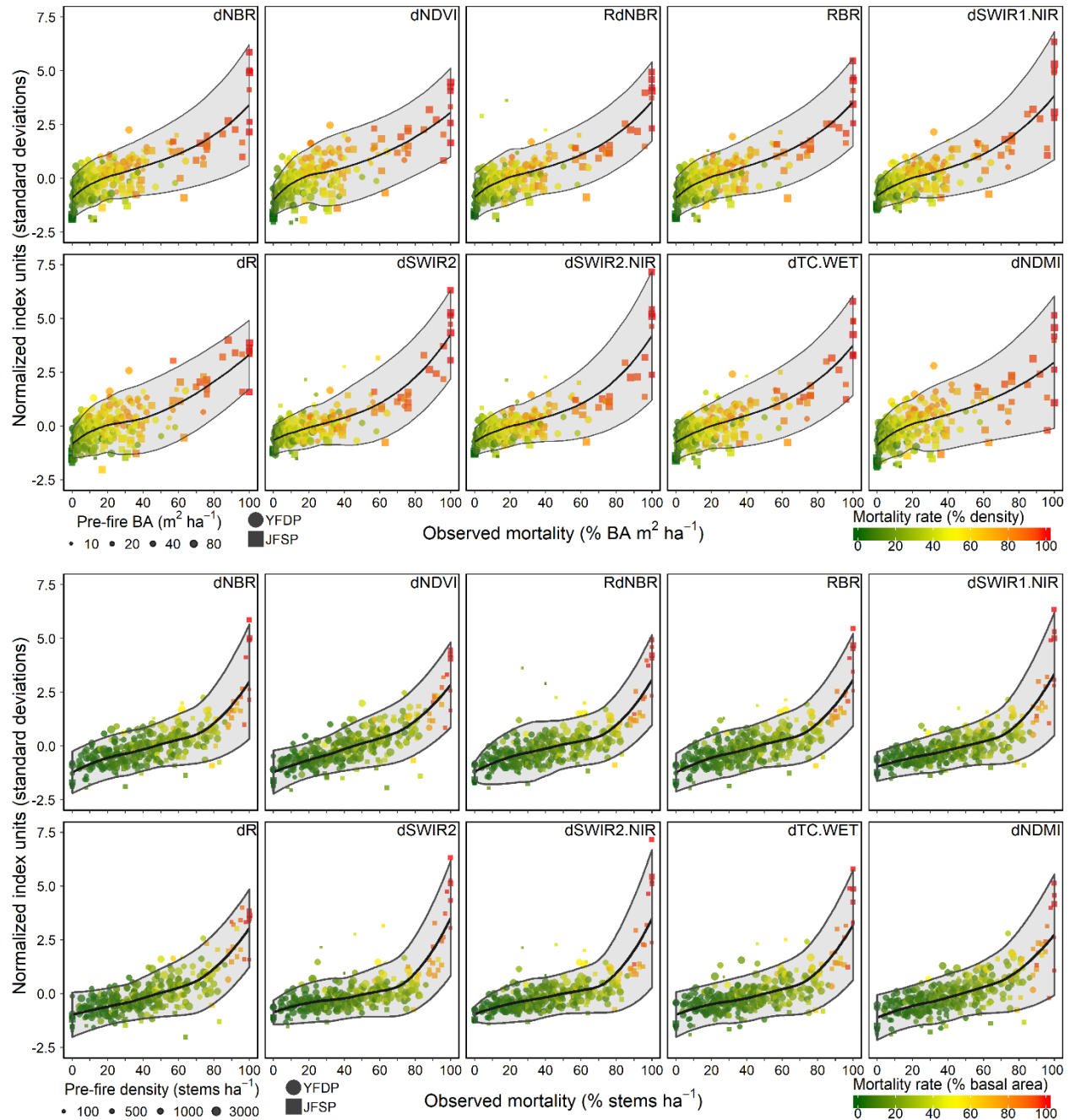
Fig. 4. Variance in satellite-derived spectral indices as a function of observed mortality of trees ≥10 cm diameter at breast height (1.37 m). Points represent 53 individual plots (JFSP) and 260 Landsat pixels with in the Yosemite Forest Dynamics Plot (YFDP). Shaded envelopes represent the variability in observed mortality as a function of each spectral index. Envelopes were derived from continuous estimates of the mean and standard deviation of the data generated with locally-weighted scatterplot smoothing (LOESS) regression models and scaled to capture 95% of the variability in observations. Spectral index units were normalized to enable comparison between different spectral index values.
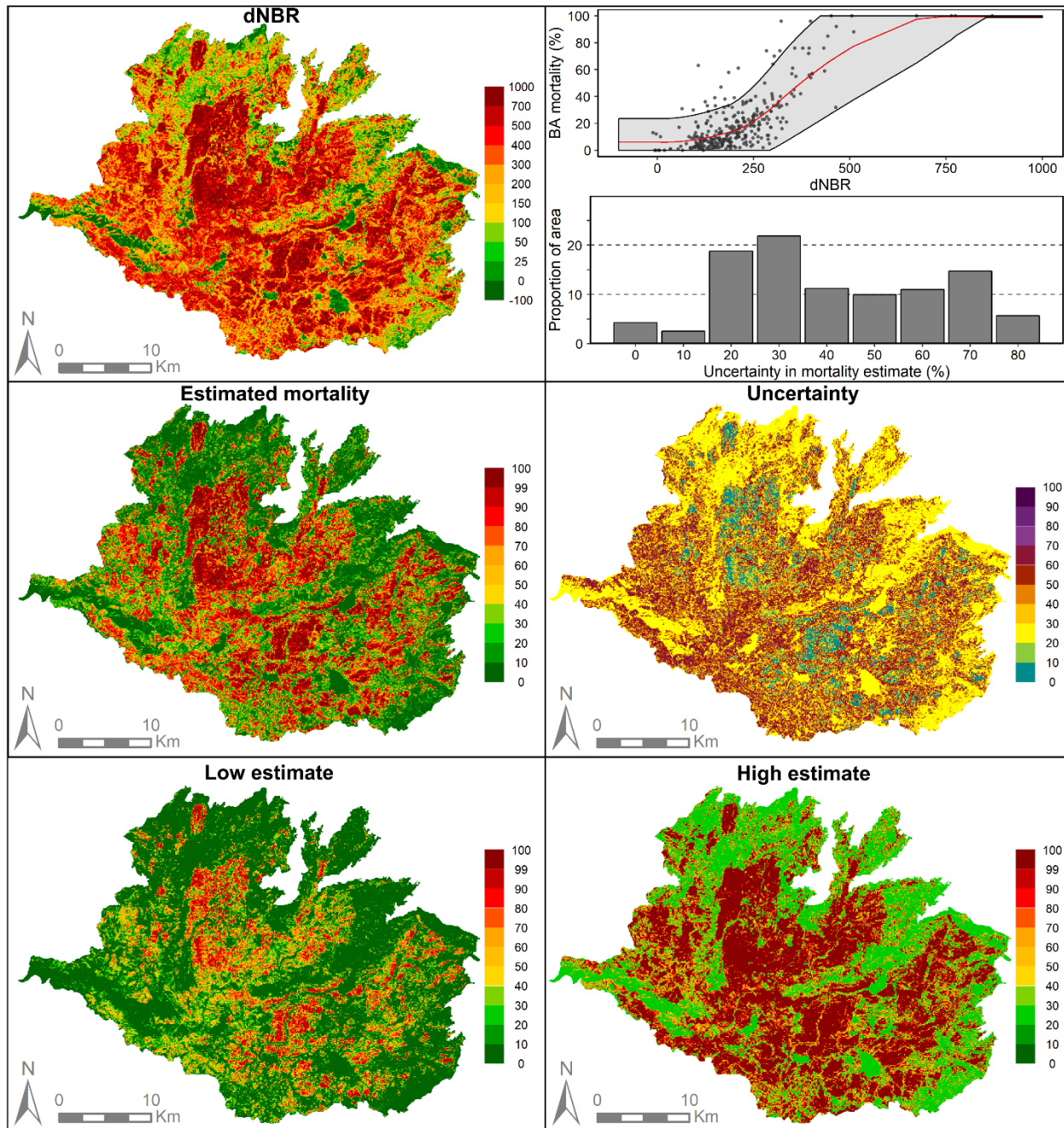
Fig. 5. Satellite-derived burn severity (dNBR) of the California Rim Fire. The scatterplot displays the relationship between dNBR and observed percent basal area mortality (red line is predicted mortality rate of LOESS model; shaded region is 95% confidence envelope). The histogram shows proportion of area within the Rim Fire at various levels of uncertainty. Maps show mean, low, and high estimates of percent basal area mortality. The uncertainty map displays the range in predicted percent basal area mortality necessary to capture the true mortality rate, 95% of the time (i.e., a 40% uncertainty level indicates a ±20% range in expected mortality levels). Iterations of this figure depicting 95% confidence envelopes around percent stem mortality and 68% confidence envelopes around basal area and stem mortality may be found in the supplemental information (Fig. S5&S6&S7).

Discussion

The satellite's perspective

Spectral indices most accurately detected mortality of stems ≥50 cm DBH, an intuitive result considering that medium-diameter stems are the most prominent component of the upper canopy and therefore dominate the satellite's view of the forest, simultaneously concealing many of the smaller diameter stems in the understory (see also Jeronimo et al. 2019). We might also expect spectral indices to be highly sensitive to large-diameter (≥100 cm DBH) tree mortality, but the lower relative abundance and reduced susceptibility to immediate fire-induced mortality (Furniss et al. 2019) appears to have reduced their contribution to the overall spectral response. The prominence of stems ≥50 cm DBH in the satellite view of the canopy also explains the strong correlation between spectral indices and %BA mortality, as trees in this diameter class are primary contributors to overall forest BA in many forests (Lutz et al. 2012, 2013, 2018a).

Spectral index comparison

No single index estimated all forms of observed mortality most accurately. Perhaps the most commonly used spectral index, dNBR, was among the top three indices for most mortality categories, but it was rarely the most accurate. RdNBR and RBR, two indices introduced to improve upon dNBR (Miller and Thode 2007, Parks et al. 2014), did not offer consistent improvement. dNBR was more accurate than RdNBR in every category other than stem mortality of trees ≥100 cm DBH, while dNBR and RBR were more similar. dNBR has been recognized as outperforming RdNBR under some conditions (Cansler and McKenzie 2012, Meddens et al. 2016, McCarley et al. 2017, 2018, Veraverbeke and Hook 2013), and although relativized versions of dNBR may increase accuracy in some contexts, they were not unconditionally better in this study. The benefits of relativized indices may be less dramatic for individual fires at the stand- to landscape-scale, but they may still enhance accuracy when applied to multiple fires at the regional scale. Other indices that have been developed to detect burn severity were less accurate; dNDMI and CSI had accuracies of 41% and 46% for percent stem and percent BA mortality, respectively, while dMIRBI had 0% VE for both types of mortality (Fig. S3&S4, Table S2).

dNBR is likely the best index for general use as it was within the top three indices for both percent stem and percent BA mortality (Table 2), but specific aspects of fire-induced structural changes may be estimated more accurately by using dSWIR1.NIR to estimate BA mortality and dNDVI to estimate stem mortality. If large-diameter trees are of specific concern, dR or dSWIR2 may be more useful. Compared to dNBR, dR improved accuracy by 15% for stems ≥50 cm DBH, and dSWIR2 improved accuracy by 26% for stems ≥100 cm DBH. This result may be of particular interest to carbon modeling research as large-diameter trees contribute disproportionately to above ground biomass (Lutz et al. 2017b, 2018a) and carbon sequestration (Stephenson et al. 2014). Landsat-derived spectral indices did not accurately detect small-diameter tree mortality in our study sites because the upper canopy obscured the satellite view of the sub-canopy trees. In forests with multi-layered canopies, Landsat-derived spectral indices may be an inadequate tool to assess aspects of burn severity that are related to small-diameter tree mortality (e.g., wildlife habitat, advanced regeneration, biodiversity).

A surprising result was that dNDVI was more accurate than dNBR for detecting percent stem mortality for trees ≥10 cm DBH (48% versus 40% VE). A plausible explanation is that dNBR uses the SWIR2 band which is primarily sensitive to newly exposed ash and mineral soil (Miller and Yool 2002, van Wagtendonk et al. 2004), and a high percent stem mortality can be achieved without actually exposing much forest floor if mortality is comprised mostly of trees <10 cm

DBH (Figs. 1&6). High BA mortality will more reliably expose forest floor compared to high stem mortality (because medium- and large-diameter trees occupy more canopy surface area; Fig. 6), commanding a tighter relationship between BA mortality and post-fire reflectance in the SWIR2 band (Fig. S8). Conversely, dNDVI replaces the SWIR2 band with the red band which is more sensitive to the density and health of vegetation (Tucker 1979). dNDVI may therefore be less sensitive to newly exposed forest floor and more sensitive to the reduction in canopy density and depth associated with mortality of shrubs and trees <50 cm DBH that are more numerically abundant (Lutz et al. 2014, 2017a) yet contribute little to overall BA. dNDVI may also be more sensitive to the mortality associated with non-crowning fire behavior that leaves many of the needles in the canopy intact, as was the case for the majority of our study area (Figs. 2,S1&S2). As needles may be killed but not consumed by surface (non-crowning) fire, they continue to obscure the forest floor post-fire which reduces the spectral response of the SWIR2 band while simultaneously inducing a strong response in the red band due to their senescence and loss of chlorophyll.
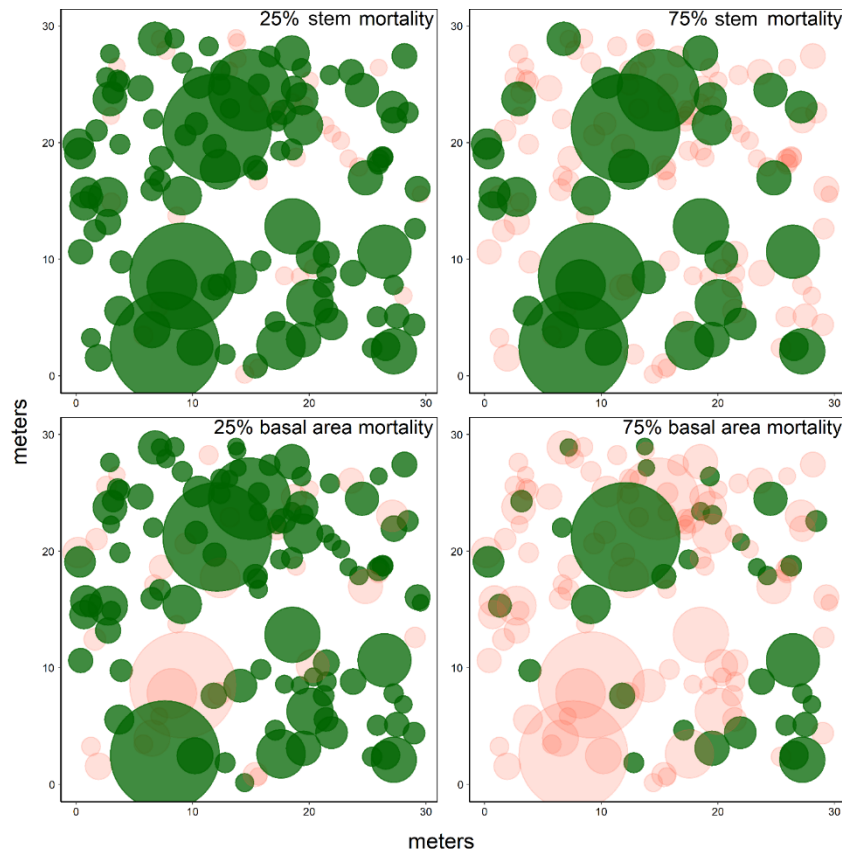

Fig. 6. Inherent uncertainties in interpreting tree mortality from Landsat-derived data. Changes in canopy cover associated with fire-induced mortality are shown as a percentage of pre-fire stem density (top panels) and basal area (bottom panels) mortality. The points represent the location of individual trees contained within a single Landsat pixel within the Yosemite Forest Dynamics Plot (pixel 120 in Fig. S1). The ecological effect of mortality depends on both the number and size of trees killed, but different combinations of mortality can elicit the same spectral response when averaged over a 30-m Landsat pixel. Mortality was simulated by identifying either the smallest trees (top panels) or a stratified selection from the full range of diameters (bottom panels). The left panels show changes in canopy cover associated with low-severity fire, while the right panels show changes associated with moderate- to high-severity fire. Green circles represent the two-dimensional crown footprint of surviving trees, red circles represent fire-killed trees. Canopy diameters were scaled according to the diameter at breast height (1.37 m) for each tree.

An additional explanation is that mortality of large trees requires greater fire intensity compared to mortality of smaller trees (due to thicker bark and higher crown base height; Hood et al. 2018), and this higher fire intensity may be associated with more duff consumption and exposed mineral soil post-fire. Both of these explanations are supported by post-hoc analyses demonstrating that the red band was most strongly correlated with percent stem mortality (compared to NIR and SWIR2), while the SWIR2 band was more strongly correlated with percent BA mortality (Fig. S8). Based on these correlations, one would expect the index containing the Red band (dNDVI) to have a higher sensitivity to stem mortality and the index containing the SWIR2 band (dNBR) to be more sensitive to BA mortality.

The best spectral index may depend on forest type, successional stage, and the history of disturbance. dNDVI had high accuracy in the mixed conifer forests within our study area (density between 42 and 1233 stems $\geq$10 cm ha^{-1}; Table S1), but indices that incorporate a SWIR band may be more appropriate in more open forests because of their sensitivity to ash and mineral soil on the forest floor. RdNBR and RBR were not consistently better than dNBR within our study area, but they may enhance accuracy when applied to broad spatial or temporal scales that contain greater variability in biophysical conditions (e.g., Miller and Thode 2007, Parks et al. 2014, 2018, Harvey et al. 2019).

Challenges in detecting tree mortality from space

The 30-m spatial resolution of Landsat pixels imposes some inescapable uncertainties associated with calibrating satellite-derived spectral indices to field-based metrics of burn severity, and this may be especially pronounced in ecosystems with heterogeneous vegetation and fire behavior (Morgan et al. 2014). Landsat pixels are broader than the scale of individual trees; the crown spread of even the largest *P. lambertiana* is <20 m (Van Pelt 2001), which combined with the dispersed spatial pattern of large-diameter trees (Lutz et al. 2012) represents a maximum of 37% of the area contained within a single pixel. Furthermore, fire behavior in the Sierra Nevada can be heterogeneous at very fine spatial scales (Kolden et al. 2012, Meddens et al. 2018a, Blomdahl et al. 2019, Furniss et al. 2019). Spectral reflectance for each pixel in a low- to moderate-severity fire therefore represents a mixture of patches experiencing different levels of mortality; the mortality of one large-diameter tree may elicit the same spectral response as the mortality of many smaller trees (Fig. 6).

Consider a pair of pixels chosen to illustrate how these ecologically disparate scenarios that can appear spectrally identical (Fig. 1). These two pixels experienced basal area mortality rates of 16% and 96%, but dNBR was the same (dNBR ~ 323; Fig. 1, top row). The plot with the low mortality rate had a high pre-fire density (644 stems ha^{-1}), and mortality was entirely stems <41 cm DBH. Conversely, the plot with a high basal area mortality rate had a lower absolute mortality rate (299 vs 411 mortalities $\geq$10 cm DBH ha^{-1}), but a much higher mortality rate of stems $\geq$50 cm DBH (109 vs 0 mortalities $\geq$10 cm DBH ha^{-1}). Relativized indices such as RdNBR or RBR may reduce this variability, but neither can entirely eliminate this problem of scale. Both RdNBR and RBR were more closely related to observed percent BA mortality compared to dNBR for the present pair of plots (RdNBR = 405 vs. RdNBR = 905; Fig. 1, top row), but this relationship was not consistent. Consider another pair of plots with similar dNBR values (dNBR ~ 155; Fig. 1, bottom row); the plot with high pre-fire density (505 stems >10 cm ha^{-1}) experienced 39% BA mortality with RdNBR = 186 and RBR = 91, while the plot with low pre-fire density (42 stems >10 cm ha^{-1}) experienced only 4% BA mortality with RdNBR = 733 and RBR = 148. In the high-density plot the surviving overstory trees obscured the satellite's view of actual mortality and resulted in an under-estimate of severity (i.e., low index values), while shrub

mortality and soil scorching in the low-density plot elicited a strong spectral response and an over-estimate of severity (Fig. 1). Both RdNBR and RBR inflated the satellite-derived index value for the low-density plots, but this did not consistently improve their correspondence to field-based measurements of severity. These case studies were not chosen to be representative of average conditions; rather, they were selected to reveal some of the confounding factors that can compromise the accuracy of satellite-derived spectral indices in heterogeneous environments. If one is interested in average conditions, much of the uncertainty illustrated by this case study will average out. If one is interested in heterogeneity, however, these plots serve as a useful example of the range in fire effects that can be indistinguishable with 30-m Landsat pixels.

Higher-resolution imagery can be used to reduce uncertainty associated with the 30-m pixel scale (e.g., Meng et al. 2017), but the temporal resolution of the Landsat and Sentinel satellites remains unmatched by sources of satellite imagery with finer spatial resolutions. The high temporal resolution (16 days for Landsat-8 and 10 days for Sentinel-2) and 35-year legacy of the Landsat program have made Landsat imagery indispensable for the remote detection of fire severity, rendering the uncertainty associated with a 30-m pixel scale a persistent feature of most fire severity maps. Analytical techniques such as spectral mixture analysis (e.g., Quintano et al. 2013) may be used to reduce this uncertainty without relying on higher resolution imagery, but there remains a need for more validations of novel analytical approaches with field-based data.

The large range in spectral index values at high levels of observed mortality (Fig. 4) was likely due to the fact that 100% mortality is the maxima for tree-based metrics of severity, but it does not necessarily represent an endpoint of potential fire effects (e.g., soil scorching, shrub and herb mortality). All trees may be killed without incinerating everything else within pixel (i.e., shrubs, grasses, and organic soil), resulting in a lower spectral index value compared to a pixel in which more of the understory, surface fuels, and organic material in the soil are consumed.

Accounting for uncertainty

This study exposed the high amount of variability in observed mortality levels associated with all spectral indices, especially at intermediate burn severity levels (Fig. 3). Although this variability may appear inconsistent with previous studies that have reported strong correlations between spectral indices and field-based measures of severity (e.g., Miller and Thode 2007, Parks et al. 2014, Meddens et al. 2016, Veraverbeke and Hook 2013), it is critical to differentiate between classification accuracy derived using categorical severity classes (unburned vs. burned; low vs. moderate vs. high) and accuracy metrics based on continuous data. The variability we observed is consistent with other spectral index calibrations based on continuous, field-based measures of burn severity (e.g., Miller et al. 2009, Whitman et al. 2018, Harvey et al. 2019), and it emphasizes the persistent need for more quantitative, field-based evaluations of satellite-derived severity products that has existed for over a decade (Lentile et al. 2006).

We suggest a few ways for users of satellite-derived severity products to cope with this uncertainty: 1) use continuous estimates of severity and associated variance (e.g., tree mortality is predicted to be between 50% and 75%), 2) associate a probability level to categorical severity classes (e.g., a moderate-severity pixel may have a 75% chance of being correctly classified), and 3) supplement satellite-derived estimates of burn severity with field-based observations. These approaches will facilitate a more nuanced and ecologically-relevant interpretation of satellite-derived severity metrics. Variability that is not detected by satellite-derived spectral indices represents actual heterogeneity in fire effects that can be of great ecological significance (Kolden et al. 2015, Cansler et al. 2018, Meddens et al. 2018b, Blomdahl et al. 2019), and this could have broad implications for applications ranging from quantifying spatial patterns in burn

severity to planning post-fire management actions.

Accounting for uncertainty is particularly important in areas that burn at intermediate severities, particularly when burn severity indices are used to parameterize further calculations (e.g., carbon emissions; Stenzel et al. 2019), examine spatial patterns in burn severity (e.g., Collins et al. 2017, Stevens et al. 2017, Meddens et al. 2018), evaluate landscape change and restoration strategies (e.g., Kane et al. 2014, Becker and Lutz 2016, Blomdahl et al. 2019), or to examine future fire vulnerability (e.g., Smith et al. 2014). These areas can represent large proportions of the area within a fire footprint (Fig. 5, S2). In the case of the Rim Fire, half of the fire footprint (>50,000 ha) had an uncertainty in predicted mortality of over 37% (±19%) for percent BA mortality and over 46% (±23%) for percent stem mortality (Figs. 5&S5). This is consistent with other studies that have shown uncertainty can severely compromise the accuracy of ecological models when they are applied across heterogeneous landscapes (Hunsaker et al. 2013, Wu et al. 2006, Harmon et al. 2015, Lutz et al. 2017b, Furniss et al. 2017, 2019), and this can limit their utility to resource managers and policy makers (Funtowicz and Ravetz 1990).

Some of this uncertainty will average out when fire effects are assessed at a large enough scale, potentially yielding estimates of mortality that accurately capture average conditions. The higher accuracy we observed for the JFSP plots compared to the YFDP (Table 2) supports this point; we calculated area-weighted averages for the 50×50 m JFSP plots which contained 4 to 9 partial pixels (Fig. S1), and this reduced the potential range of variability in observed mortality compared to the YFDP pixels which were assessed at a 30×30 m scale. We conducted a post-hoc test of this hypothesis by aggregating the YFDP pixels into groups of 4 (2×2 pixels), 9 (3×3 pixels), and 16 (4×4 pixels) and assessing %VE with random forest. We found that accuracy increased as we aggregated pixels up to the 3×3 pixel scale (67% compared to 38% variance explained at the individual-pixel scale), but accuracy declined as we continued to aggregate (50% at the 4×4 pixel scale).

If the metric of interest is simply the amount of area burned at a given severity, and if the scope of inference is broad enough, then the uncertainty that this study exposed may be of little importance. However, for applications that rely on burn areas that have been characterized as homogeneous (e.g., patch metrics *sensu* Keane et al. [2008], distance to seed source, etc.), uncertainty in fire effects should be carefully considered. *While error will average out at large scales, the heterogeneity that this uncertainty represents does not disappear.* In other words, a 46% range in mortality does not mean that mortality predictions will be off by 46%; it means that an area that appears homogeneous based on Landsat-derived spectral indices may actually contain pixels in which mortality was 23% higher or lower than the average conditions.

The single-fire approach

With a single-fire study there is always the question about generalizability to other fires and biogeographical regions, but the single-fire approach can provide a unique perspective that is more akin to how managers use satellite-derived severity maps to guide post-fire management and restoration. This approach allowed us to isolate variability due to intrinsic, within-landscape factors without the additional variance due to extrinsic factors associated with fires that burn in different biogeographical regions and under different climates (e.g., Harvey et al. 2019).

We note that neither approach is inherently better; the most optimal approach will depend on the nature of the desired inferences. Multiple-fire studies offer greater capacity to parameterize models and generalize across broad regions, while single-fire studies permit a more precise evaluation of the maximum accuracy that may be attained with a Landsat-derived severity map for any given fire. By choosing a fire characteristic of the new fire regime in the Sierra Nevada

(van Wagtendonk and Lutz 2007, Lutz et al. 2009, Lydersen et al. 2014) and establishing a large calibration dataset with a wide range of severities (Fig. S2), we were able to consistently evaluate the relationship between spectral indices, tree-based mortality, and the uncertainty of the estimates. This provides a rigorous estimate of maximal severity index accuracy for Sierra Nevada mixed-conifer forests.

This study is also unique in that we relied heavily on tree mortality data from a single large study plot that contained many contiguous Landsat pixels; this is distinctly different from the 30-m Composite Burn Index (CBI) plots that have been widely used to calibrate spectral indices for over a decade (Key and Benson 2006, Eidenshink et al. 2007, Miller and Thode 2007, Parks et al. 2014, 2018). CBI calibration plots are typically located in patches of homogeneous fire severity to minimize uncertainty due to co-registration error between the field plots and Landsat pixels, and spectral response values are calculated based on an area-weighted average of up to four different Landsat pixels. Two key problems with the CBI approach are that it may artificially inflate perceived accuracy by aggregating pixel values (i.e., the aggregation problem, Marceau et al. 1994), and it precludes the comparison of field measured severity between adjacent pixels (because field plots are not contiguous). We chose to evaluate spectral index accuracy using a large, fully censused forest plot because it enabled us to minimize potential bias associated with pixel aggregation and to quantify the correlation between field-based measures of fire effects and the spectral response of *individual* Landsat pixels. There are indeed both benefits and limitations to our 'big plot' design (Lutz 2015), but it provides a valuable contrast with the existing remote detection of fire severity literature.

Another consideration with our study design potential for spatial auto-correlation due to patchy fire behavior within the YFDP (e.g., Fig. S1). Semivariograms revealed that while mortality was spatially auto-correlated at spatial scales <100 m, uncertainty was spatially random (Fig. S9). In other words, the difference between predicted mortality rate and observed mortality rate for any given pixel was independent of surrounding pixels. We did not explicitly control for the autocorrelation in mortality because we were primarily interested in the variance between observed mortality and spectral response (which was not autocorrelated), and previous research has shown that the spatial autocorrelation of fire has a negligible influence on assessment of fire effects in the Sierra Nevada (van Mantgem and Schwik 2009).

The results of the rarefied spectral index validation provide further evidence that spatially auto-correlated fire behavior did not compromise our results. The rarefied sampling procedure greatly reduced spatial auto-correlation by sampling 20 plots at a time; greatly reducing the chance that neighboring pixels would be sampled. The average distance between any two YFDP pixels was 296 m, far greater than the 100 m scale at which mortality was auto-correlated.

Advancing the ecological relevance of future spectral index calibrations

Continuing to perform calibrations with continuous, field-based metrics of severity will enhance the ecological interpretation of satellite-derived severity maps (e.g., Whitman et al. 2018, Harvey et al. 2019). Tree-based metrics such as percent stem mortality, change in canopy cover, and basal area mortality can provide more well-defined indicators burn severity (Morgan et al. 2014); they are preferable to semi-quantitative metrics such as the Composite Burn Index (Key and Benson 2006) because they are more directly applicable to post-fire management (Kolden et al. 2015), and they allow uncertainty to be assessed independently for different elements of forest ecosystems (i.e., soil, shrubs, large trees, etc.).

Traditional accuracy metrics such as R^2 and AIC will continue to be useful ways to compare indices (e.g., Miller et al. 2009, Whitman et al. 2018, Parks et al. 2018), but we suggest

continuing to explore novel ways of quantifying accuracy that reveal error in ecologically-relevant terms (e.g., Figs. 2 – 4; Harvey et al. 2019) such as the range in tree mortality that should be expected based on spectral index values. The development of region-specific variance models (e.g., Figs. 4, S5 – S7) may be used to create uncertainty maps which will permit more ecologically accurate interpretation of satellite-derived severity; the variance model we developed in this study may be adequate for Sierra Nevada mixed-conifer forests, but additional models would improve performance for other forest types.

Finally, there is a general need to enhance the theoretical foundation from which we develop, calibrate, and compare spectral indices. Improving our understanding of how various ecosystem attributes influence spectral reflectance will facilitate a more cohesive synthesis of the inconsistencies in the literature regarding which spectral index is best (e.g., Roy et al. 2006, Miller and Thode 2007, Cansler and McKenzie 2012, Parks et al. 2014, McCarley et al. 2018).

Conclusion

This study makes three key contributions to the field of remote sensing of fire effects. First, we conducted a comprehensive comparison of previously published severity indices using the largest set of contiguous, field-based, georeferenced individual-tree level mortality data to date. This revealed that both absolute and relative spectral index accuracy depends on the tree mortality metric of interest, and different indices may be optimal for different ecological objectives. Second, we conducted this study in a single, large fire that enabled us to isolate variability due to intrinsic, within-landscape factors without the additional variance due to extrinsic factors associated different biogeographies or climatic conditions. This permitted a detailed evaluation of the maximum accuracy that may be attained with Landsat-derived spectral indices for any given fire, and this revealed a great deal of persistent uncertainty that may reflect a fundamental accuracy limit due to the spatial and spectral resolution of the Landsat 8 OLI sensor. Finally, we identified the range in tree mortality that may be indistinguishable based on spectral indices derived from Landsat satellites, and we demonstrated how this variability translates into uncertainty in fire effects and patterns in burn severity at the landscape scale.

The range in observed tree mortality was highest at intermediate spectral index values, with a range in expected mortality as high as 70% (±35%) of stems and 80% (±40%) of basal area (Fig. 5&S5). This uncertainty in observed tree mortality reveals that apparently homogeneous patches may actually contain a considerable amount of variability in fire effects and post-fire dynamics. This may be dealt with by estimating severity with tree-based metrics of fire effects (*sensu* Kolden et al. 2015, Harvey et al. 2019) and associated estimates of variance, calculating the probability that a pixel is classified in the correct categorical severity class, or supplementing remotely-sensed data with field-based observations.

Explicitly accounting for uncertainty in satellite-derived estimates of burn severity will facilitate a more ecologically-nuanced and accurate interpretation of fire effects. The high degree of uncertainty in actual tree mortality that we observed challenges the confidence with which Landsat-derived spectral indices have been used to measure fire effects, and this has broad implications for any studies or management actions that rely on accurate assessments of patterns in fire severity, distribution of seed sources, persistence of fire refugia, or post-fire landscape complexity.

Data access

Complete data for the 53 Joint Fire Science Program plots (Macriss et al. 2019) is available at Utah State University DigitalCommons.

Declaration of interest

The authors declare no conflicts of interest.

Literature cited

Baig, M. H. A., L. Zhang, T. Shuai, and Q. Tong. 2014. Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. Remote Sensing Letters 5:423–431.

Barbosa, P. M., D. Stroppiana, J.-M. Grégoire, and J. M. Cardoso Pereira. 1999. An assessment of vegetation fire in Africa (1981-1991): Burned areas, burned biomass, and atmospheric emissions. Global Biogeochemical Cycles 13:933–950.

Barth, M. A. F., A. J. Larson, and J. A. Lutz. 2015. A forest reconstruction model to assess changes to Sierra Nevada mixed-conifer forest during the fire suppression era. Forest Ecology and Management 354:104–118.

Becker, K. M. L., and J. A. Lutz. 2016. Can low-severity fire reverse overstory compositional change in montane forests of the Sierra Nevada, USA? Ecosphere 7(12): e01484. http://dx.doi.org/10.1002/ecs2.1484

Blomdahl, E. M., C. A. Kolden, A. J. H. Meddens, and J. A. Lutz. 2019. The importance of small fire refugia in the central Sierra Nevada, California, USA. Forest Ecology and Management 432:1041–1052.

Blomdahl, E. M., C. M. Thompson, J. R. Kane, V. R. Kane, D. J. Churchill, L. M. Moskal, and J. A. Lutz. 2019. Forest structure predictive of fisher (*Pekania pennanti*) dens exists in recently burned forest in Yosemite, California, USA. Forest Ecology and Management 444: 174-186.

Brewer, C. K., J. C. Winne, R. L. Redmond, D. W. Opitz, and M. V. Mangrich. 2005. Classifying and mapping wildfire severity: a comparison of methods. Photogrammetric Engineering & Remote Sensing:10.

CAL FIRE. 2018. Top 20 largest California wildfires.

Cansler, C. A., and D. McKenzie. 2012. How Robust Are Burn Severity Indices When Applied in a New Region? Evaluation of Alternate Field-Based and Remote-Sensing Methods. Remote Sensing 4:456–483.

Cansler, C. A., D. McKenzie, and C. B. Halpern. 2018. Fire enhances the complexity of forest structure in alpine treeline ecotones. Ecosphere 9:e02091.

Cansler, C. A., M. E. Swanson, T. J. Furniss, A. J. Larson, and J. A. Lutz. 2019. Fuel dynamics after reintroduced fire in an old-growth Sierra Nevada mixed-conifer forest. Fire Ecology 15:16. doi: https://doi.org/10.1186/s42408-019-0035-y

Collins, B. M., J. D. Miller, E. E. Knapp, and D. B. Sapsis. 2019. A quantitative comparison of forest fires in central and northern California under early (1911–1924) and contemporary (2002–2015) fire suppression. International Journal of Wildland Fire.

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random Forests for classification in ecology. Ecology 88:2783–2792.

Eidenshink, J., B. Schwind, K. Brewer, Z.-L. Zhu, B. Quayle, and S. Howard. 2007. A project for monitoring trends in burn severity. Fire Ecology 3:3–21.

Epting, J., D. Verbyla, and B. Sorbel. 2005. Evaluation of remotely sensed indices for assessing burn severity in interior Alaska using Landsat TM and ETM+. Remote Sensing of Environment 96:328–339.

Funtowicz, S. O., and J. R. Ravetz. 1990. Uncertainty and quality in science for policy. Springer Science & Business Media.

Furniss, T. J., A. J. Larson, and J. A. Lutz. 2017. Reconciling niches and neutrality in a subalpine temperate forest. Ecosphere 8(6).

Furniss, T. J., A. J. Larson, V. R. Kane, and J. A. Lutz. 2019. Multi-scale assessment of post-fire tree mortality models. International Journal of Wildland Fire 28:46–61.

Gao, B. 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment 58:257–266.

Harmon, M. E., B. Fasth, C. B. Halpern, and J. A. Lutz. 2015. Uncertainty analysis: an evaluation metric for synthesis science. Ecosphere 6:art63.

Harris, L., and A. H. Taylor. 2015. Topography, fuels, and fire exclusion drive fire severity of the Rim fire in an old-growth mixed-conifer forest, Yosemite National Park, USA. Ecosystems 18:1192–1208.

Harvey, B. J., R. A. Andrus, and S. C. Anderson. 2019. Incorporating biophysical gradients and uncertainty into burn severity maps in a temperate fire-prone forested region. Ecosphere.

Hijmans, R. J. 2016. raster: Geographic Data Analysis and Modeling. R package version 2.5-8.

Hood, S. M., J. M. Varner, P. van Mantgem, and C. A. Cansler. 2018. Fire and tree death: understanding and improving modeling of fire-induced tree mortality. Environmental Research Letters 13:113004.

Huete, A. . 1988. A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment 25:295–309.

Hunsaker, C. T., M. F. Goodchild, M. A. Friedl, and T. J. Case, editors. 2013. Spatial uncertainty in ecology: implications for remote sensing and GIS applications. Springer, New York.

Jeronimo, S. M. A., V. R. Kane, D. J. Churchill, J. A. Lutz, M. P. North, G. P. Asner, and J. F. Franklin. 2019. Forest structure and pattern vary by climate and landform across active-fire landscapes in the montane Sierra Nevada. Forest Ecology and Management 437:70–86.

Jeronimo, S. M., J. A. Lutz, V. R. Kane, A. J. Larson, and J. F. Franklin. In press. Burn weather and three-dimensional fuel structure determine post-fire tree mortality. Landscape Ecology.

Kane, V. R., C. A. Cansler, N. A. Povak, J. T. Kane, R. J. McGaughey, J. A. Lutz, D. J. Churchill, and M. P. North. 2015. Mixed severity fire effects within the Rim fire:

　　　　Relative importance of local climate, fire weather, topography, and forest structure. Forest Ecology and Management 358:62–79.

Kane, V. R., M. North, J. A. Lutz, D. Churchill, S. L. Roberts, D. F. Smith, R. J. McGaughey, J. T. Kane, and M. L. Brooks. 2014. Assessing fire-mediated change to forest spatial structure using a fusion of Landsat and airborne LiDAR data in Yosemite National Park. Remote Sensing of Environment 151: 89-101.

Keane, R. E., J. K. Agee, P. Fulé, J. E. Keeley, C. Key, S. G. Kitchen, R. Miller, and L. A. Schulte. 2008. Ecological effects of large fires on US landscapes: benefit or catastrophe? International Journal of Wildland Fire 17:696.

Key, C. H., and N. C. Benson. 2006. Landscape assessment (LA). Gen. Tech. Rep. RMRS-GTR-164-CD, Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Kolden, C. A., J. T. Abatzoglou, J. A. Lutz, C. A. Cansler, J. T. Kane, J. W. V. Wagtendonk, and C. H. Key. 2015. Climate contributors to forest mosaics: ecological persistence following wildfire. Northwest Science 89:219–238.

Kolden, C. A., J. A. Lutz, C. H. Key, J. T. Kane, and J. W. van Wagtendonk. 2012. Mapped versus actual burned area within wildfire perimeters: characterizing the unburned. Forest Ecology and Management 286:38–47.

Kolden, C. A., A. M. S. Smith, and J. T. Abatzoglou. 2015. Limitations and utilisation of Monitoring Trends in Burn Severity products for assessing wildfire severity in the USA. International Journal of Wildland Fire 24:1023–1028.

Kushla, J. D., and W. J. Ripple. 1998. Assessing wildfire effects with Landsat thematic mapper data. International Journal of Remote Sensing 19:2493–2507.

Larson, A. J., C. A. Cansler, S. G. Cowdery, S. Hiebert, T. J. Furniss, M. E. Swanson, and J. A. Lutz. 2016. Post-fire morel (*Morchella*) mushroom abundance, spatial structure, and harvest sustainability. Forest Ecology and Management 377:16–25.

Lentile, L. B., Z. A. Holden, A. M. S. Smith, M. J. Falkowski, A. T. Hudak, P. Morgan, S. A. Lewis, P. E. Gessler, and N. C. Benson. 2006. Remote sensing techniques to assess active fire characteristics and post-fire effects. International Journal of Wildland Fire 15:319.

Liaw, A., and M. Wiener. 2002. Classification and Regression by randomForest. R News 2:18–22.

Lutz, J. A. 2015. The evolution of long-term data for forestry: large temperate research plots in an era of global change. Northwest Science 89:255–269.

Lutz, J. A., T. J. Furniss, S. J. Germain, K. M. L. Becker, E. M. Blomdahl, S. M. A. Jeronimo, C. A. Cansler, J. A. Freund, M. E. Swanson, and A. J. Larson. 2017a. Shrub communities, spatial patterns, and shrub-mediated tree mortality following reintroduced fire in Yosemite National Park, California, USA. Fire Ecology 13:104–126.

Lutz, J. A., T. J. Furniss, D. J. Johnson, S. J. Davies, D. Allen, A. Alonso, K. J. Anderson-Teixeira, A. Andrade, J. Baltzer, K. M. L. Becker, E. M. Blomdahl, N. A. Bourg, S. Bunyavejchewin, D. F. R. P. Burslem, C. A. Cansler, K. Cao, M. Cao, D. Cárdenas, L.-W. Chang, K.-J. Chao, W.-C. Chao, J.-M. Chiang, C. Chu, G. B. Chuyong, K. Clay, R. Condit, S. Cordell, H. S. Dattaraja, A. Duque, C. E. N. Ewango, G. A. Fischer, C. Fletcher, J. A. Freund, C. Giardina, S. J. Germain, G. S. Gilbert, Z. Hao, T. Hart, B. C. H. Hau, F. He, A. Hector, R. W. Howe, C.-F. Hsieh, Y.-H. Hu, S. P. Hubbell, F. M. Inman-Narahari, A. Itoh, D. Janík, A. R. Kassim, D. Kenfack, L. Korte, K. Král, A. J. Larson, Y. Li, Y. Lin, S. Liu, S. Lum, K. Ma, J.-R. Makana, Y. Malhi, S. M. McMahon, W. J.

McShea, H. R. Memiaghe, X. Mi, M. Morecroft, P. M. Musili, J. A. Myers, V. Novotny, A. de Oliveira, P. Ong, D. A. Orwig, R. Ostertag, G. G. Parker, R. Patankar, R. P. Phillips, G. Reynolds, L. Sack, G.-Z. M. Song, S.-H. Su, R. Sukumar, I.-F. Sun, H. S. Suresh, M. E. Swanson, S. Tan, D. W. Thomas, J. Thompson, M. Uriarte, R. Valencia, A. Vicentini, T. Vrška, X. Wang, G. D. Weiblen, A. Wolf, S.-H. Wu, H. Xu, T. Yamakura, S. Yap, and J. K. Zimmerman. 2018a. Global importance of large-diameter trees. Global Ecology and Biogeography:849–864.

Lutz, J. A., and C. B. Halpern. 2006. Tree mortality during early forest development: a long-term study of rates, causes, and consequences. Ecological Monographs 76(2): 257-275.

Lutz, J. A., C. H. Key, C. A. Kolden, J. T. Kane, and J. W. van Wagtendonk. 2011. Fire frequency, area burned, and severity: a quantitative approach to defining a normal fire year. Fire Ecology 7:51–65.

Lutz, J. A., A. J. Larson, J. A. Freund, M. E. Swanson, and K. J. Bible. 2013. The importance of large-diameter trees to forest structural heterogeneity. PLoS ONE 8:e82784.

Lutz, J. A., A. J. Larson, M. E. Swanson, and J. A. Freund. 2012. Ecological importance of large-diameter trees in a temperate mixed-conifer forest. PLoS ONE 7:e36131.

Lutz, J. A., A. Larson, and M. Swanson. 2018b. Advancing fire science with large forest plots and a long-term multidisciplinary approach. Fire 1:5.

Lutz, J. A., J. Matchett, L. Tarnay, D. Smith, K. M. L. Becker, T. J. Furniss, and M. Brooks. 2017b. Fire and the distribution and uncertainty of carbon sequestered as aboveground tree biomass in Yosemite and Sequoia & Kings Canyon National Parks. Land 6:10.

Lutz, J. A., K. A. Schwindt, T. J. Furniss, J. A. Freund, M. E. Swanson, K. I. Hogan, G. E. Kenagy, and A. J. Larson. 2014. Community composition and allometry of Leucothoe davisiae, Cornus sericea, and Chrysolepis sempervirens. Canadian Journal of Forest Research 44:677–683.

Lutz, J. A., J. W. van Wagtendonk, A. E. Thode, J. D. Miller, and J. F. Franklin. 2009. Climate, lightning ignitions, and fire severity in Yosemite National Park, California, USA. International Journal of Wildland Fire 18:765–774.

Lydersen, J. M., M. P. North, and B. M. Collins. 2014. Severity of an uncharacteristically large wildfire, the Rim Fire, in forests with relatively restored frequent fire regimes. Forest Ecology and Management 328:326–334.

Macriss, N., T. J. Furniss, S. M. A. Jeronimo, E. Crowley, O. W. Germain, S. J. Germain, V. R. Kane, A. J. Larson, and J. A. Lutz. 2019. Data for tree mortality calibration of satellite and LiDAR-derived fire severity estimates. https://doi.org/10.26078/jsz1-3980

Marceau, D. J., P. J. Howarth, and D. J. Gratton. 1994. Remote sensing and the measurement of geographical entities in a forested environment. 1. The scale and spatial aggregation problem. Remote Sensing of Environment 49:93–104.

McCarley, T. R., C. A. Kolden, N. M. Vaillant, A. T. Hudak, A. M. S. Smith, and J. Kreitler. 2017. Landscape-scale quantification of fire-induced change in canopy cover following mountain pine beetle outbreak and timber harvest. Forest Ecology and Management 391:164–175.

McCarley, T. R., A. M. S. Smith, C. A. Kolden, and J. Kreitler. 2018. Evaluating the Mid-Infrared Bi-spectral Index for improved assessment of low-severity fire effects in a conifer forest. International Journal of Wildland Fire 27:407–412.

Meddens, A. J. H., C. A. Kolden, and J. A. Lutz. 2016. Detecting unburned areas within wildfire perimeters using Landsat and ancillary data across the northwestern United States. Remote Sensing of Environment 186:275–285.

Meddens, A. J. H., C. A. Kolden, J. A. Lutz, J. T. Abatzoglou, and A. T. Hudak. 2018a. Spatiotemporal patterns of unburned areas within fire perimeters in the northwestern United States from 1984 to 2014. Ecosphere 9.

Meddens, A. J. H., C. A. Kolden, J. A. Lutz, A. M. S. Smith, C. A. Cansler, J. T. Abatzoglou, G. W. Meigs, W. M. Downing, and M. A. Krawchuk. 2018b. Fire refugia: what are they, and why do they matter for global change? BioScience.

Meng, R., J. Wu, K. L. Schwager, F. Zhao, P. E. Dennison, B. D. Cook, K. Brewster, T. M. Green, and S. P. Serbin. 2017. Using high spatial resolution satellite imagery to map forest burn severity across spatial scales in a Pine Barrens ecosystem. Remote Sensing of Environment 191:95–109.

Miller, J. D., E. E. Knapp, C. H. Key, C. N. Skinner, C. J. Isbell, R. M. Creasy, and J. W. Sherlock. 2009. Calibration and validation of the relative differenced Normalized Burn Ratio (RdNBR) to three measures of fire severity in the Sierra Nevada and Klamath Mountains, California, USA. Remote Sensing of Environment 113:645–656.

Miller, J. D., and A. E. Thode. 2007. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). Remote Sensing of Environment 109:66–80.

Miller, J. D., and S. R. Yool. 2002. Mapping forest post-fire canopy consumption in several overstory types using multi-temporal Landsat TM and ETM data. Remote Sensing of Environment 82:481–496.

Parks, S. A., L. M. Holsinger, M. A. Voss, R. A. Loehman, and N. P. Robinson. 2018. Mean Composite Fire Severity Metrics Computed with Google Earth Engine Offer Improved Accuracy and Expanded Mapping Potential. Remote Sensing 10:879.

Parks, S., G. Dillon, and C. Miller. 2014. A new metric for quantifying burn severity: the relativized burn ratio. Remote Sensing 6:1827–1844.

Quintano, C., A. Fernández-Manso, and D. A. Roberts. 2013. Multiple Endmember Spectral Mixture Analysis (MESMA) to map burn severity levels from Landsat images in Mediterranean countries. Remote Sensing of Environment 136:76–88.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rouse, W., and R. H. Haas. 1974. Monitoring vegetation systems in the Great Plains with ERTS:9.

Roy, D. P., L. Boschetti, and S. N. Trigg. 2006. Remote sensing of fire severity: assessing the performance of the normalized burn ratio. IEEE Geoscience and Remote Sensing Letters 3:112–116.

Scrucca, L. 2011. Model-based SIR for dimension reduction. Computational Statistics & Data Analysis 5:3010–3026.

Smith, A. M. S., N. A. Drake, M. J. Wooster, A. T. Hudak, Z. A. Holden, and C. J. Gibbons. 2007. Production of Landsat ETM+ reference imagery of burned areas within Southern African savannahs: comparison of methods and application to MODIS. International Journal of Remote Sensing 28:2753–2775.

Smith, A. M. S., C. A. Kolden, W. T. Tinkham, A. F. Talhelm, J. D. Marshall, A. T. Hudak, L. Boschetti, M. J. Falkowski, J. A. Greenberg, J. W. Anderson, A. Kliskey, L. Alessa, R. F.

Keefe, and J. R. Gosz. 2014. Remote sensing the vulnerability of vegetation in natural terrestrial ecosystems. Remote Sensing of Environment 154:322–337.

Stenzel, J. E., K. J. Bartowitz, M. D. Hartman, J. A. Lutz, C. A. Kolden, A. M. S. Smith, M. E. Swanson, A. J. Larson, W. J. Parton, and T. W. Hudiburg. 2019. Fixing a snag in carbon emissions estimates from wildfires. Global Change Biology. https://doi.org/10.1111/gcb.14716

Stephenson, N. L., A. J. Das, R. Condit, S. E. Russo, P. J. Baker, N. G. Beckman, D. A. Coomes, E. R. Lines, W. K. Morris, N. Rüger, E. Álvarez, C. Blundo, S. Bunyavejchewin, G. Chuyong, S. J. Davies, á. Duque, C. N. Ewango, O. Flores, J. F. Franklin, H. R. Grau, Z. Hao, M. E. Harmon, S. P. Hubbell, D. Kenfack, Y. Lin, J.-R. Makana, A. Malizia, L. R. Malizia, R. J. Pabst, N. Pongpattananurak, S.-H. Su, I.-F. Sun, S. Tan, D. Thomas, P. J. van Mantgem, X. Wang, S. K. Wiser, and M. A. Zavala. 2014. Rate of tree carbon accumulation increases continuously with tree size. Nature 507:90–93.

Trigg, S., and S. Flasse. 2001. An evaluation of different bi-spectral spaces for discriminating burned shrub-savannah. International Journal of Remote Sensing 22:2641–2647.

Tucker, C. J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment:24.

USGS. 2007. 10m Digital Elevation Model for Yosemite National Park. U.S. Geological Survey, Department of the Interior.

USGS. 2016. LANDSAT 8 (L8) Data Users Handbook Version 2.0. 106 pages. U.S. Geological Survey, Department of the Interior. Sioux Falls, South Dakota, USA.

van Mantgem, P. J., and D. W. Schwilk. 2009. Negligible influence of spatial autocorrelation in the assessment of fire effects in a mixed conifer forest. Fire Ecology 5:10.

Van Pelt, R. 2001. Forest giants of the Pacific coast.

Vermote, E., C. Justice, M. Claverie, and B. Franch. 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. Remote Sensing of Environment 185:46–56.

Vogelmann, J. E. 1990. Comparison between two vegetation indices for measuring different types of forest damage in the north-eastern United States. International Journal of Remote Sensing 11:2281–2297.

van Wagtendonk, J. W., and J. A. Lutz. 2007. Fire regime attributes of wildland fires in Yosemite National Park, USA. Fire Ecology 3:34–52.

van Wagtendonk, J. W., R. R. Root, and C. H. Key. 2004. Comparison of AVIRIS and Landsat ETM+ detection capabilities for burn severity. Remote Sensing of Environment 92:397–408.

Veraverbeke, S., and S. J. Hook. 2013. Evaluating spectral indices and spectral mixture analysis for assessing fire severity, combustion completeness and carbon emissions. International Journal of Wildland Fire 22:707–720.

Whitman, E., M.-A. Parisien, D. K. Thompson, R. J. Hall, R. S. Skakun, and M. D. Flannigan. 2018. Variability and drivers of burn severity in the northwestern Canadian boreal forest. Ecosphere 9:e02128.

Wilson, E. H., and S. A. Sader. 2002. Detection of forest harvest type using multiple dates of Landsat TM imagery. Remote Sensing of Environment 80:385–396.

Wu, J., B. Jones, H. Li, and O. L. Loucks. 2006. Scaling and uncertainty analysis in ecology. Springer.

Detecting tree mortality with Landsat-derived spectral indices: Improving ecological accuracy by examining uncertainty

Tucker J. Furniss[1,†], Van R. Kane[2], Andrew J. Larson[3], James A. Lutz[1]

[1]*Wildland Resources Department and Ecology Center, Utah State University, Logan, UT*
[2]*School of Environmental and Forest Sciences, University of Washington, Seattle, WA*
[3]*Wilderness Institute and Department of Forest Management, University of Montana, Missoula, MT*

[†] E-mail: tucker.furniss@usu.edu

Table S1. Stand structure, spectral index values, and topographic variables for the study areas. Values represent individual Landsat pixels (Yosemite Forest Dynamics Plot) or area-weighted average value for pixels within 0.25-ha plots (Joint Fire Science Program [JFSP] plots).

	Code	Yosemite Forest Dynamics Plot				JFSP plots			
		Min	Max	Mean	Median	Min	Max	Mean	Median
	Stems ≥1 cm DBH ha^{-1}	200	3133	1357	1311	–	–	–	–
	Stems ≥10 cm DBH ha^{-1}	44	1222	532	511	42	1233	391	387
	Basal area (m^2 ha^{-1})	8.5	129.6	64.8	65.5	9.7	102.9	61.3	67.1
Snapshot	NDVI	-0.86	-0.45	-0.65	-0.65	0.22	0.85	0.52	0.50
	MIRBI	-5067	-2153	-3344	-3300	-5401	1173	-3264	-3432
	NBR	135	664	451	452	-250	767	236	255
	CSI	0.99	2.73	1.71	1.66	0.62	3.40	1.29	1.11
	SAVI	0.68	1.29	0.97	0.97	0.33	1.28	0.78	0.76
	NDMI	-7	463	250	250	-239	545	77	49
	SWIR1:NIR	0.37	1.01	0.61	0.60	0.29	1.63	0.92	0.91
	SWIR2:NIR	0.20	0.76	0.39	0.38	0.13	1.67	0.70	0.60
	SWIR2:SWIR1	0.52	0.79	0.63	0.62	0.45	1.02	0.71	0.67
	NIR:G	3.84	9.73	5.44	5.34	1.86	9.24	4.52	4.42
	NIR:R	2.66	13.00	4.93	4.68	1.57	12.66	3.93	3.09
	TCBRI	1343	2535	1762	1738	1536	4802	2273	1989
	TCGRE	535	1760	859	838	107	1500	667	640
	TCWET	-992	102	-235	-225	-2001	257	-706	-565
Bi-temporal	dB	-66.0	8.0	-20.7	-19.0	-256.2	16.1	-46.3	-33.7
	dG	-65	53	-17	-18	-206	32	-39	-25
	dR	-277	5	-93	-88	-474	74	-149	-111
	dNIR	190	943	497	484	-28	2022	495	428
	dSWIR1	-495	171	-92	-84	-1013	88	-251	-132
	dSWIR2	-448	24	-156	-149	-1714	55	-362	-304
	dNDVI	-343	-25	-136	-125	-0.02	0.51	0.19	0.17
	dMIRBI	-2860	665	-667	-525	-7213	1113	-1157	-610
	dNBR	34	464	198	192	-11	870	278	232
	RdNBR	41.6	542.5	249.1	246.0	-37.7	1056.7	404.1	417.9
	RBR	20.3	267.7	121.1	120.0	-10.4	517.9	180.3	177.6
	dCSI	-0.12	1.80	0.60	0.58	-0.23	1.86	0.54	0.38
	dSAVI	0.02	0.50	0.19	0.17	-0.05	0.74	0.26	0.23
	dNDMI	7	445	161	152	-21	675	215	151
	dSWIR1:NIR	-0.54	0.01	-0.18	-0.16	-1.21	0.05	-0.33	-0.29
	dSWIR2:NIR	-0.42	-0.02	-0.16	-0.16	-1.30	0.03	-0.34	-0.31
	dSWIR2:SWIR1	-0.25	0.01	-0.12	-0.11	-0.46	0.00	-0.14	-0.13
	dNIR:G	-0.15	3.29	1.35	1.28	-1.19	4.94	1.08	0.80
	dNIR:R	-0.21	8.85	2.96	2.87	-1.50	8.14	2.13	1.54
	dTCBRI	-26	458	143	129	-402	285	-18	24
	dTCGRE	132	799	406	404	-54	2022	469	423
	dTCWET	-20	720	246	237	-84	1933	426	365
Topo	Slope	5.6	24.5	17.2	17.4	3.9	26.5	13.1	12.3
	Aspect	-0.67	1.00	0.76	0.89	-0.98	1.00	0.04	0.11
	TPI	-3.35	126.06	66.14	65.89	-102.2	177.4	1.04	17.44
	Solar irradiance	1.38	1.89	1.62	1.61	1.50	1.95	1.76	1.77

Table S2. Correlations between spectral indices and observed mortality of trees ≥10 cm diameter at breast height (1.37 m) determined with individual random forest models and summarized using percent variance explained. Values represent "percent variance explained" determined with individual random forest models. Observed mortality was quantified as percent of pre-fire live stems (or basal area [BA]) that was killed by fire. "Initial" columns represent correlations based on a post-fire Landsat scene from immediately after the fire (September 16, 2013), while "extended" columns use a post-fire scene taken one year following the fire (July 1, 2014). Column titles indicate which structural subsets were used to calculate observed mortality. The data were also subset by plot to compare accuracy between the two datasets.

| | Initial | | Extended | | | | | % BA | | |
| | Stems ≥10 cm | Basal area | Basal area | Stems 1-10 cm | Stems ≥10 cm | Stems ≥50 cm | Stems ≥100 cm | YFDP (Δ buffer) | | JFSP |
Index										
NDVI	3.4	32.7	43.4	**9.0**	33.2	34.9	8.3	28.2	-	28.1
dNDVI	26.3	54.8	42.8	-	**49.4**	31.9	6.4	21.0	-	76.6
MIRBI	-	9.4	-	-	-	-	-	-	-	-
dMIRBI	-	-	-	-	-	-	-	-	-	-
NBR	8.5	42.9	35.6	-	18.7	34.5	14.9	9.3	-	13.8
dNBR	26.5	**57.6**	56.7	-	41.5	42.6	3.8	**38.3**	-	**76.8**
RdNBR	18.7	40.7	49.7	-	30.7	40.3	15.2	**36.9**	-	44.3
RBR	**26.9**	**56.3**	57.3	-	34.9	**51.8**	10.3	28.3	-	68.0
CSI	1.0	36.1	45.5	-	19.2	43.4	21.2	8.3	(1)	38.2
dCSI	20.0	2.7	-	-	3.0		-	-	-	38.9
SAVI	2.4	33.2	43.5	5.2	32.2	34.6	8.9	27.8	-	26.8
dSAVI	**26.8**	**54.7**	41.7	-	**48.8**	30.1	7.6	19.2	-	76.5
NDMI	-	32.4	44.4	-	16.6	41.1	23.4	9.3	-	41.3
dNDMI	23.9	47.3	36.9	-	40.5	23.1	8.9	25.7	-	72.2
SWIR1.NIR	-	37.1	46.9	-	20.9	43.6	**25.4**	8.3	(11)	42.4
dSWIR1.NIR	14.1	49.3	**63.1**	-	**42.5**	50.6	8.9	**28.3**	-	73.9
SWIR2.NIR	2.9	39.0	32.6	-	18.3	33.8	9.6	**1.87**	(11)	15.5
dSWIR2.NIR	22.9	52.0	55.2	-	33.9	45.8	**24.4**	29.5	-	60.3
SWIR2.SWIR1	-	34.9	2.7	-	0.6	1.9	-	-	-	-
dSWIR2.SWIR1	12.2	41.6	16.3	-	-	9.4	8.1	-	-	40.0
NIR.G	7.6	35.2	40.7	-	28.2	32.0	14.6	**18.5**	-	34.8
dNIR.G	26.7	37.1	-	-	-	-	-	-	-	42.6
NIR.R	5.0	32.4	47.5	**7.1**	35.5	35.3	15.6	**29.6**	-	32.8
dNIR.R	2.2	-	-	-	-	-	-	-	-	27.4
TC.BRI	-	-	-	-	-	-	-	-	-	-
dTC.BRI	-	-	23.7	-	-	37.6	3.0	-	-	33.7
TC.GRE	2.9	27.8	28.7	-	28.5	19.9	12.0	10.0	-	18.8
dTC.GRE	10.5	33.9	41.5	-	22.6	22.4	13.1	**26.3**	-	**78.7**
TC.WET	-	39.6	27.1	**7.3**	7.8	31.3	9.1	**1.05**	(3)	17.9
dTC.WET	**27.9**	48.4	52.4	-	20.1	48.0	5.2	16.8	-	73.3
dB	-	44.4	49.5	-	29.2	**56.3**	19.7	-	-	67.7
dG	-	38.7	31.5	-	-	48.2	16.0	-	-	64.6
dR	3.1	45.1	54.9	-	40.9	**57.4**	20.6	4.2	(6)	**83.1**
dNIR	-	-	3.0	-	-	-	-	-	-	60.0
dSWIR1	-	34.3	48.7	-	18.1	48.0	21.8	-	(14)	63.5
dSWIR2	4.1	47.2	56.6	-	30.2	47.7	**33.0**	24.1	-	54.8

Table S3. Correlations between spectral indices and observed mortality of trees ≥10 cm diameter at breast height (1.37 m) based on a **rarefied dataset**. Observed mortality was quantified as percent of pre-fire live stems (or basal area [BA]) that was killed by fire. Values represent percent variance explained (%VE) determined with random forest models, averaged among all rarefied datasets. Values in parentheses indicate standard deviation of %VE among rarefied datasets. Column titles indicate which structural subsets were used to calculate observed mortality. Superscripts indicate significant differences between indices as determined with Tukey's HSD test. The columns for YFDP (Yosemite Forest Dynamics Plot) and JFSP (Joint Fire Science Program plots) represent percent BA mortality.

Index	Basal area		Stems ≥10 cm		Stems ≥50 cm		Stems ≥100 cm		YFDP		JFSP	
dNBR	52.4	(2.4)	46.5	(2.4)	38.3	(3.0)	5.6	(4.3)	14.6	(5.6)	34.7	(4.1)
dNDVI	50.7[a]	(2.3)	52.9[a]	(2.7)	28.7[a]	(3.6)	0[a]	(4.6)	10.8	(4.1)	32.3	(4.0)
RdNBR	56.9	(2.3)	42.6[b]	(2.7)	47.3[b]	(3.3)	18.8[b]	(4.5)	11.6	(5.6)	32.4	(4.3)
RBR	58.2	(1.8)	44.6	(2.4)	47.6[b]	(3.0)	18.4[b]	(4.4)	11.6	(5.7)	41.2	(2.9)
dSWIR1.NIR	59.8[b]	(2.0)	48.1	(2.3)	44.4[b]	(2.8)	8.2	(4.4)	19.0	(4.9)	36.4	(4.2)
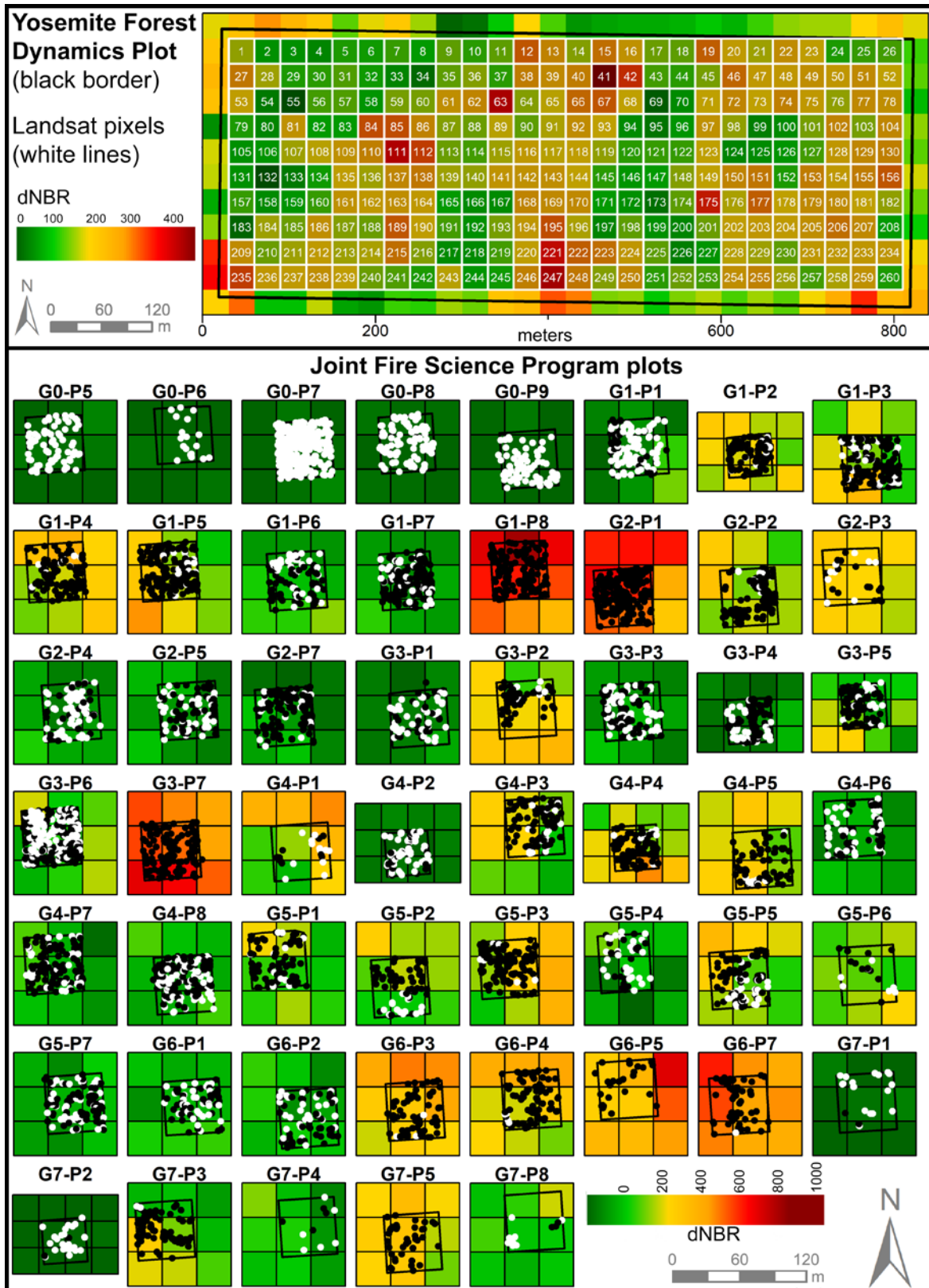
Fig. S1. Landsat pixels within the Yosemite Forest Dynamics Plot (numbers, top panel) and Joint Fire Science Plot network (Gx-Px titles, bottom panel). Points in the bottom panels represent surviving (white) and fire-killed (black) trees that were live pre-fire.
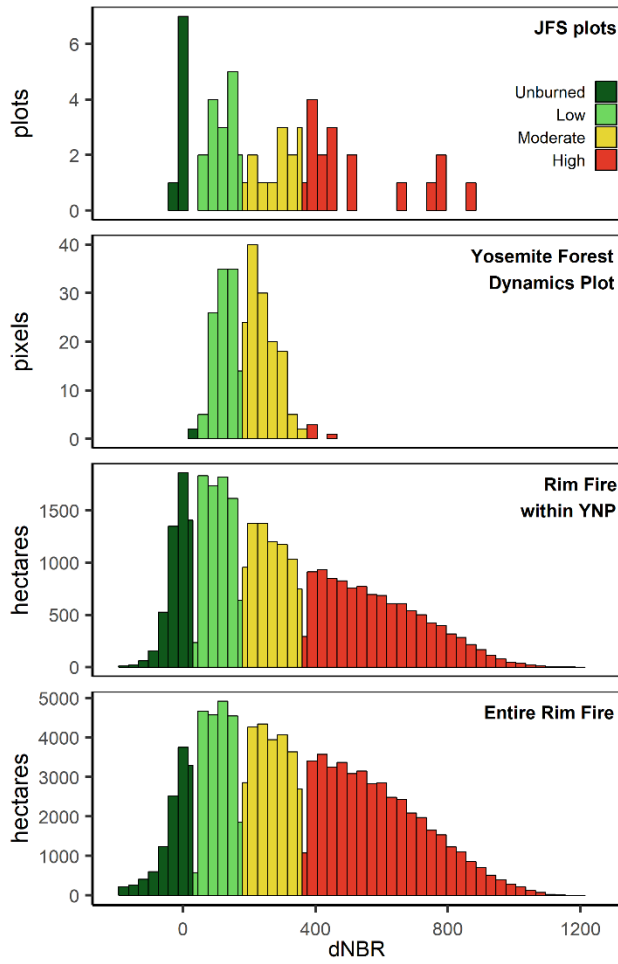
Fig. S2. Distribution of fire severity (differenced Normalized Burn Ratio [dNBR], calculation details in *Methods*) within the Joint Fire Science Program plots, Yosemite Forest Dynamics Plot, Rim Fire within Yosemite National Park, and within the entire Rim Fire.
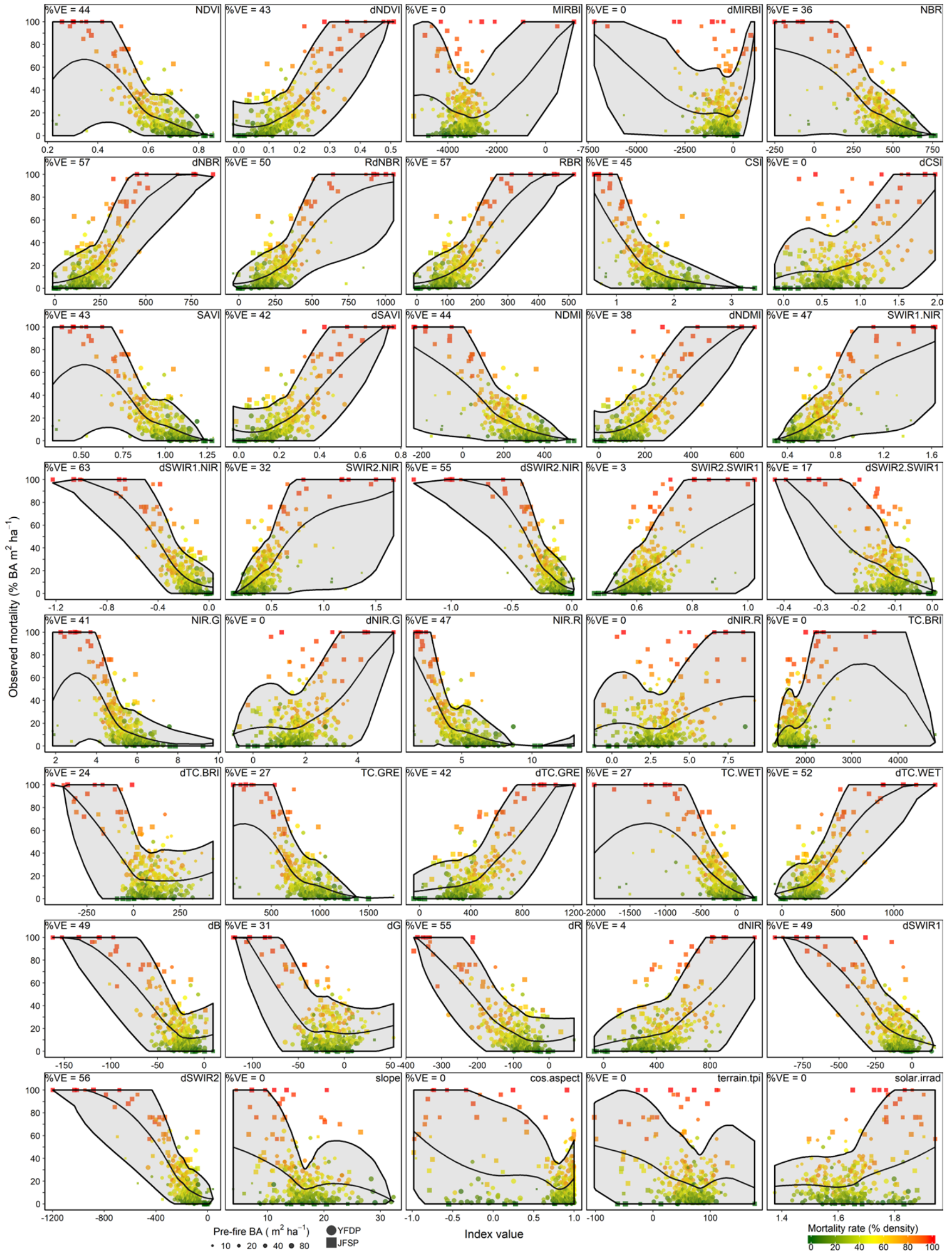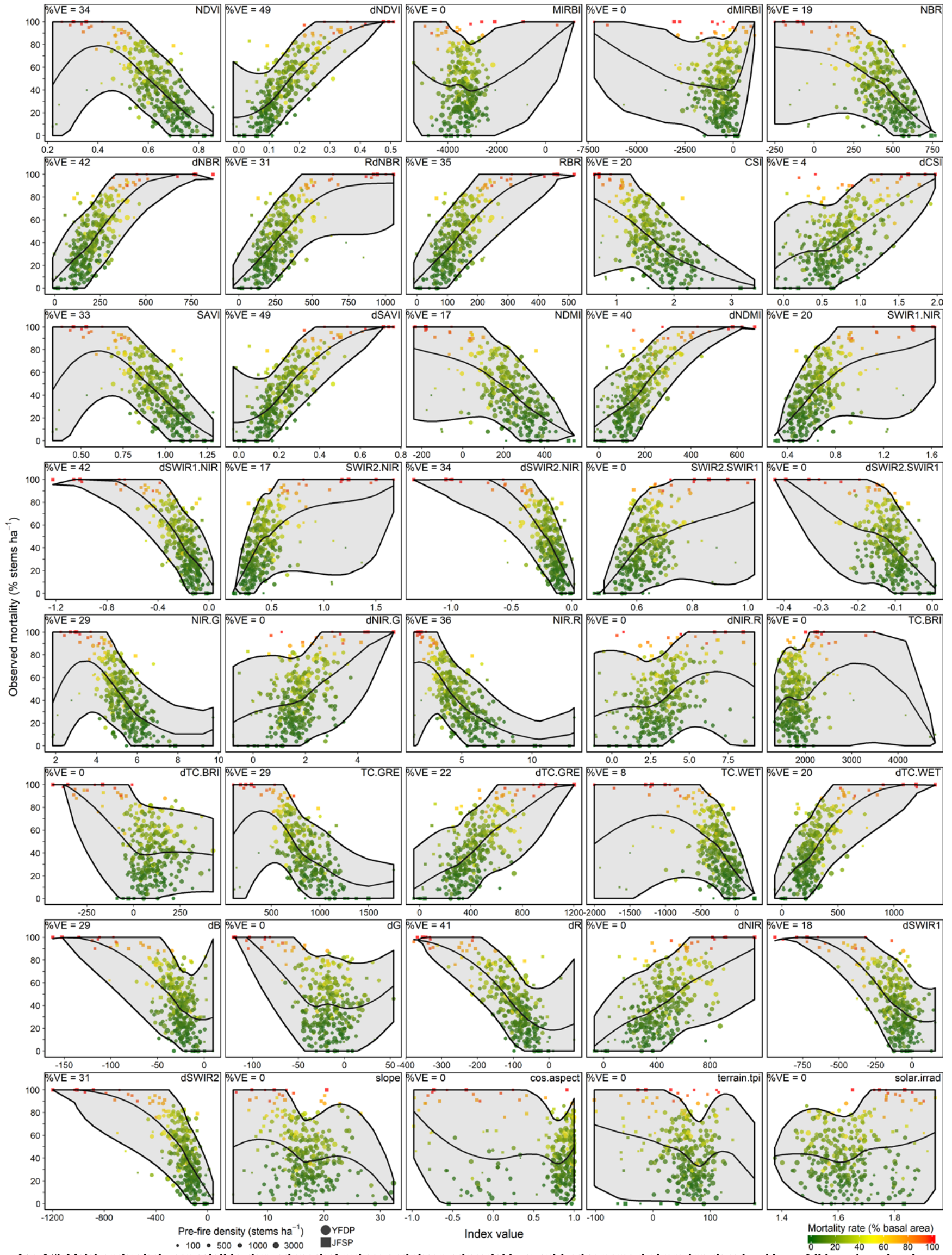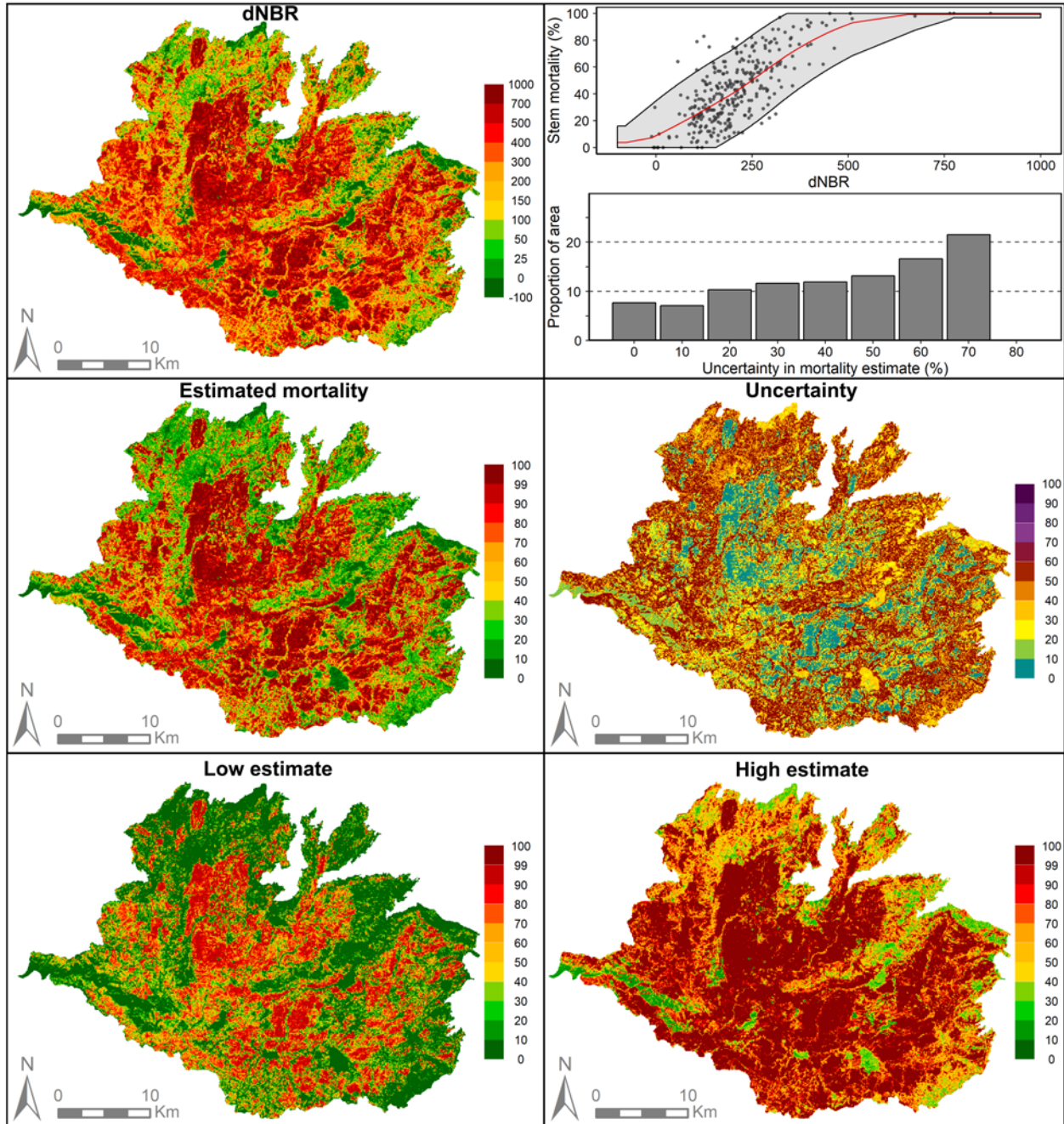
Fig. S3. Relationship between satellite-derived spectral indices and observed mortality quantified as percent of pre-fire basal area of trees >10 cm diameter at breast height (1.37 m) that was killed within one year of the fire. Points represent 53 individual plots (JFSP) and 260 Landsat pixels within in the Yosemite Forest Dynamics Plot (YFDP). Shaded envelopes represent the variability in observed mortality as a function of each spectral index. Envelopes were derived from continuous estimates of the mean and standard deviation of the data generated with locally-weighted scatterplot smoothing (LOESS) regression models and scaled to capture 95% of the variability in observations. The %VE indicates the percent variance explained using random forest models (Table 2, S2).
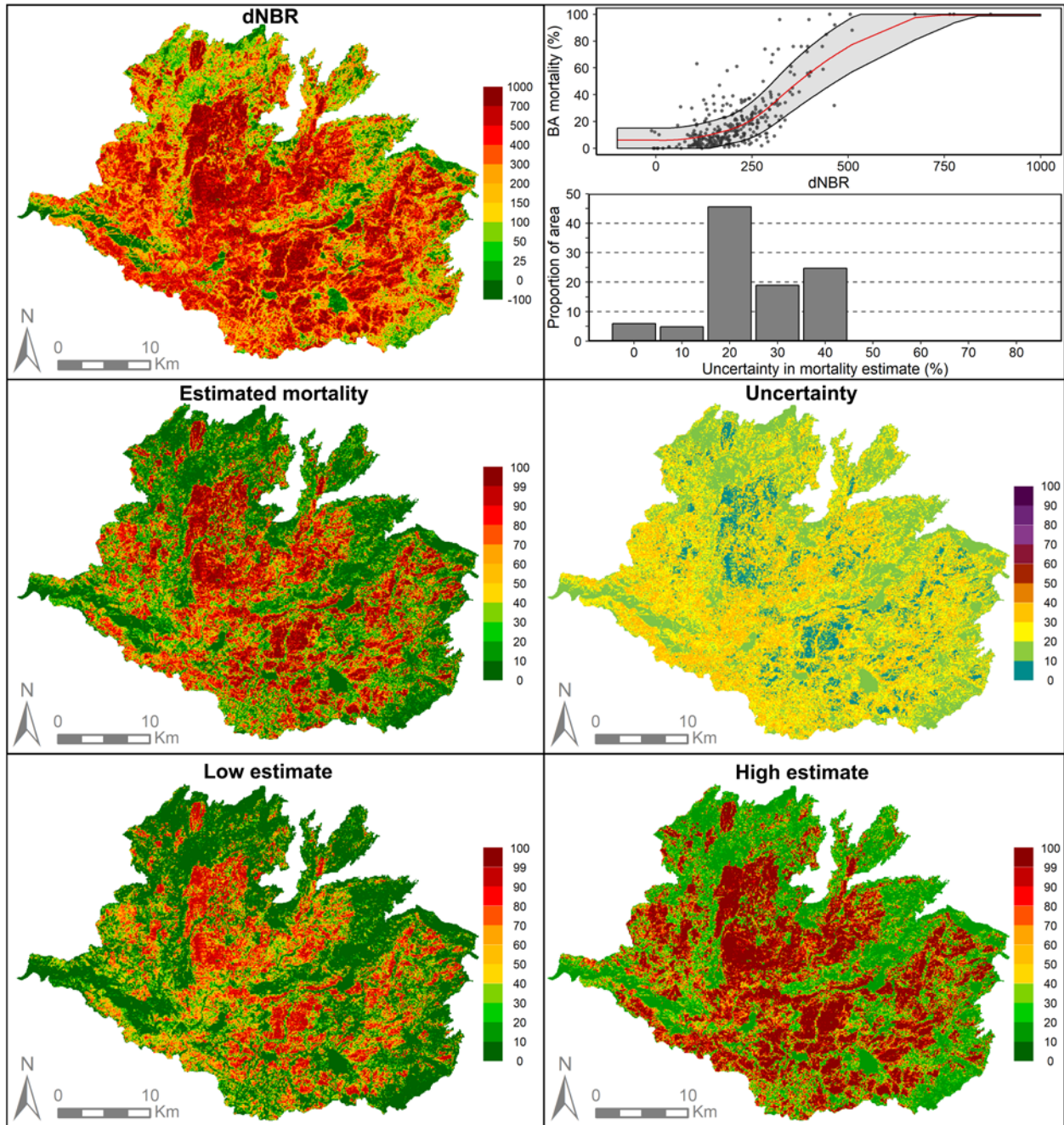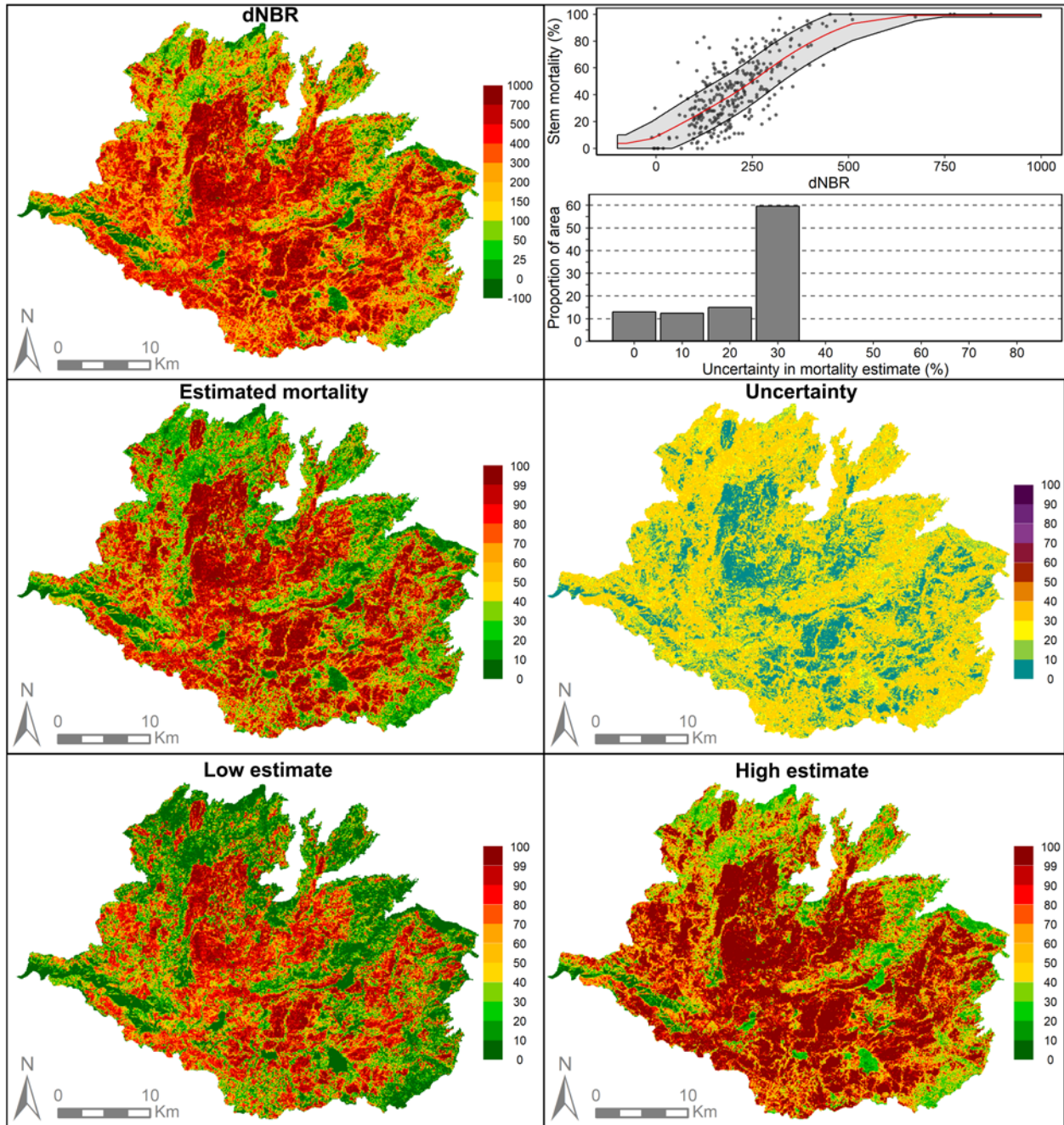
Fig. S4. Relationship between satellite-derived spectral indices and observed mortality quantified as percent of pre-fire density of trees ≥10 cm diameter at breast height (1.37 m) that was killed within one year of the fire. Points represent 53 individual plots (JFSP) and 260 Landsat pixels within the Yosemite Forest Dynamics Plot (YFDP). Shaded envelopes represent the variability in observed mortality as a function of each spectral index. Envelopes were derived from continuous estimates of the mean and standard deviation of the data generated with locally-weighted scatterplot smoothing (LOESS) regression models and scaled to capture 95% of the variability in observations. The %VE indicates the percent variance explained using random forest models (Table 2, S2).

Fig. S5. Satellite-derived burn severity (dNBR) of the California Rim Fire. The scatterplot displays the relationship between dNBR and observed percent mortality of trees ≥10 cm diameter at breast height (red line is predicted mortality rate of LOESS model; shaded region is 95% confidence envelope). The histogram shows proportion of area within the Rim Fire at various levels of uncertainty. Maps show mean, low, and high estimates of percent stem mortality. The uncertainty map displays the range in predicted percent stem mortality necessary to capture the true mortality rate, 95% of the time (i.e., a 40% uncertainty level indicates a ±20% range in expected mortality levels).

Fig. S6. Satellite-derived burn severity (dNBR) of the California Rim Fire. The scatterplot displays the relationship between dNBR and observed percent basal area mortality of trees ≥10 cm diameter at breast height (red line is predicted mortality rate of LOESS model; shaded region is 68% confidence envelope). The histogram shows proportion of area within the Rim Fire at various levels of uncertainty. Maps show mean, low, and high estimates of percent basal area mortality. The uncertainty map displays the range in predicted percent basal area mortality necessary to capture the true mortality rate, 68% of the time (i.e., a 20% uncertainty level indicates a 10% range in expected mortality levels).

Fig. S7. Satellite-derived burn severity (dNBR) of the California Rim Fire. The scatterplot displays the relationship between dNBR and observed percent stem mortality of trees ≥10 cm diameter at breast height (red line is predicted mortality rate of LOESS model; shaded region is 68% confidence envelope). The histogram shows proportion of area within the Rim Fire at various levels of uncertainty. Maps show mean, low, and high estimates of percent stem mortality. The uncertainty map displays the range in predicted percent stem mortality necessary to capture the true mortality rate, 68% of the time (i.e., a 20% uncertainty level indicates a 10% range in expected mortality levels).
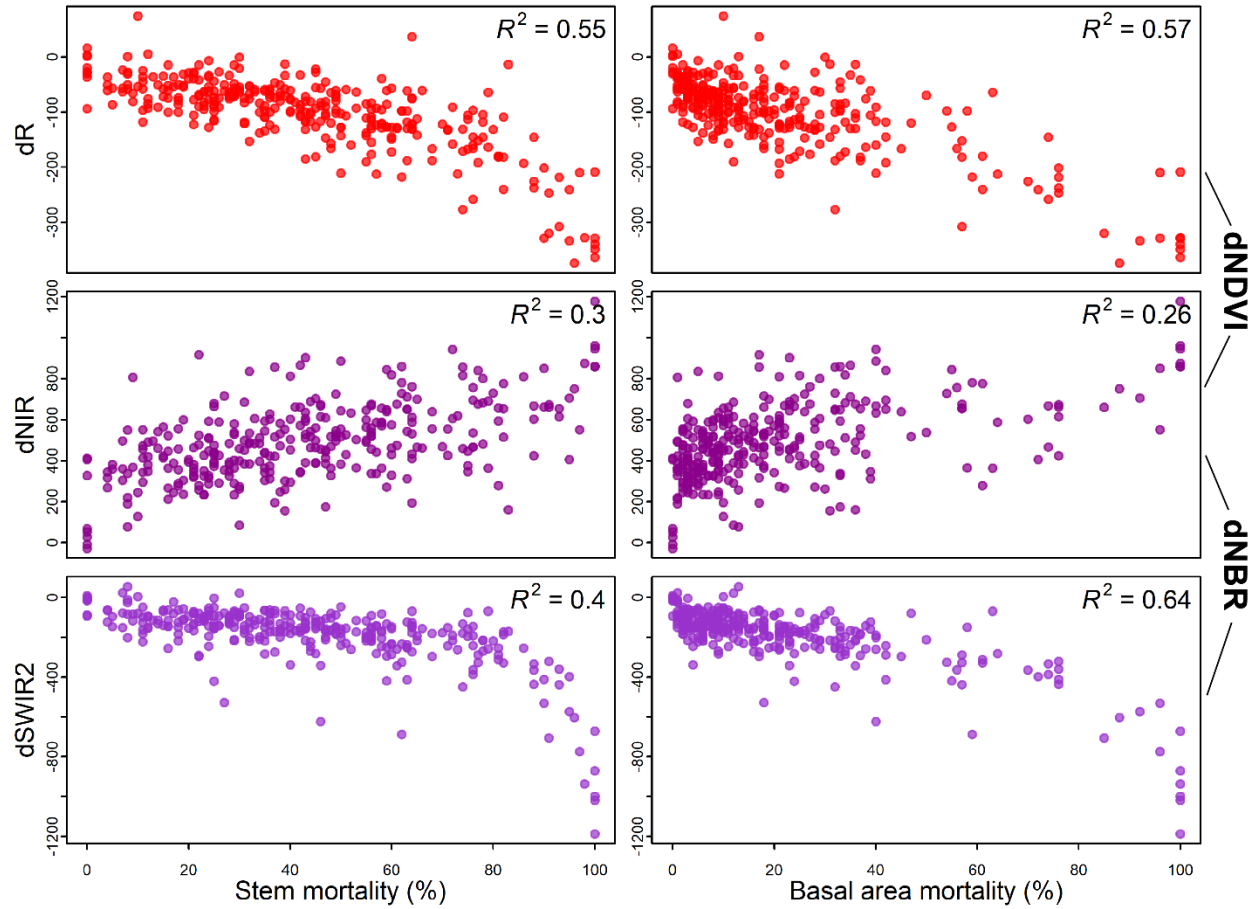
32   Fig. S8. Scatterplots of differenced band reflectance values versus mortality of trees ≥10 cm DBH. Each row
33 represents a different band (red band is the top row, near-infrared is the middle row, and short-wave infra-red is the
34 bottom row). The left column contains observed mortality measured as percent mortality of stems ≥10 cm DBH
35 while the right column contains observed mortality measured as percent of pre-fire basal area.
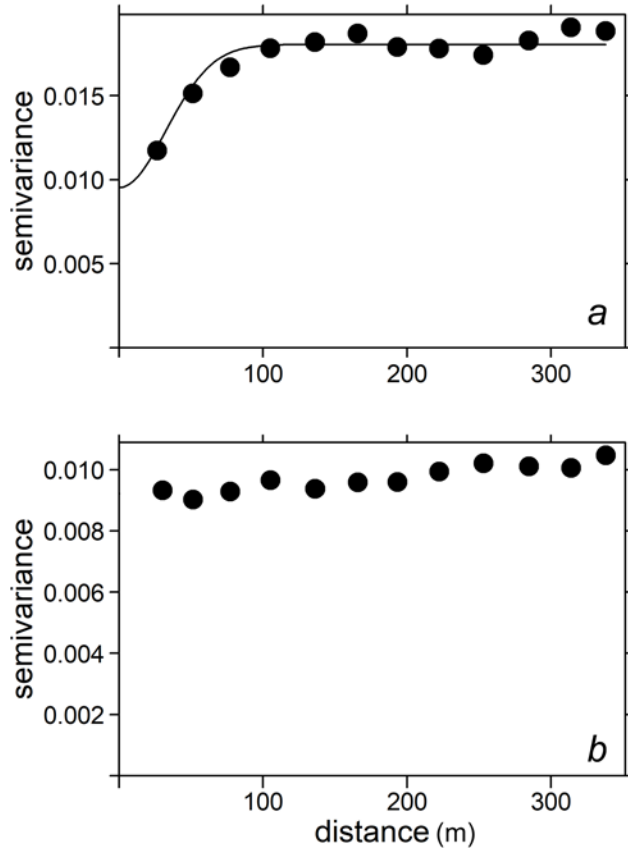
Fig. S9. Semivariograms depicting spatial autocorrelation of basal area mortality (*a*) and residuals (predicted minus observed) of the LOESS model (details in *Methods*) for dNBR (*b*). The top panel reveals that mortlaity was spatially auto-correlated at spatial scales <100 m, while error in Landsat-derived severity indices was spatially random (i.e., the difference between predicted and observed mortality for any given pixel was independent of surrounding pixels).