Graduate Student Theses, Dissertations, & Professional Papers

Graduate School

2019

NEW, MULTI-SCALE APPROACHES TO CHARACTERIZE PATTERNS IN VEGETATION, FUELS, AND WILDFIRE

Christopher Jacob Moran

Let us know how access to this document benefits you.

Follow this and additional works at: https://scholarworks.umt.edu/etd

NEW, MULTI-SCALE APPROACHES TO CHARACTERIZE PATTERNS IN

VEGETATION, FUELS, AND WILDFIRE

By

CHRISTOPHER JACOB MORAN

B.S. Biology, South Dakota State University, Brookings, SD, 2009
M.S. Biological Sciences, South Dakota State University, Brookings, SD, 2011

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in Forestry and Conservation Sciences
The University of Montana
Missoula, MT

Official Graduation Date December 2019

Approved by:

Scott Whittenburg,
Graduate School Dean

Dr. Carl Seielstad, Chair
Department of Forest Management

Dr. LLoyd Queen
Department of Forest Management

Dr. David Affleck
Department of Forest Management

Dr. Russell Parsons
USDA Forest Service Rocky Mountain Research Station

Dr. Dave Opitz
Department of Computer Science

Moran, Christopher, PhD, December 2019   Forestry and Conservation Sciences

New, multi-scale approaches to characterize patterns in vegetation, fuels, and wildfire

Chairperson: Carl Seielstad

   Pattern and scale are key to understanding ecological processes. My dissertation research aims for novel quantification of vegetation, fuel, and wildfire patterns at multiple scales and to leverage these data for insights into fire processes. Core to this motivation is the 3-dimensional (3-D) characterization of forest properties from light detection and ranging (LiDAR) and structure-from-motion (SfM) photogrammetry. Analytical methods for extracting useable information currently lag the ability to collect such 3-D data. The chapters that follow focus on this limitation blending interests in machine learning and data science, remote sensing, wildland fuels (vegetation), and wildfire. In Chapter 2, forest canopy structure is characterized from multiple landscapes using LiDAR data and a novel data-driven framework to identify and compare structural classes. Motivations for this chapter include the desire to systematically assess forest structure from landscape to global scales and increase the utility of data collected by government agencies for landscape restoration planning. Chapter 3 endeavors to link 3-D canopy fuels attributes to conventional optical remote sensing data with the goal of extending the reach of laser measurements to the entire western US while exploring geographic differences in LiDAR-Landsat relationships. Development of predictive models and resulting datasets increase accuracy and spatial variation over currently used canopy fuel datasets. Chapters 4 and 5 characterize fire and fuel variability using unmanned aerial systems (UAS) and quantify trends in the influence of fuel patterns on fire processes.

Acknowledgements

Table of Contents

Introduction

Pattern and scale are key to understanding ecological processes (Levin 1992). My dissertation research aims for novel quantification of vegetation, fuel, and wildfire patterns at multiple scales and to leverage these data for insights into fire processes. Core to this motivation is the 3-dimensional (3-D) characterization of forest properties from light detection and ranging (LiDAR) and structure-from-motion (SfM) photogrammetry. Analytical methods for extracting useable information currently lag the ability to collect such 3-D data. The dissertation focuses on this limitation, blending interests in machine learning and data science, remote sensing, wildland fuels (vegetation), and wildfire. In Chapter 2, forest canopy structure is characterized from multiple landscapes using LiDAR data and a novel data-driven framework to identify and compare structural classes. Motivations for this chapter include the desire to systematically assess forest structure from landscape to global scales and increase the utility of data collected by government agencies for landscape restoration planning. Chapter 3 endeavors to link 3-D canopy fuels attributes to conventional optical remote sensing data with the goal of extending the reach of laser measurements to the entire western US while exploring geographic differences in LiDAR-Landsat relationships. Development of predictive models and resulting datasets increase accuracy and spatial variation over currently used canopy fuel datasets. Chapters 4 and 5 characterize fire and fuel variability using unmanned aerial systems (UAS) and quantify trends in the influence of fuel patterns on fire processes.

The common threads relating each chapter are the collection and analysis of 3-D point cloud representations of vegetation using new technology and novel analytical methods with the final chapter relating the representations to wildfire processes. At the time of this writing, Chapter 2 has been published in the journal *Remote Sensing of Environment* (Moran et al. 2018). Chapter 4 has been published in the journal *Fire* (Moran et al. 2019). Chapter 3 will be submitted for peer-review, and Chapter 5 is complete but needs collection of additional data across a broader range of environmental conditions before is submission for publication is warranted. For this reason, Chapter 5 can be thought of as ongoing and future research.

For the remainder of this introductory chapter, abstracts for each chapter are presented. The following chapters are each stand-alone in terms of having introduction, methods, results, discussion, conclusions, and literature cited sections.

Chapter 2: A data-driven framework to identify and compare forest structure classes using LiDAR

As LiDAR datasets increase in availability and spatial extent, demand is growing for analytical frameworks that allow for robust comparison and interpretation among ecosystems. We utilize data-driven classification in a hierarchical design to estimate forest structure classes with parsimony, flexibility, and consistency as priorities. We use an *a priori* selection of six input features derived from small-footprint (32 cm), high density (17 returns/m^2) airborne LiDAR: four

L-moments to describe the vertical distribution of canopy structure, canopy density as a measure of vegetation coverage, and standard deviation of canopy density to characterize within-cell horizontal variability. We identify 14 statistically-separated meta-classes characterizing six ecoregions over 168,117 ha in Montana, USA. Meta-classes follow four general vertical shapes: tall and continuous, short-single strata, tall-single strata, and broken strata over short strata. Structure classes that dominate locally but are rare overall are also identified. The approach outlined here allows for intuitive comparison and assessment of forest structure from any number of landscapes and forest types without need for field training data.

Chapter 3: Mapping forest canopy fuels over the western US with LiDAR-Landsat relationships

Complete coverage of forest canopy fuels spatial data over the western US is necessary for pre-fire, during-fire, and post-fire decision-making in land management. LiDAR provides accurate, comprehensive characterization of canopy fuels but lacks complete spatial coverage. Landsat satellite imagery provides complete coverage but lacks the data for direct canopy fuel characterizations. Here, a collection of LiDAR datasets are fused with Landsat-derived spectral indices to map the canopy fuels necessary for wildfire predictions: canopy cover (CC), canopy height (CH), canopy base height (CBH), and canopy bulk density (CBD). Spatial variance in the LiDAR-Landsat relationships is identified. The spectral index response shapes are largely consistent among landscapes but shifted in mean response values. A single, gradient boosting machine (GBM) model using data from all datasets is able to characterize these relationships with only small reductions in model performance (mean 0.04 reduction in R^2) compared to local GBM models trained on individual LiDAR datasets. Model evaluation on completely independent LiDAR datasets shows the global model outperforming local models (mean 0.24 increase in R^2) indicating improved model generality. The global GBM model drastically improves performance over existing LANDFIRE canopy fuels data while testing on independent datasets (R^2 ranging from 0.15 to 0.61 vs. -3.94 to -0.374). The ability to automatically update canopy fuels following wildfire disturbance is also tested, and results show intuitive reductions in fuels for high and moderate fire severity classes and little to no change for unburned to low fire severity classes.

Chapter 4: Deriving fire behavior metrics from UAS imagery

The emergence of affordable unmanned aerial systems (UAS) creates new opportunities to study fire behavior and ecosystem pattern—process relationships. A rotor-wing UAS hovering above a fire provides a static, scalable sensing platform that can characterize terrain, vegetation, and fire coincidently. Here, we present methods for collecting consistent time-series of fire rate of spread (RoS) and direction in complex fire behavior using UAS-borne NIR and Thermal IR cameras. We also develop a technique to determine appropriate analytical units to improve statistical analysis of fire-environment interactions. Using a hybrid temperature-gradient threshold approach with data from two prescribed fires in dry conifer forests, the methods characterize complex

interactions of observed heading, flanking, and backing fires accurately. RoS ranged from 0–2.7 m/s. RoS distributions were all heavy-tailed and positively-skewed with area-weighted mean spread rates of 0.013-0.404 m/s. Predictably, the RoS was highest along the primary vectors of fire travel (heading fire) and lower along the flanks. Mean spread direction did not necessarily follow the predominant head fire direction. Spatial aggregation of RoS produced analytical units that averaged 3.1–35.4% of the original pixel count, highlighting the large amount of replicated data and the strong influence of spread rate on unit size.

Chapter 5: Deriving 3-D fuel variability from UAS imagery and characterizing trends in fuel pattern-fire Process spatial associations

Characterizing fuel pattern and fire process relationships remains a research priority but is hampered by a lack of datasets with coincident fuel and fire measurements. Here, unmanned aerial systems (UAS) are used to characterize 3-D fuel variability using structure-from-motion at moderate extents (100 m^2). Using 196-340 images collected per 100 m^2 plot and with custom UAS flight patterns and manual identification of reference tiepoints and scale constraints, surface fuels are characterized in exquisite detail (mean 235,816,443 3-D points/plot). Conversion of point cloud data to gradient metrics produces a direct measurement of fuel variability without supporting field data. In conjunction with coincident fire rate of spread (RoS) observations, trends in spatial associations between fuel and fire patterns are quantified. A bivariate measure of spatial association, the L-index, shows decreasing spatial co-patterning between fuel variability and fire RoS as RoS increases. Collection of additional fire and fuel data in different fuel types and RoS ranges will further elucidate trends in fuel and fire spatial associations.

Literature Cited

Levin, S.A.. 1992. The problem of pattern and scale in ecology. *Ecology* 73: 1943-1967.

Moran, C.J., E. Rowell, and C.A. Seielstad. 2018. A data-driven framework to identify and compare forest structure classes using LiDAR. *Remote Sensing of Environment* 211: 154-166.

Moran, C.J., C.A. Seielstad, M.R. Cunningham, V. Hoff, R.A. Parsons, L. Queen, K. Sauerbrey, and T. Wallace. 2019. Deriving fire behavior metrics from UAS imagery. *Fire* 2: 36.

Chapter 2. A data-driven framework to identify and compare forest structure classes using LiDAR

1) Introduction

Forest structure is both a driver and product of ecosystem processes (Spies 1998, Shugart et al. 2010). A variety of connections exist between structure and ecosystem traits including biodiversity, habitat, previous and future disturbance, successional trajectories, water interception, gas exchange, carbon storage, and productivity (Ellsworth and Reich 1993, Spies 1998, Franklin et al. 2002, Parker et al. 2004, Pregitzer and Euskirchen 2004, Bergen et al. 2009, Culbert et al. 2013, Johnstone et al. 2016). Characterizing forest structure and its variation remains a priority for research and land management engaged in conservation, restoration, and the ecological sciences.

Light detection and ranging (LiDAR) and related analyses have been used to quantify and classify forest structure for a variety of applications (Lim et al. 2003, Vierling et al. 2008, Kane et al. 2010a, Miura and Jones 2010, Smart et al. 2012, Simonson et al. 2014, Listopad et al. 2015). LiDAR can characterize the three-dimensional arrangement of the overstory canopy, which correlates to biomass and other structural metrics (Lefsky et al. 2002). In its raw form, a modern LiDAR point cloud contains an abundance of height measurements that are often summarized on raster grids to reduce data volumes and facilitate development of predictive models (Yu et al. 2010, Wulder et al. 2013). These so-called area-based approaches (ABAs) effectively predict forest attributes and classify forest structure (e.g., Lefsky et al. 1999, Næsset 2002, Zimble et al. 2003, Frazer et al. 2005, Lefsky et al. 2005, Coops et al. 2007, Falkowski et al. 2009, Leiterer et al. 2015).

While applicable for a particular study area, ABA techniques tend to produce features and models with unique properties, which often have low generality (Lefsky et al. 2005, Bouvier et al. 2015). Differing LiDAR sensor configurations and acquisition parameters, vegetation types, structure attributes, and field collection methods contribute to site-specific results. In addition, the widespread use of arbitrary canopy height strata and percentiles provides no consistency, thereby confusing potential comparisons, and may reduce the ability to accurately characterize structural attributes (Chen 2013, Gorgens et al. 2017). Field-sampling usually provides the training and test data for classification but adds considerable expense and rarely captures the range of variability over large spatial extents (Hawbaker et al. 2009, Maltamo et al. 2011).

The latter limitation in particular has led to development of data-driven approaches utilizing unsupervised or semi-supervised classification methods (Kane et al. 2010a, Jones et al. 2012, Leiterer et al. 2015, Vauhkonen and Imponen 2016). The defining characteristic distinguishing a data-driven approach from a standard ABA is that classes are not predefined and thus depend on the characteristics of the input datasets (e.g., Zhang et al. 2011, Dupey et al. 2013, Kane et al. 2013, Dickinson et al. 2014). Derived classes are centered on statistical groupings (Halkidi et al. 2001) and may not match existing forest structure classifications. Depending on the

features used in a classification, interpretation of identified classes may also suffer from many of the same aforementioned ABA limitations (e.g., Leiterer et al. 2015).

Much of the literature focusing on data-driven approaches is concerned with selection of variables that best characterize and differentiate forest structure types. Kane et al. (2010b) identified a subset of LiDAR variables related to both field measurements and forest structure complexity including 95th percentile height, mean height, height variance, canopy density, and rumple. Similarly, Jones et al. (2012) showed that certain structure classes (e.g., young forest versus mature forest) could only be discriminated using specific metrics — in this case the ordinary statistical moment kurtosis. The most frequently cited variables in the literature consistently fall into the general categories of forest height, height variability, and canopy cover, corresponding to the classes noted by Lefsky et al. (2005). More than a decade later, the literature has repeatedly revealed the utility of variables in these categories for characterizing a variety of related forest structure classifications based on forest age (Jones et al. 2012), complexity (Kane et al. 2010a, b), number of strata (Whitehurst et al. 2013), tree size (Kane et al. 2013, North et al. 2017), successional stage (Falkowski et al. 2009), and forest type (Zhang et al. 2011) among others (Dupey et al. 2013, Dickinson et al. 2014, Niemi and Vauhkonen 2016). The diversity of variables combined with the diversity of classification schemes adds complexity to forest characterization but still provides strong underpinnings for development of more broadly applicable approaches.

As LiDAR datasets increase in availability and spatial extent, demand is growing for unifying analytical frameworks that allow for comparison and interpretation among and between landscapes with and without supporting field data. At a minimum, such approaches could support optimization of forest surveys by systematically describing structure variability, guiding field data collection, and determining optimal plot dimensions (Frazer et al. 2011). A major challenge is the selection of a small set of LiDAR variables that not only can discriminate relevant structure classes but can be interpreted by forestry professionals without supporting field data (Kane et al. 2010b). A second challenge is development of classification methods to facilitate natural groupings of structure attributes and common interpretations of them across landscapes. These challenges provide the basis for our research, which focuses on methods to: (1) discriminate natural groupings of forest structure (classes) within landscapes using a few, interpretable LiDAR metrics without the need for field training data, and (2) aggregate structure classes across landscapes using a consistent set of features. Our approach follows Frazier et al. (2011) who suggest partitioning LiDAR into a few unique statistical classes each with relatively homogenous properties. It addresses the sensitivity of statistical classes to area and data parameters noted by Jones et al. (2012) by aggregating multiple sub-classifications of individual landscapes to create 'meta-classes'. Leiterer et al. (2015) suggests this type of spatially stratified classification to maintain localized distinctness when comparing diverse forest types.

We use machine learning to produce structure classes. Specifically, a combination of Random Forests (RF, Breiman 2001) to estimate dissimilarity and predict classes and hierarchical

clustering to group based on dissimilarity (Murtagh and Legendre 2014). Hierarchical clustering has successfully grouped forest structure into ecologically-relevant classes based on statistical distinctiveness (Latham et al. 1998, Kane et al. 2010a, Kane et al. 2013). While machine learning techniques excel at identifying complex, non-linear feature relationships (Lawrence and Moran 2015), their inherent nature and the propensity to use a large number of input features contribute to 'black-box' classification. *A priori* feature selection, with priority given to interpretable metrics and low dimensionality, allows us to exploit the power of machine learning and minimize the black-box effect.

The L-moments provide the basis of our *a priori* feature selection to characterize the vertical domain of forest structure. L-moments have strong statistical underpinnings and provide a low-dimensional solution to the complex problem of distribution characterization (Hosking 1990). Multiple studies have utilized L-moments for characterizing canopy structure with LiDAR data (Frazer et al. 2011, Ozdemir and Donoghue 2013, Valbuena et al. 2017). They are order statistics and can be used to calculate features analogous to standard deviation, skewness, and kurtosis (i.e. basic descriptors of theoretical distributions). Being linearly combined, they are less affected by outliers and variation in sample sizes than standard product moments (Hosking 1992). Furthermore, the L-moment ratios have finite theoretical bounds allowing for comparisons of shape with different location and scale (Hosking 1990). Valbuena et al. (2017) utilized two L-moment ratios describing LiDAR distributions, the L-coefficient of variation (L-CV) and L-skewness, to classify key structural features of boreal forest canopies without having to statistically link field data to LiDAR metrics. Similarly, L-CV, L-skewness, and L-kurtosis explained unique and significant structure variability in simulated forest stands (Frazer et al. 2011). Hosking and Wallis (1997) provide an in-depth treatment of L-moments and their formulation.

Alone, the L-moments are not sufficient to fully characterize forest structure because, like many LiDAR point cloud derivatives, they do not account explicitly for the abundance and horizontal distribution of canopy structure within individual cells (Popescu and Zhao 2008, Zhao et al. 2009, Bouvier et al. 2015, Leiterer et al. 2015). Canopy density (defined here as the number of first returns above 2 meters height divided by all first returns) provides a useful conception of the amount of vegetation coverage within a cell (Lefsky et al. 2002, Maltamo et al. 2016), but lacks information on the within-cell, horizontal variability of canopy material. Sub-cell metrics that characterize variability in canopy density or canopy density within different height strata have been used to describe the horizontal distribution of vegetation and as predictors of related field metrics (Lim and Treitz 2004, Hudak et al. 2006).

The primary objective of our work is to develop a consistent, interpretable, and flexible framework to identify and compare predominant forest canopy structures across diverse landscapes without the need for field training data. We rely on *a priori* selection of input features, using the four statistical L-moments, canopy density, and a sub-grid metric called horizontal standard deviation of canopy density (HSD of CD, defined in section 2.3). We take an

unsupervised classification approach to estimate forest structure classes. RF identifies natural groupings within LiDAR datasets, hierarchical clustering groups based on estimated dissimilarity, a second iteration of RF classifies landscapes using cluster labels, and equivalence testing aggregates landscape-specific classes to meta-classes. We use 168,117 ha of high-density, small-footprint LiDAR data spanning six ecoregions in the Northern Rocky Mountains, USA to develop and test our methods.

2) Methods
2.1) Study Area

The study area is in the southwest portion of the four million ha Crown of the Continent Ecosystem (CCE) in the Northern Rocky Mountains, USA. In 2014-2015, the US Forest Service acquired airborne laser scanning (ALS) data for the southwestern portion of the CCE. In 2015, the University of Montana added two additional acquisitions on its properties in the CCE. The extent of these datasets forms the boundary of our study area. The study area was divided into 17 distinct landscapes delineated by LiDAR acquisition dates, ecoregion, and total area (Fig. 1, Table 1). The landscapes cover 168,117 ha in a range of ecosystems typical of the Northern Rocky Mountains from dry, low elevation *Pinus ponderosa* forests to higher elevation, mixed-conifer forest types. Six Omernik Level IV ecoregions are present in the study area (Fig.1; Omernik and Griffith 2014). Elevations range from 1038-2544 m, annual precipitation from 405-1479 mm, and mean annual temperature from 1.3-6.4º C (Table 1; PRISM 2012). *Pseudotsuga menziesii* is the most common tree, coexisting with abundant *Larix occidentalis*, *Pinus contorta*, *Pinus ponderosa*, *Abies lasiocarpa*, and *Picea engelmannii* (Barber et al. 2012). Stands of *Pinus albicaulis* are also present at a lower abundance.

The study area has been subjected to extensive timber harvesting and fire exclusion practices. Fire cessation in the 20th century pushed portions of these ecoregions outside historical norms with consequent changes in forest structure (Schoennagel et al. 2004). Other land-uses such as logging, grazing, agriculture, and human habitation have had strong impacts on forest structure, particularly in the Stillwater-Swan Wooded Valley, depicted in landscapes 1-5 where a checkerboard of different land ownership and use is conspicuous.
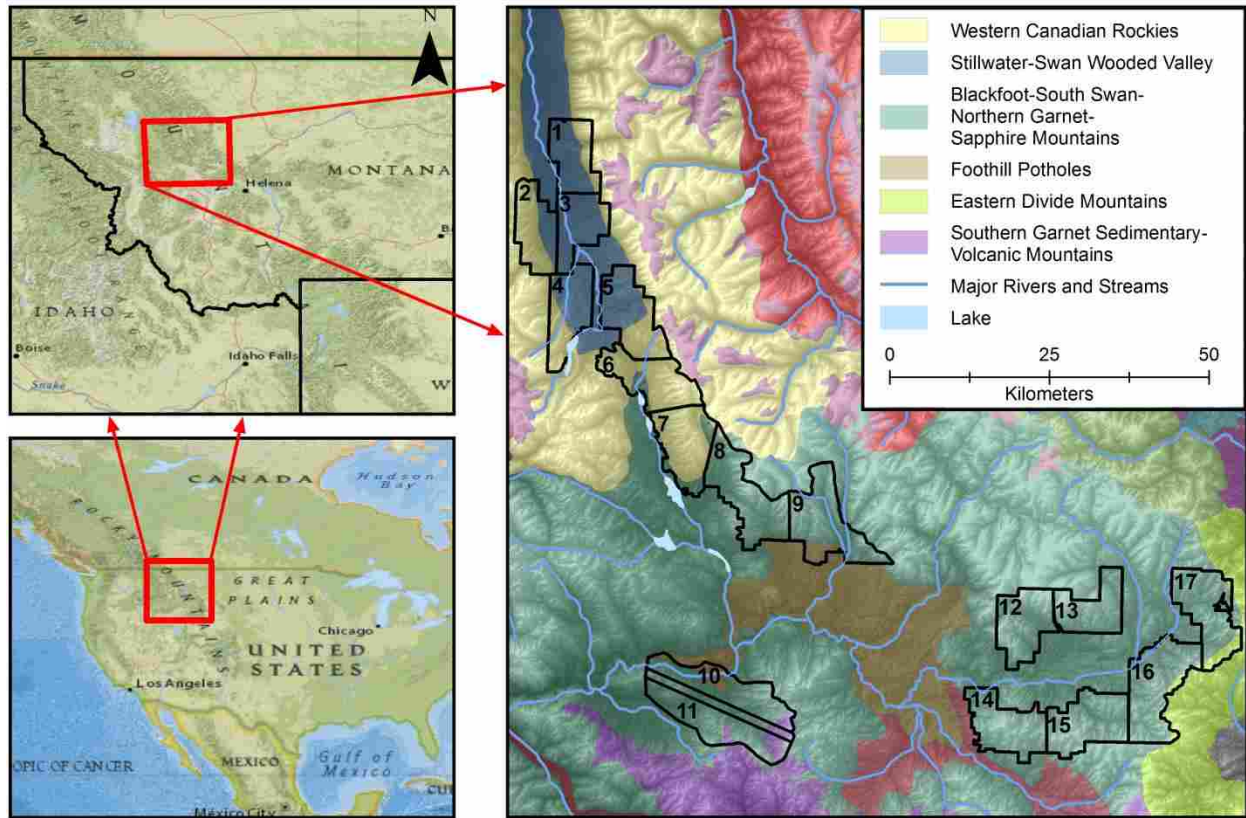
Figure 1. Map of Omernik Level IV ecoregions and delineated landscapes (numbered) comprising the study area (Omernik and Griffith 2014). Ecoregions shown in the legend are only those that are within the study area.

Table 1. Description of landscapes in this study. Ranges and means are spatially derived from 1981-2010 climatic normals (PRISM 2012).

#	Name	Area (ha)	Elevation (m) Min-Max	Mean	Mean Annual Temperature (^{0}C) Min-Max	Mean	Mean Annual Precipitation (mm) Min-Max	Mean
1	North Meadow	7458	1038-2305	1226	2.6-6.4	5.8	528-1469	679
2	North Hemlock	7774	1089-1965	1364	2.9-6.1	5.3	516-1269	796
3	South Meadow	7332	1068-1993	1156	4.4-6.4	6.0	495-1206	558
4	South Hemlock	8546	1155-1984	1375	2.5-5.8	5.2	526-1359	763
5	Buck	9956	1170-	1347	3.9-5.4	5.0	539-1167	670

			1949					
6	Colt	9296	1237-2449	1552	1.3-5.1	4.3	613-1479	888
7	Rice	9661	1218-2103	1522	2.8-5.1	4.5	485-1133	672
8	Lower Trail	11531	1283-2450	1613	1.3-5.1	4.2	466-1309	750
9	Monture	12006	1238-2200	1609	2.9-4.9	4.2	418-1078	645
10	North Lubrecht	13105	1080-2149	1426	2.6-6.1	5.2	411-742	507
11	South Lubrecht	13386	1157-2145	1525	2.9-6.1	5.1	405-747	552
12	West Stonewall	8654	1334-2544	1798	1.3-4.8	3.7	452-1277	766
13	East Stonewall	7523	1451-2520	1837	1.3-4.6	3.4	520-1244	761
14	West Dalton	9306	1325-2258	1770	2.3-4.8	4.1	423-811	608
15	East Dalton	10016	1403-2330	1793	2.1-4.8	3.8	477-768	627
16	Stemple	12226	1433-2312	1806	2.1-4.5	3.6	487-782	607
17	Blackfoot	10341	1536-2203	1773	2.8-4.3	3.6	496-690	570
	All:	*168117*	*1038-2544*	*1558*	*1.3-6.4*	*4.5*	*405-1479*	*672*

2.2) LiDAR Data

Quantum Spatial (https://quantumspatial.com) conducted three discrete return, small-footprint LiDAR collection campaigns in 2014 and 2015 (Table 2). Acquisition parameters were designed for optimal vegetation characterization: low flight altitude, high pulse rate, maximum angle of 15° off nadir, multiple returns recorded per pulse, 50% side-lap (100% overlap) of flightlines, and opposing flight line directions. These parameters intended to reduce laser shadowing, reduce edge-of-scan error, and prevent data gaps. Fundamental vertical accuracy averaged 0.06 m (n=963) for all acquisitions. First-return point density averaged 17.74 points/m^2 for the 2014 acquisitions and 19.82 points/m^2 for 2015 acquisitions.

Sensor and acquisition parameters were nearly identical between campaigns (Table 2). However, acquisition dates were in the fall for 2014 and early summer in 2015. Major differences between acquisitions due to phenology in the canopy were unlikely as evergreen, coniferous trees dominate, and the foliage of deciduous conifer *Larix occidentalis* was still attached in all acquisitions. We did not include any points below two meters above ground (except in canopy density calculations) to focus on forest canopy structure and to reduce any differences due to acquisition dates in the sub-canopy. Gorgens et al. (2017) found exclusion of points close to the ground strengthened relationships between most LiDAR metrics and stand volume estimates.

Table 2. Description of LiDAR parameters.

Acquisition Parameters	1st Campaign	2nd Campaign
Dates	Sep 24, 14 - Nov 1, 14	Jun 22, 15 - Jun 30, 15
Landscapes Scanned	1-9, 14-17	10-13, ~6% of 9
Laser Sensor	Leica ALS70	Leica ALS70
Altitude	1450m	1400-1450m
Laser Pulse Rate	180-193 kHz	180-193 kHz
Field of View	30º	30º
Laser Pulse Diameter	32 cm	32 cm
Mean First Return Pulse Density	17.72 points/m^2	19.82 points/m^2
Maximum Returns	Up to 4/pulse	Up to 5/pulse
Fundamental Vertical Accuracy	0.06 m	0.07 m

LiDAR processing to attribute point location, filter erroneous points, perform calibration corrections, and classify ground points with error estimations was completed by Quantum Spatial using a proprietary workflow including TerraScan, TerraMatch, and TerraModeler (versions 14 and 15). A digital terrain model (DTM) was derived from a triangulated irregular network using the mean elevation at a 1 m scale. Each LiDAR point's orthometric elevation was differenced from the underlying DTM to normalize to height above ground.

2.3) Feature Selection

We utilized FUSION software (McGaughey 2015) to calculate the four L-moment features and canopy density (CD, Table 3). Horizontal standard deviation of canopy density (HSD of CD) was calculated with R software (raster package; Hijmans 2016). All features were calculated at a 20 m cell resolution. Cell size was informed by Leiterer et al.'s (2015) findings on the effects of spatial resolution on information loss in data-driven canopy structure classification. CD is defined as the percentage of first laser returns above 2 m. HSD of CD was derived by first splitting the original 20 x 20 m cell into 400 sub-cells (1 m^2 each) and calculating the CD for each sub-cell. The standard deviation of all the sub-cell CDs was then attributed to the original 20 x 20 m cell. In other words, HSD of CD is a focal statistic of neighboring sub-cells that correspond to the original cell resolution. HSD of CD does not explicitly account for spatial arrangement, only variability, but Fig. 2 shows this metric can characterize important structural arrangements related to clump-gap patterns. The use of a 1 m^2 sub-cell size to calculate HSD of CD requires dense LIDAR data, and more typical acquisitions with lower point densities may need larger sub-cell sizes.

Table 3. Description of input features for classification of canopy structure.

Feature Name	Units	Theoretical Bounds	Range Observed	Description
L-location (mean), λ_1	Meters	*	2.01 – 50.53	Mean height of all points in the sample distribution.
L-scale, λ_2	Meters	≥ 0	0.00 – 12.13	Measurement of scale or dispersion, similar to standard deviation but with less weight given to outliers (Hosking 1990).
L-skewness, τ_3	Dimensionless (ratio)	-1 to 1	-0.98 – 1.00	Ratio of 3rd L-moment to L-scale (λ_3/λ_2). Analogous to skewness, a measure of distribution shape, high or low values indicate departure (or skew) from a symmetrical distribution (Hosking 1990).
L-kurtosis, τ_4	Dimensionless (ratio)	$\frac{1}{4}(5\tau_3^2 - 1)$ to 1	-0.25 – 1.00	Ratio of 4th L-moment to L-scale (λ_4/λ_2). Analogous to kurtosis, a measure of distribution shape, but with less weight given to extremes. Generally, high values indicate heavy-tailed and/or peaked distributions with low values describing the inverse (Hosking 1990).
Canopy Density, CD	Percent	0 to 100	0.00 – 100.00	Ratio of first laser pulse returns above 2 meters to the total, similar to canopy cover (field measurement), and used in this study as a measure of the amount of vegetation coverage in the canopy.
Horizontal Standard Deviation of Canopy Density, HSD of CD	Percent	0 to $(\max(CD) - \min(CD))/2$	0.00 – 50.00	Standard deviation of 1 m^2 canopy density within 20x20 m grid cells, used to describe dispersion/variability in the horizontal plane.

*L-location is theoretically unconstrained, but the sample distribution must have a finite mean for L-moments to exist.
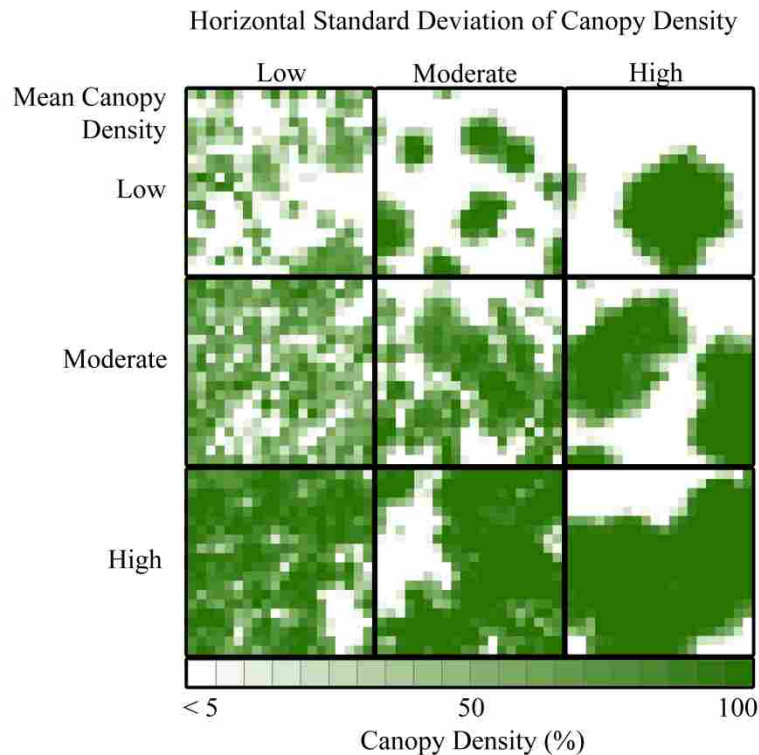
Figure 2. Examples of the spatial configuration of 1 m² canopy density at varying levels of canopy density (CD) and horizontal standard deviation of canopy density (HSD of CD) at the 20 m cell size (black boxes). 'Low' refers to CD's from 24.8-27.9% and HSD of CD's from 27-27.9%. 'Moderate' ranges from 49.3-50.5% and 36.0-36.4%. 'High' ranges from 72.5-75.6% and 41.9-45.1%. Examples taken from landscape 10.

2.4) Plot Data

Following Kane et al. (2010b)'s analysis, we tested whether selected features corresponded to field metrics of canopy structure. 113 plots were measured throughout the summer of 2015 coincident with the LiDAR campaigns. The field plots were located on a regular grid spaced at approximately 1 plot per 80 ha within landscapes 10, 11, and the southern portion of 9 (Fig 1). Trees greater than 12.7 cm diameter at breast height (DBH, mature trees) were measured on a 0.1 ha fixed-radius plot with total tree height, height of live crown, and DBH recorded for each individual. Tree saplings, defined as greater than 2 m in height and less than 12.7 cm in DBH, were measured in an 81 m² area at plot center and tallied within 2.5 cm DBH classes. A mean height was also estimated for each 2.5 cm DBH class. Seedlings were not considered as they were below the 2 m canopy threshold. We calculated Spearman's rank correlation coefficient between our input features and plot metrics. We also classified each plot as a meta-class for validation outlined in section 2.7.

2.5) Structure Classification

Our approach utilized Random Forests (RF) to find similarities between sample cells, hierarchical clustering to group based on estimated similarity, and then RF again to classify the landscape based on cluster labels. The algorithm was applied individually to each landscape within the study area (Fig. 1). We took this approach to explore the variation in structure classes across a large geographical area while maintaining the ability to identify relatively rare structural classes with localized areas of high density. Every landscape had several structural classes defined, and we then grouped these classes into meta-classes through equivalence testing. Analysis was conducted using R statistical software (version 3.3) with the following packages: randomForest (model development; Liaw and Wiener 2015), raster (model prediction; Hijmans 2016), cluster (hierarchical clustering; Maechler et al. 2016), maptree (tree pruning; White and Gramacy 2012), gstat (variogram modeling; Pebesma and Graeler 2016), and ggplot2 (visualization; Wickham 2009).

We first used RF in an unsupervised fashion to assess dissimilarity between samples. Many unsupervised methods, including hierarchical clustering used here, require a measure of dissimilarity or distance between samples. RF dissimilarity has been shown to naturally weigh feature contributions, to display non-parametric qualities including robustness to outliers and skewed distributions, and to allow features with different units and ranges (Breiman and Cutler 2003, Shi and Horvath 2006, Seligson et al. 2005), all attributes important to a generalized framework. We randomly sampled 10% of each landscape for training data. We chose 10% as a reasonable number of training samples (mean of 22,772 training samples for each landscape) balanced by the computational costs to calculate the similarity matrices. Synthetic datasets were created by randomly sampling reference distributions of the input features. We then trained an RF model to distinguish observed data from the synthetic (two classes). A similarity (or proximity) matrix was developed by calculating the frequency real observations appear in the same terminal tree node divided by the total number of trees. We grew five separate forests for each landscape, each containing 1000 trees, to minimize the random effects of the generated synthetic datasets (Shi and Horvath 2006). Three of the six variables were randomly selected as candidates for splitting at each tree node. Seventy-five percent of the training data was randomly selected for use in each individual tree. Minimum terminal node size was kept to 1% of the tree's training dataset to reduce overfitting and increase stability of the similarity measure. We averaged the five resulting similarity matrices to calculate a final measure of dissimilarity (defined as one minus similarity).

Hierarchical clustering then grouped training samples using dissimilarity as the metric for aggregation. It affords flexibility in cluster number if project objectives desire a different number of clusters. We used the Kelley-Gardner-Sutcliffe (KGS) penalty function to determine the appropriate number of clusters. The function seeks to minimize the within-cluster mean dissimilarity relative to the mean dissimilarity across all clusters with a penalty function for a larger number of clusters (Kelley et al. 1996).

Each training sample was assigned a cluster label and used to train a subsequent RF model in supervised mode for class prediction in each landscape. One forest was trained with 10,000 trees for each landscape. Seventy-five percent of the training data was used for each tree, and minimum node size kept to 1% of the total sample size to prevent overfitting. The clusters found and the resulting RF model's predictions were then considered 'landscape classes' from which 'meta-classes' could be derived.

2.6) Aggregation to Meta-classes

We aggregated landscape classes into meta-classes with pairwise equivalence testing. While meta-classes could be aggregated by any combination (or all) of the six input features, we chose to use the four that corresponded to the vertical distribution of vegetation — the L-moments. Our reasoning was threefold-fold: (1) *within-cell* forest structure is traditionally defined in the vertical plane (Lefsky et al. 2002, Lim et al. 2003); (2) we desired to explore how many structural classes could have equivalent vertical distributions but different CD and HSD of CD values; (3) we wanted to offer an example of one of the multiple ways to aggregate in order to highlight the utility and flexibility of meta-class aggregation. In this way, landscape classes were derived using the comprehensive set of features to create base data from which a variety of meta-classes could be derived dependent on objectives.

Each feature variable was first tested for spatial autocorrelation to maintain assumptions of independence in all subsequent statistical tests. We computed semivariograms for each feature in each landscape using a 20 m distance interval corresponding to cell size. We overlaid the 17 subsequent response curves for each feature and also estimated the semivariogram range values. A 300 m separation of samples (15 cells), corresponding to 0.4 % of the total sample count, was determined optimal to maintain a reasonable expectation of no spatial dependence while also keeping an adequate number of samples for robust statistical testing.

After sampling, each landscape class was compared to every other for each of the four vertical features (L-moments). Non-parametric Mann-Whitney tests for equivalence were used with a 20% symmetrical equivalence interval, the higher end of suggested interval values (Wellek 1996, Wellek 2010). The equivalence interval can be used as a tuning parameter for the number of meta-classes dependent on objectives (i.e. small or large number of resulting meta-classes). Equivalence testing was completed in R using scripts from Wellek (2010). Landscape classes that were determined equivalent ($p = 0.05$) in all four vertical features were then aggregated to a meta-class. Pairwise comparison allows for the possibility that a landscape class would test equivalent to multiple meta-classes. For these cases, we attributed the landscape class to the meta-class that was closest in terms of the landscape class's median values compared to the meta-class's median values.

2.7) Validation

We aimed to determine whether our method was suitable for its intended purpose, thereby following a broad definition of validation (Mayer and Butler 1993). A data-driven framework to identify and compare structure classes must be consistent, statistically robust, and relate to field-measured canopy structure. We correlated our input features to field-measured plot data (described in section 2.4). We then conducted a three-step validation to test:

1) *If the method showed consistency in derived landscape classes and predictions over similar landscapes.* We overlapped the extent of two adjacent landscapes by approximately 10% of their respective area and assessed the predictions from each classification through inspection of a matching matrix and calculation of the Kappa coefficient (Cohen 1960). We chose landscapes 10 and 11 because they fell within Lubrecht Experimental Forest and have similar species composition, environmental gradients, and disturbance histories. Thus, we expected derived classes and predictions to be similar. We used table 2 from Vierra and Garrett (2005) as the means to interpret the Kappa coefficient's value and selected a threshold of greater than 0.6 ('substantial agreement') to demonstrate consistency.

2) *If the meta-classes were statistically distinct in the native point cloud-derived probability density function (PDF).* We used two-sample Kolmogorov-Smirnov (KS) tests to assess differences in both location and shape of the empirical distribution functions (Conover 1971). Fedrigo et al. (2018) showed that differences in vertical distributions of LiDAR points relate to ecologically relevant forest structure classes. If feature-driven classification leads to statistically differentiated vertical PDFs, then support is present for both the *a priori* feature selection and the method of meta-class aggregation. For each meta-class, we randomly selected up to 500 cells (after sampling to remove spatial autocorrelation), extracted the raw LiDAR points, and then fit PDFs to each cell's points. We also created 2D kernel density heat maps of the vertical PDFs to visually assess their distinctiveness.

3) *If the meta-classes showed differences in field-sampled plot data.* After extracting the LiDAR points in 0.1 ha circles equivalent to the field plot size and calculating input features using the same process described in section 2.3, we used the RF model from each respective landscape to predict the class for each plot and then assigned them to the respective meta-class. With a small sample size, we first used non-parametric Kruskal-Wallis multiple comparison tests to ensure Mann-Whitney pairwise comparisons of LiDAR class and field metrics were robust. We also applied a Bonferroni correction to the significance level.

For the plot data, we compared the metrics of mean tree height, the ratio of sapling to mature tree basal area, quadratic mean diameter of mature trees, and the ratio of live crown to total height of mature trees. While our classification considers more than these metrics, three-dimensional structure is difficult to assess with traditional field sampling and would require more

extensive measurement than was available for this study. Also, we acknowledge that the field plots measured only a small portion of the meta-classes identified. These caveats highlight the difficulty in sampling forest structure variability over large areas and support the continued development of data-driven approaches.

3) Results

Each feature selected showed significant correlation to at least one field metric (Table 4). Canopy density showed the strongest correlation (0.78) to trees per hectare. L-scale showed the most significant correlations with six of eight comparisons being significant at $p < 0.01$. L-kurtosis had the least number of significant correlations (1 of 8), which was the ratio of live crown to total height of live trees.

Table 4. Spearman rank correlation of LiDAR features and field metrics derived from plot data.

	LiDAR Features					
Field Metrics	Mean Height	L-scale	L-skewness	L-kurtosis	Canopy Density	SD of Canopy Density
Trees per ha		-0.37	-0.28		0.78	
Quad. Mean DBH	0.29	0.46			-0.48	
SD of DBH		0.40	0.35			0.28
Mean Tree Height	0.63	0.52	-0.36			-0.26
SD of Tree Height	0.34	0.60	0.2			
Sapling Mean Height	-0.26					-0.33
Sapling:Mature Basal Area Ratio		-0.49			0.38	
Live Crown:Total Height Ratio	-0.40		0.43	-0.34	-0.50	0.18

Values shown are significant at $p < 0.01$.

Seventy landscape classes were identified from the seventeen landscapes tested. Fifteen landscapes had four classes each, and two landscapes had five classes each. While predicting over each landscape, the mean RF out-of-bag error rate was 0.11 for all classes in all landscapes. The maximum error rate of 0.27 was for class 3 in landscape 7 while the minimum was 0.04 for class 1 also in landscape 7. Mean prediction error rates ranged from 0.08 to 0.15 for each landscape. The matching matrix showed 'substantial' agreement in class prediction for the overlapping areas of landscapes 10 and 11 with a Kappa coefficient of 0.72 (Table 5).

Table 5. Matching matrix of class predictions in overlapped portions of landscapes 10 and 11 (see Fig. 1 and Table 1 for landscape descriptions).

Landscape 11	Landscape 10 Classes			
	1	2	3	4
1	7368	3529	0	189
2	88	8589	3016	50
3	362	958	21246	9584
4	2017	200	1596	27582

Kappa Coefficient: 0.72

After aggregation, fourteen meta-classes were derived from the seventy landscape classes (Table 6). The top nine classes represented 94.7% of the study area (Figs. 3 and 4). The bottom five were unique to a particular landscape (i.e. were not equivalent to any other landscape class), and each represented less than 1.5% of the total area (Fig. 5). KS tests showed significant differences ($p < 0.01$) in vertical distributions for every meta-class (Fig. 6). D-statistics of each pairwise KS test ranged from very similar (0.04) for meta-classes F and C to highly different (0.84) for meta-classes K and N.

Table 6. Description of meta-class median feature values and area coverage following aggregation of individual landscape classes. See Figs. 4 and 5 for graphical depiction of feature distributions for each meta-class. Mean height (L-location), L-scale, L-skewness, and L-kurtosis were used to aggregate to meta-classes. HSD of CD refers to horizontal standard deviation of canopy density.

Meta-Class	Coverage of Study Area (%)	Mean Height (m)	L-scale (m)	L-skewness	L-kurtosis	Canopy Density (%)	HSD of CD (%)
A	24.3	8.3	2.2	0.08	0.07	51.4	37.0
B	16.7	10.1	2.6	0.03	0.05	53.6	37.3
C	13.9	15.7	3.5	-0.09	0.07	70.4	34.7
D	8.3	12.6	3.4	0.01	0.04	58.8	37.8
E	7.4	6.4	1.6	0.14	0.07	33.6	36.2
F	6.9	15.6	4.1	-0.03	0.05	64.9	37.1
G	6.3	10.8	2.2	-0.10	0.07	77.0	29.7
H	5.7	4.9	1.1	0.18	0.09	27.4	34.1
I	5.2	13.9	3.2	-0.06	0.05	66.4	36.4
J	1.5	14.2	2.5	-0.08	0.10	29.9	33.5
K	1.2	15.8	2.6	-0.11	0.13	32.8	34.4
L	1.0	15.7	5.0	0.03	0.02	64.0	38.4
M	0.8	17.0	2.9	-0.19	0.13	42.3	33.2
N	0.7	2.9	0.4	0.28	0.12	3.5	10.7
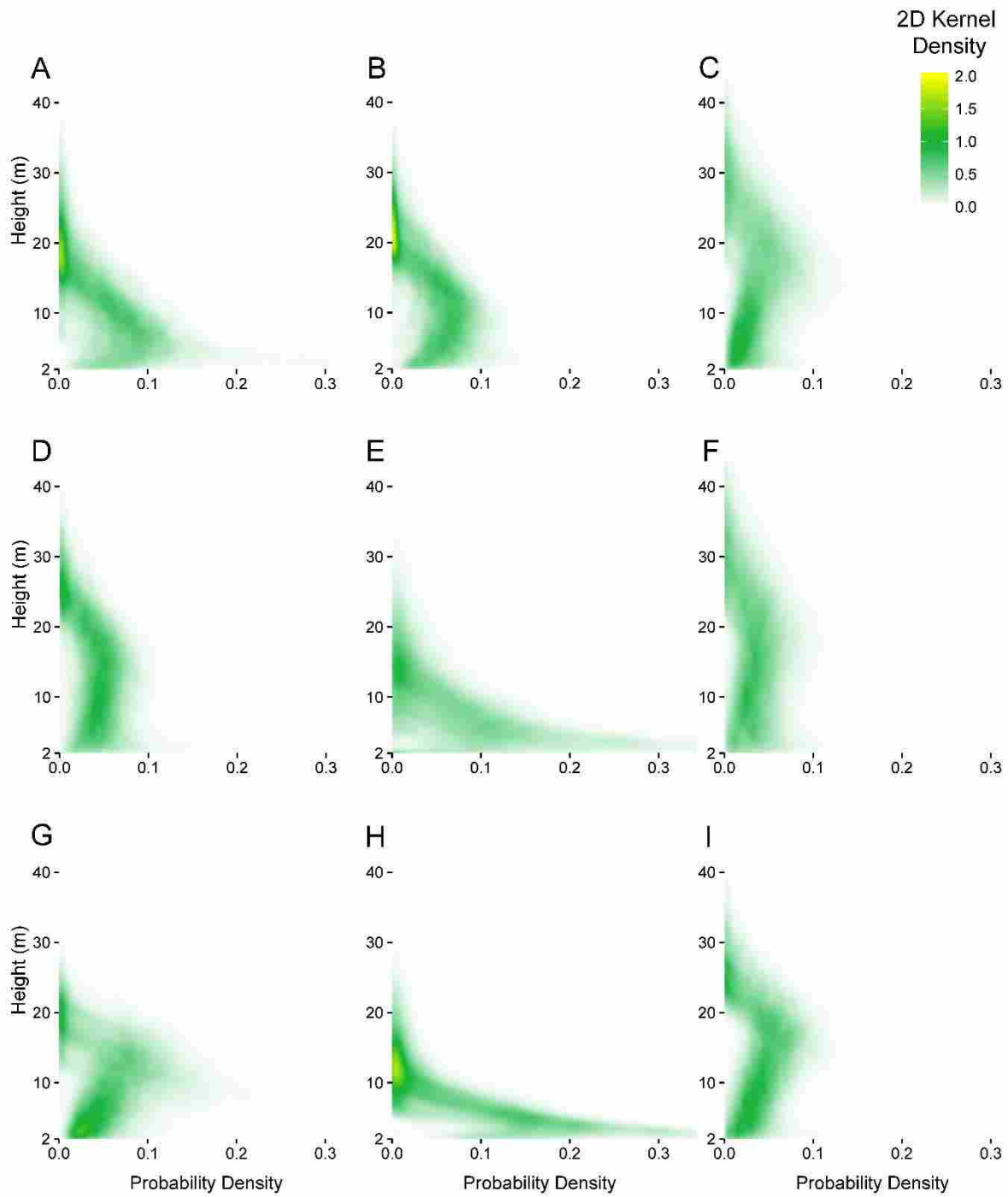
Figure 3. Heat map (2D kernel density) of vertical probability density derived from sample cells of the native LiDAR point clouds for the top nine meta-classes (letter designations) ordered by percent coverage (descending) in the study area.
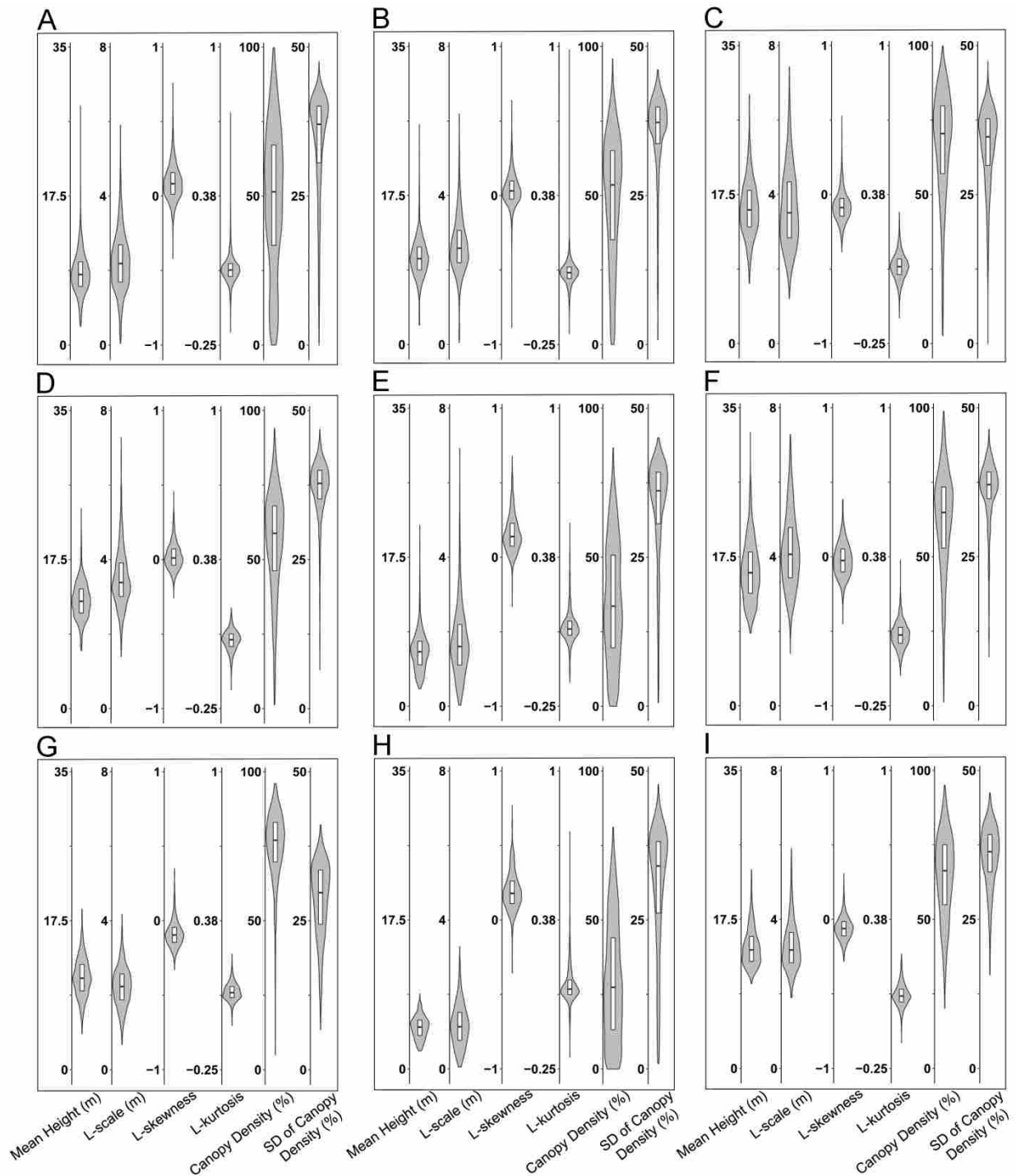
Figure 4. Violin plots of the top nine meta-classes (letter designations) sorted by area covered (descending) showing the median (bold horizontal line) and interquartile ranges (white boxplots) along with probability density of values over the full range (gray area) of input features used to classify canopy structure.
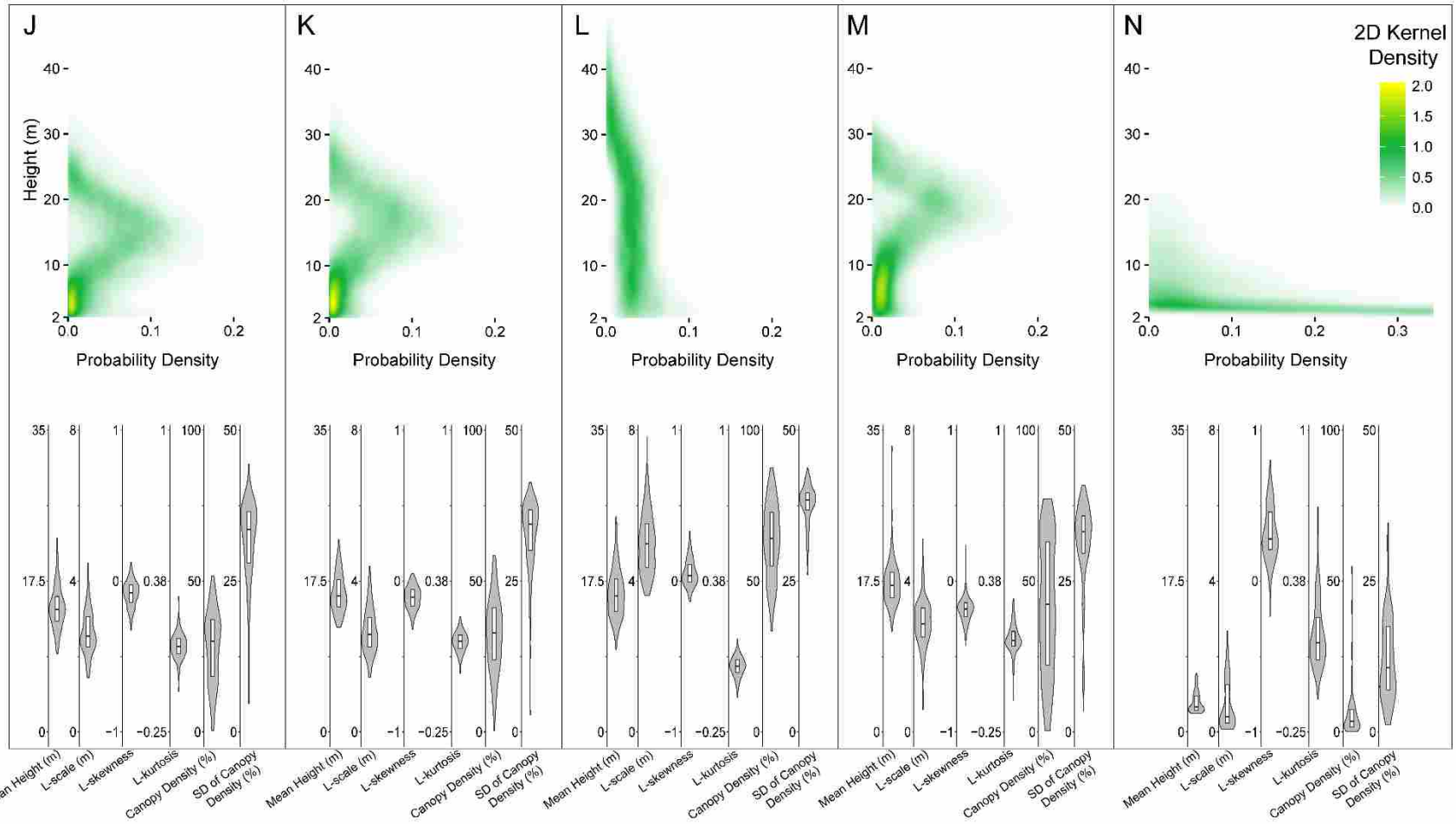
Figure 5. Heat map (2D kernel density) of vertical probability density derived from native LiDAR point clouds for the five rare classes (letter designations) along with violin plots (bottom row) showing the median (bold horizontal line) and interquartile ranges (white boxplots) along with probability density of values over the full range (gray area) of input features used to classify canopy structure.

Meta-Class

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A		0.14	0.46	0.30	0.18	0.44	0.21	0.36	0.41	0.48	0.56	0.41	0.59	0.60
B			0.35	0.18	0.31	0.35	0.08	0.48	0.30	0.34	0.44	0.34	0.49	0.66
C				0.17	0.58	0.04	0.36	0.70	0.10	0.16	0.11	0.08	0.14	0.77
D					0.43	0.19	0.19	0.57	0.12	0.19	0.26	0.21	0.31	0.70
E						0.54	0.39	0.18	0.54	0.62	0.68	0.50	0.69	0.48
F							0.37	0.66	0.14	0.20	0.14	0.05	0.15	0.74
G								0.55	0.29	0.31	0.42	0.36	0.49	0.72
H									0.67	0.75	0.80	0.62	0.80	0.39
I										0.08	0.15	0.18	0.20	0.76
J											0.14	0.23	0.24	0.81
K												0.18	0.10	0.84
L													0.19	0.71
M														0.82
N														

Figure 6. Meta-class comparison matrix showing the D-statistic of the two-sample Kolmogorov-Smirnov (KS) test. Vertical probability distribution functions were derived from native point clouds randomly sampled from each meta-class. All tests rejected the null hypothesis of being drawn from the same distribution at $p < 0.01$.

Visual interpretation of the meta-classes indicated four basic shapes of vertical distributions. Meta-classes H and N had nearly all their canopy material low to the ground with low mean height, low L-scale, high L-skewness, and high L-kurtosis (Table 6, Fig. 3). Meta-classes A and B distributions also had a majority of points low but with taller tails implying a sparse overstory. Meta-class A's high sapling to mature tree basal area ratio (0.2, Table 7) found in the field data supports this interpretation. These classes were typified by moderate mean heights, moderate L-scale, positive L-skewness, and moderate L-kurtosis. Meta-classes C, F, and L exhibited tall, vertically dispersed structure with high mean heights, high L-scale, variable L-skewness, and low to moderate L-kurtosis. The last shape was depicted by meta-classes J, K, and M showing heavy overstory dominance with little understory (Fig. 5). These classes were typified by high mean heights, moderate L-scale, high negative L-skewness, and high L-kurtosis. The remaining meta-classes fit between or were variations of these four basic shapes.

Although not directly used to aggregate meta-classes, canopy density (CD) and horizontal standard deviation of canopy density (HSD of CD) showed several trends. Dominant meta-classes A and B's CDs spanned nearly the entire range with median values

of 51.4% and 53.6%, respectively (Fig. 4). Meta-class G had the highest median CD (77%) paired with low HSD of CD (median 29.7%), suggestive of high uniformity in the horizontal plane. Meta-classes E and H had predominantly low CD (median 33.6% and 27.4%) and high HSD of CD (median 36.2% and 34.1%), indicating more variability in the horizontal domain. Meta-class N was distinct with low CD (3.5%) and relatively high HSD of CD (10.7%), suggesting sparse, variable vegetation in the horizontal plane. In general, HSD of CD had a restricted range of variation in 12 of the 14 meta-classes with median values from 33.2-38.4%.

The field sampling data allowed comparison of two similar meta-classes, C and I, and the most prevalent, meta-class A. Significant differences (Bonferroni-corrected $p = 0.017$) were found for at least one metric among all three classes (Table 7). Meta-class A had significantly lower mean tree height and a larger live crown ratio compared to meta-classes C and I. The ratio of sapling to mature tree basal area was higher for meta-class A compared to I but not significantly different from C. Meta-classes C and I were the same for all metrics except the sapling to mature tree basal area ratio. The Kruskal-Wallis test showed no difference among all three meta-classes for tree sapling mean height and mature tree quadratic mean diameter. Thus, no Mann-Whitney pairwise comparison tests were performed for those metrics.

Table 7. Comparison of field sampled data for three meta-classes. Different superscript letters denote significant differences as determined first by Kruskal-Wallis multiple comparison tests followed by Mann-Whitney pairwise comparisons (Bonferroni-corrected significance level of $p = 0.017$).

Meta-class	Plots Sampled	Mean Height of Mature Trees (m)	Sapling:Mature Basal Area Ratio	Quadratic Mean Diameter (cm)	Live Crown Ratio	Sapling Mean Height (m)
A	32	15.3[A]	0.20[A]	30.2[A]	0.63[A]	4.6[A]
C	20	17.3[B]	0.17[A]	26.9[A]	0.50[B]	5.9[A]
I	32	18.2[B]	0.04[B]	30.7[A]	0.51[B]	4.7[A]

4) Discussion

Our primary objective was to create a parsimonious framework for determining predominant forest structure over large and diverse geographical areas without pre-defining classes. The approach is explicitly flexible at multiple levels through a hierarchical and stratified design recognizing that no rigid classification system will meet all potential objectives (O'Hara et al. 1996). The input features comprehensively describe vertical and

horizontal distribution of canopy material and are interpretable. RF-based similarity naturally weighs feature contributions (Shi and Horvath 2006, Seligson et al. 2005), and hierarchical clustering allows for selection of any number of landscape clusters. Meta-class aggregation can then use any combination of input features.

Two important concerns for this approach are how well the structural variation is characterized and the ecological relevancy of identified classes. O'Hara et al. (1996) identified seven structural classes defined by successional traits characteristic of Inland Northwest forests. Falkowski et al. (2009) mapped these classes using a large suite of LiDAR metrics with a high degree of accuracy. Comparison of height profiles from these studies to the meta-classes derived here shows that our approach captures the range of variability, and the meta-classes can be intuitively grouped to the seven major successional types of O'Hara (1996).

Structure development proceeds along a continuum forever in transition (Franklin et al. 2002). Binning this spectrum of variation can be as specific as necessary. We aggregated meta-classes on vertical distribution features, the L-moments, and identified fourteen significantly different structural meta-classes. However, evaluation of the height profiles and associated variance along with comparison of the distributions of input features shows overlap in several of the classes (Figs. 3, 4, and 5). A primary cause is the clustering by landscape design. A cell in two different landscapes with the same feature values could be aggregated to different meta-classes. For example, two adjacent landscapes cells near their connected edge are likely similar but may be on the fringes of their respective class centers. Each landscape class may then be aggregated into a different meta-class. Thus, the fringes of the meta-classes must be viewed as fuzzy with no discrete delineation. This example also highlights the existence of edge effects in the design with the potential for arriving at different landscape classes with different landscape delineations, inherent in any data-driven approach. However, our analysis of landscapes 10 and 11 shows stability in the classes when run over similar landscapes (Table 5). A simple solution to reduce edge effects and have discrete delineations between classes is to derive feature thresholds from the meta-classes. From this perspective, our approach discovers and compares existing structure classes over broad areas, and thresholds could then be used to define discreet classes for mapping or other product-oriented applications.

A key question, then, is why produce sub-classifications first and aggregate to meta-classes instead of sampling all landscapes together and applying a single classification? Our primary objective was to develop a framework that can be applied at any scale, with any number of landscapes, and with consistency and flexibility. Sub-classification followed by aggregation ensures that each landscape has representation with an equal sampling rate

regardless of size, prioritizes identification and comparison of locally dominant, but globally rare structure classes, allows for meta-class aggregation with any combination of features, and facilitates geographic comparisons by giving each landscape its own set of classes derived from the same features and methods. Leiterer et al. (2015) suggested a spatially stratified clustering approach to maintain the distinctness of localized structure classes (i.e. locally dominant, but globally rare classes). We envision the development of a database of classified landscapes that is built incrementally over time as new LiDAR datasets become available without having to reclassify previous landscapes. Additionally, the flexibility inherent in meta-class aggregation is amenable to diverse or transient conservation and ecological objectives.

The selection of features was driven by the project design goal of consistent and parsimonious characterization of LiDAR point distributions. Feature selection based on search algorithms or optimization criteria such as Akaike information criterion (AIC) will likely choose different features when presented different datasets, especially when a large number of features and/or models are evaluated. For this reason (and others described above), we instead relied on *a priori* selection of features.

Naturally, other features that would meet our design goals could have been selected to characterize forest structure. For example, many studies derive the same features at multiple height strata (e.g., CD at 2-4 m, 4-6 m, etc.; Hudak et al. 2008, Falkowski et al. 2009). While effective in certain situations, the large number of resulting features, arbitrary thresholds, and potential lack of consistency can confuse interpretation and limit the ability to compare (Chen 2013). Similarly, features such as the rumple index (ratio of canopy outer surface area to ground surface area), which integrate three-dimensional horizontal and vertical heterogeneity (Parker et al. 2004, Kane et al. 2010a, b), can be difficult to interpret as variation in either the vertical or horizontal plane can drive value change. Even the L-moments present multiple options for feature selection. For example, the L-coefficient of variation (L-CV, λ_2/λ_1) can be used in place of L-scale (λ_2) as a measure of dispersion. L-CV has the advantage of discreet theoretical bounds (Hosking 1990, Valbuena et al. 2017). However, we chose L-scale over L-CV for three reasons: 1) Hosking (1990) preferred it as a characterization of dispersion; 2) L-CV had much higher Spearman correlation ($\rho > 0.9$) to L-location (mean) than L-scale in our test landscapes (9, 10, 11); 3) Tall forests should inherently have more potential dispersion than short forests. We argue that measures of location and scale of height distributions of forest canopy material should depend on the absolute values of the sample data (although arguments for or against application of this principle to dispersion metrics border on the philosophical). In any case, there is no one collection that captures forest structure in its entirety, and our broader framework is amenable to a different selection of features. However, consistency and low dimensionality

should be maintained for robust comparison and interpretation with any set of features in this framework.

Because the L-moments do not explicitly account for the abundance and horizontal distribution of vegetation, we included the features CD and HSD of CD. Spies (1998) stated that horizontal variation in canopy density is one of the four major components of forest structure. Larson and Churchill (2012) found inherent horizontal aggregation of trees at scales less than 20 m in the western US and noted its ecological importance. Additionally, Hudak et al. (2006) discovered that sub-plot variation in canopy cover was an important predictor of tree basal area and tree density. We also observed complimentary correlations of CD and HSD of CD to field metrics in our plot data (Table 4). Although alternative spatial metrics such as Ripley's K are attractive due to their spatial explicitness, they are not feasible to calculate on every cell for large area analysis because every LiDAR point must be compared to every other. Since HSD of CD measures sub-cell variability rather than spatial arrangement of CD, there are certain structural arrangements that may have comparable HSD of CD values without being structurally similar (e.g., alternating vs. symmetrically-split high and low sub-cell CD values). Despite this limitation, we argue that HSD of CD meets our objective of parsimony, is based on a conventional measure of variability (interpretable), and would rarely have enough similarity in the other structural features for those confounding cases to create confusion in a classification. Altogether, the six features selected represent a comprehensive, interpretable set of features to describe canopy structure.

The study area is dominated by younger, largely multi-strata forest structure. L-skewness (L-skew) has been used successfully to divide forest stands into oligophotic (low light, negative L-skew) and euphotic (high light, positive L-skew) structural types (Valbuena et al. 2013, Valbuena et al. 2017). Using L-skew = 0 as a threshold, 64.1% of the study area is classified into meta-classes with understory dominance (large euphotic zone) and 35.9% with overstory dominance (large oligophotic zone). Although the work of Valbuena et al. (2017) was restricted to Boreal forest types, the meta-classes in our classification with high, positive L-skew also have lower mean heights and canopy densities; generally supporting the inference of understory dominance and the ability of L-skew to delineate structure in diverse forest types.

The less common meta-classes M, K, and J are exceptions, exhibiting high negative L-skew and low CD. They represent the open, tall, single-strata structural type, which is the climax structure of dry, low-elevation forests traditionally maintained by frequent, low-severity fire (O'Hara et al. 1996, Agee 1998). Today, these meta-classes are largely created by mechanical treatments in the study area as evidenced by the clumping of meta-classes K

and J in the heavily-managed Lubrecht Experimental Forest (landscapes 10 and 11). Mean height largely separates these three meta-classes from each other, and a relatively low L-scale, negative L-skew, and highly positive L-kurtosis distinguishes them from other meta-classes.

A strength of the stratified clustering method is its ability to identify and group globally prevalent classes along with locally predominant, but rare overall structure types. The rarest meta-class, N, represents a severely burned portion of the 2002 Crazy Horse wildfire. Analysis of high resolution aerial imagery shows an abundance of shrubs and regenerating conifers. Low mean height, low L-scale, high L-skew, and high L-kurtosis typify this structure class, which is essentially an extreme version of meta-class H. These structural types have important, early-successional ecological considerations (Swanson et al. 2010). Indeed, many of the rare classes are similar in shape to prevalent meta-classes, but with one or more of their L-moment features pushed to an extreme. Class L, representing 1% of the study area, has the lowest L-kurtosis, a slightly positive L-skew, and the largest L-scale indicating a tall, continuous canopy structure typical of mixed-conifer forests in the high elevation Western Canadian Rockies ecoregion from which it's derived (landscape 6).

By aggregating only on the features that characterize the vertical domain, we show that multiple meta-classes exist on nearly the entire range of CDs especially the two most prevalent, meta-classes A and B, representing 41% of the study area (Fig. 4). Both of these classes have low to moderate mean heights, moderate L-scales, positive L-skews, and low to moderate L-kurtosis. Their vertical distributions imply broken overstories over a stratum of understory trees, characteristic of forests often targeted for restoration (Graham et al. 2004, Noss et al. 2006). Other meta-classes span a large range of CDs, but their distributions imply a self-organization of certain vertical canopy structures at particular CDs, such as the aforementioned L-skew and CD relationships.

The HSD of CD shows moderate values within nearly all meta-classes (Fig. 4). These values, however, must be interpreted with respect to the value of the CD metric. For example, meta-classes G and H have a 4.4% difference in median HSD of CD but meta-class G has a 77% median CD compared to H's 27.4%. This indicates that meta-class H has more relative horizontal variability. Scale likely influences the HSD of CD more than any other input feature. Leiterer et al. (2015) found a 0.12 absolute mean difference in input data correlation comparing 1 m to 20 m horizontal cell sizes. However, differences reached an asymptote near 20 m resolution. There likely exists a minimum cell size at which HSD of CD is not relevant potentially below the average crown diameter of canopy trees. We speculate that HSD of CD would become increasingly relevant for structure classification

as cell size increases and therefore includes more potential horizontal structure variation. The point density and laser footprint size must also be considered while determining both sub-cell and cell resolutions. The LiDAR data used here had unusually high point density, and typical acquisitions may need to use larger sub-cells (e.g., 5 m). Further research is warranted for within-cell horizontal variation and scaling.

As they are related, both CD and HSD of CD can be used to create sub-classes within the meta-classes similar to the matrix classification shown in Fig. 2. We suggest first grouping based on vertical distribution and after achieving the desired number of meta-classes, splitting these classes into low, moderate, and high CDs along with variable (high) or uniform (low) HSD of CD. A large number of classes will ensue, but the logical and hierarchical design should allow for relatively easy interpretation and application for a variety of objectives. This again highlights the strength of a flexible approach as it supports an *a posteriori* ability to select and modify the defining characteristics of a structure class. For example, meta-classes J, K, and M have similar vertical distributions with different mean heights driving the class separations (Fig. 5). These could logically be grouped into one meta-class for practical applications ignoring height differences of a few meters. For other meta-classes, height may be a defining characteristic, such as C and G, which share similar shape features.

5) Conclusions

Resource management is increasingly recognizing the need for consistent and comparable assessments of forest structures and spatial patterns as the pace of environmental change accelerates. The growing abundance of LiDAR and other remotely-sensed datasets necessitate new methods capable of efficiently deriving actionable information. We build on previous data-driven approaches of structural classification, such as Kane et al. (2010a) and Leiterer et al. (2015), to create a framework of identification along with comparison of multiple, diverse landscapes.

We identify several strengths to the approach:

1) Statistically robust design allowing for comparability with interpretable and consistent *a priori* feature selection.
2) Field sampling not required to develop structure classes.
3) Characterizes and maps globally dominant meta-classes across ecosystems.
4) Characterizes and maps locally dominant, but globally rare structure classes.
5) Provides inherent flexibility at multiple levels including number of clusters at the landscape level, features selected for meta-class aggregation, and an *a posteriori* ability to select and modify defining features of a meta-class.

Weaknesses:

1)   Meta-classes are fuzzy in nature with overlapping ranges of feature values.

2)   Assumes airborne LiDAR produces unbiased samples of 3-dimensional forest structure.

3)   Can be computationally expensive with certain parameterizations.

4)   Despite low-dimensionality, six input features still allows for a potentially confusing variety of combinations for aggregating to meta-classes.

5)   Scale affects results from size of landscapes to cell size. While L-moment features are resistant and meta-class aggregation methods tolerate differences in landscape size, HSD of CD is likely susceptible to changes in cell and sub-cell size.

6)   Derived classes and meta-classes are statistical in nature, and their ecological significance must be demonstrated with independent validation (e.g., field sampling).

We argue that a cursory knowledge of statistical moments and forest structure will allow for interpretation of the derived classes in this study's approach. As such, consistent comparisons of forest canopy structure that consider a diversity of landscapes and ecoregions should expand the perspectives of resource managers and improve their ability to assess and respond to environmental change.

Literature Cited

Agee, J.K. 1998. The landscape ecology of western forest fire regimes. Northwest Science 72: 24-34.

Barber, J., S.R. Brown Jr., and R. Ahl. 2012. Mid-level and base-level databases of the R1 existing vegetation (VMap) products. USDA Forest Service, Northern Region, Numbered Report 12-38.

Bergen, K.M., S.J. Goetz, R.O. Dubayah, G.M. Henebry, C.T. Hunsaker, M.L. Imhoff, R.F. Nelson, G.G. Parker, and V.C. Radeloff. 2009. Remote sensing of vegetation 3-D structure for biodiversity and habitat: review and implications for lidar and radar spaceborne missions. Journal of Geophysical Research 114: G00E06.

Bouvier, M., S. Durrieu, R. A. Fournier, and J. Renaud. 2015. Generalizing predictive models of forestry inventory attributes using an area-based approach with airborne LiDAR data. Remote Sensing of Environment 156: 322-334.

Breiman, L. 2001. Random forests. Machine Learning 45: 5-32.

Breiman, L. and A. Cutler. 2003. Random Forests Manual v4.0. Technical Report, University of California- Berkeley. Available online at ftp://ftp.stat.berkeley.edu/pub/users/breiman/Using.random.forests.v4.0.pdf.

Chen, Q. 2013. Lidar remote sensing of vegetation biomass. Taylor & Francis: Remote Sensing of Natural Resources. CRC Press, pp. 399–420.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20: 37-46.

Conover, W.J. 1971. Practical Nonparametric Statistics. New York: John Wiley & Sons, pp. 309-314.

Coops, N.C., T. Hilker, M.A. Wulder, B. St-Onge, G. Newnham, A. Siggins, J.A. Trofymow. 2007. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. Trees 21:295-310.

Culbert, P.D., V.C. Radeloff, C.H. Flather, J.M. Kellndorfer, C.D. Rittenhouse, and A.M. Pidgeon. 2013. The influence of vertical and horizontal habitat structure on nationwide patterns of avian biodiversity. The Auk 130: 656-665.

Dickinson, Y., E.K. Zenner, and D. Miller. 2014. Examining the effect of diverse management strategies on landscape scale patterns of forest structure in Pennsylvania using novel remote sensing techniques. Canadian Journal of Forest Research 44: 301-312.

Dupuy, S., G. Lainé, J. Tassin, and J-M. Sarrailh. 2013. Characterization of the horizontal structure of the tropical forest canopy using object-based LiDAR and multispectral image analysis. International Journal of Applied Earth Observation and Geoinformation 25: 76-86.

Ellsworth, D.S., and P.B. Reich. 1993. Canopy structure and vertical patterns of photosynthesis and related leaf traits in a deciduous forest. Oecologia 96: 169-178.

Falkowski, M.J., J.S. Evans, S. Martinuzzi, P.E. Gessler, and A.T. Hudak. 2009. Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. Remote Sensing of Environment 113: 946-956.

Franklin, J.F., T.A. Spies, R. Van Pelt, A.B. Carey, D.A. Thornburgh, D.R. Berg, D.B. Lindenmayer, M.E. Harmon, W.S. Keeton, D.C. Shaw, K. Bible, and J. Chen. 2002. Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. Forest Ecology and Management 155: 399-423.

Frazer, G.W., M.A. Wulder, and K.O. Niemann. 2005. Simulation and quantification of the fine-scale spatial pattern and heterogeneity of forest canopy structure: A lacunarity-based methods to analysis of continuous canopy heights. Forest Ecology and Management 214: 65-90.

Frazer, G.W., S. Magnussen, M.A. Wulder, and K.O. Niemann. 2011. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. Remote Sensing of Environment 115: 636-649.

Fredrigo, M., G.J. Newnham, N.C. Coops, D.S. Culvenor, D.K. Bolton, and C.R. Nitschke. 2018. Predicting temperate forest stand types using only structural profiles from discrete return airborne lidar. ISPRS Journal of Photogrammetry and Remote Sensing 136: 106-119.

Graham, R.T., S. McCaffrey, and T.B. Jain (tech. eds.). 2004. Science basis for changing forest structure to modify wildfire behavior and severity. General Technical Report. RMRS-GTR-120. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 43 p.

Gorgens, E.B., R. Valbuena, and L.C.E. Rodriguez. 2017. A method for optimizing height threshold when computing airborne laser scanning metrics. Photogrammetic Engineering & Remote Sensing 83: 343-350.

Halkidi, M., Y. Batistakis, M. Vazirgiannis. 2001. On clustering validation techniques. Journal of Intelligent Information Systems 17: 107-145.

Hawbaker, T.J., N.S. Keuler, A.A. Lesak, T. Gobakken, K. Contrucci, and V.C. Radeloff. 2009. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. Journal of Geophysical Research 114: G00E04.

Hijmans, R. J. 2016. raster: geographic data analysis and modeling. R package version 2.5-8.

Hosking, J.R.M. 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. Journal of the Royal Statistical Society 52: 105-124.

Hosking, J.R.M. 1992. Moments or L moments? An example comparing two measures of distributional shape. The American Statistician 46: 186-189.

Hosking, J. R. M., and J. R. Wallis. 1997. Regional frequency analysis: an approach based on L-moments. New York: Cambridge University Press, pp. 14-41.

Hudak, A.T., N.L. Crookston, J.S. Evans, M.J. Falkowski, A.M.S. Smith, P.E. Gessler, and P. Morgan. 2006. Regression modeling and mapping of coniferous forest basal area and

tree density from discrete-return lidar and multispectral satellite data. Canadian Journal of Remote Sensing 32: 126-138.

Hudak, A.T., N.L. Crookston, J.S. Evans, D.E. Hall, and M.J. Falkowski. 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sensing of Environment 112: 2232-2245.

Johnstone, J.F., C.D. Allen, J.F. Franklin, L.E. Frelich, B.J. Harvey, P.E. Higuera, M.C. Mack, R.K. Meentemeyer, M.R. Metz, G.L. Perry, T. Schoennagel, and M.G. Turner. 2016. Changing disturbance regimes, ecological memory, and forest resilience. Frontiers in Ecology and the Environment 14: 369-378.

Jones, T.G., N. C. Coops & and T. Sharma. 2012. Assessing the utility of LiDAR to differentiate among vegetation structural classes. Remote Sensing Letters 3: 231-238.

Kane, V.R., J.D. Bakker, R.J. McGaughey, J.A. Lutz, R.F. Gersonde, and J.F. Franklin. 2010a. Examining conifer canopy structural complexity across forest ages and elevations with LiDAR data. Canadian Journal of Forest Research 40: 774-787.

Kane, V.R., R.J. McGaughey, J.D. Bakker, R.F. Gersonde, J.A. Lutz, and J.F. Franklin. 2010b. Comparisons between field- and LiDAR-based measures of stand structural complexity. Canadian Journal of Forest Research 40: 761-773.

Kane, V.R., J. James A. Lutz, S.L. Roberts, D.F. Smith, R.J. McGaughey, N.A. Povak, M.L. Brooks. 2013. Landscape-scale effects of fire severity on mixed-conifer and red fir forest structure in Yosemite National Park. Forest Ecology and Management 287: 17-31.

Kelley, L.A., S. P. Gardner, and M.J. Sutcliffe. 1996. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. Protein Engineering 9: 1063-1065.

Latham, P.A., H.R. Zuuring, and D.W. Coble. 1998. A method for quantifying vertical forest structure. Forest Ecology and Management 104: 157-170.

Larson, A.J., and D. Churchill. 2012. Tree spatial patterns in fire-frequent forests of western North America, including mechanisms of pattern formation and implications for designing fuel reduction and restoration treatments. Forest Ecology and Management 267: 74-92.

Lawrence, R.L., and C.J. Moran. 2015. The AmericaView classification methods accuracy comparison project: a rigorous approach for model selection. Remote Sensing of Environment 170: 115-120.

Lefsky, M.A., W.B. Cohen, S.A. Acker, G.G. Parker, T.A. Spies, and D. Harding. 1999. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. Remote Sensing of Environment 70: 339-361.

Lefsky, M.A., W.B. Cohen, G.G. Parker, and D.J. Harding. 2002. Lidar remote sensing for ecosystem studies. BioScience 52: 19-30.

Lefsky, M.A., A.T. Hudak, W.B. Cohen, and S.A. Acker. 2005. Geographic variability in lidar predictions of forest stand structure in the Pacific Northwest. Remote Sensing of Environment 95: 532-548.

Leiterer, R., R. Furrer, M.E. Schaepman, and F. Morsdorf. 2015. Forest canopy-structure characterization: a data-driven approach. Forest Ecology and Management 358: 48-61.

Liaw, A. and M. Wiener. 2015. randomForest: Breiman and Cutler's random forests for classification and regression. R package version 4.6-12.

Lim, K., P. Treitz, M. Wulder, B. St-Onge, and M. Flood. 2003. LiDAR remote sensing of forest structure. Progress in Physical Geography 27: 88-106.

Lim, K. and P.M. Treitz. 2004. Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. Scandinavian Journal of Forest Research 19: 558-570.

Listopad, C., R.E. Masters, J. Drake, J. Weishampel, and C. Branquinho. 2015. Structural diversity indices based on airborne LiDAR as ecological indicators for managing highly dynamic landscapes. Ecological Indicators 57: 268-279.

Maechler, M., P. Rousseeuw, A. Struyf, and M. Hubert. 2016. cluster: finding groups in data. R package version 2.0.5.

Maltamo, M., Packalén, P., Kallio, E., Kangas, J., Uuttera, J. and Heikkilä, J., 2011. Airborne laser scanning based stand level management inventory in Finland. In Proceedings of SilviLaser 2011, 11th International Conference on LiDAR Applications for Assessing Forest Ecosystems, University of Tasmania, Australia, 1-10. Conference Secretariat.

Maltamo, M., O.M. Bollandsås, T. Gobakken, and E. Næsset. 2016. Large-scale prediction of aboveground biomass in heterogeneous mountain forests by means of airborne laser scanning. Canadian Journal of Forest Research 46: 1138-1144.

Mayer, D.G., and D.G. Butler. 1993. Statistical validation. Ecological Modeling 68: 21-32.

McGaughey, R.J. 2015. FUSION v3.5. USDA Forest Service, Pacific Northwest Research Station, Olympia, WA. http://forsys.cfr.washington.edu/

Miura, N., and S.D. Jones. 2010. Characterizing forest ecological structure using pulse types and height of airborne laser scanning. Remote Sensing of Environment 114: 1069-1076.

Murtagh, F., and P. Legendre. 2014. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? Journal of Classification 31: 274-295.

Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80: 88-99.

Niemi, M.T., and J. Vauhkonen. 2016. Extracting Canopy Surface Texture from Airborne Laser Scanning Data for the Supervised and Unsupervised Prediction of Area-Based Forest Characteristics. Remote Sensing 8: 582.

North, M.P., J.T. Kane, V.R. Kane, G.P. Asner, W. Berigan, D.J. Churchill, S. Conway, R.J. Gutiérrez, S. Jeronimb, J. Keane, A. Koltunov, T. Mark, M.Moskal, T. Munton, Z. Peery, C. Ramirez, R.Sollmann, A.M. White, S. Whitmore. 2017. Cover of tall trees best predicts California spotted owl habitat. Forest Ecology and Management 405: 166-178.

Noss, R.F., J.F. Franklin, W.L. Baker, T. Schoennagel, and P.B. Moyle. 2006. Managing fire-prone forests in the western United States. Frontiers in Ecology and the Environment 4: 481-487.

O'Hara, L.O., P.A. Latham, P. Hessburg, and B.G. Smith. 1996. A structural classification for inland Northwest forest vegetation. Western Journal of Applied Forestry 11: 97-102.

Omernik, J.M. and G.E. Griffith. 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environmental Management 54: 1249-1266.

Ozdemir, I., and N.M. Donoghue. 2013. Modelling tree size diversity from airborne laser scanning using canopy height models with image texture measures. Forest Ecology and Management 2013: 28-37.

Parker, G.G., M.E. Harmon, M.A. Lefsky, J. Chen, R. Van Pelt, S.B. Weis, S.C. Thomas, W.E. Winner, D.C. Shaw, and J.F. Franklin. 2004. Three-dimensional structure of an old-growth Pseudotsuga-Tsuga canopy and its implications for radiation balance, microclimate, and gas exchange. Ecosystems 7: 440-453.

Pebesma, E. and B. Graeler. 2016. gstat: spatial and spatio-temporal geostatistical modeling, prediction and simulation. R package version 1.1-4.

Popescu, S. C. and K. Zhao. 2008. A voxel-based lidar method for estimating crown base height for deciduous and pine trees. Remote Sensing of Environment 112: 767-781.

Pregitzer, K. S. and E.S. Euskirchen. 2004. Carbon cycling and storage in world forests: biome patterns related to forest age. Global Change Biology 10: 2052−2077.

PRISM Climate Group. 2012. Oregon State University, http://prism.oregonstate.edu.

Seligson, D. B., S. Horvath, T. Shi, H. Yu, S. Tze, M. Grunstein, and S. K. Kuristani. 2005. Global histone modification patterns predict risk of prostate cancer recurrence. Nature 435: 1262-1266.

Schoennagel, T., T.T. Veblen, and W.H. Romme. 2004. The interaction of fire, fuels, and climate across Rocky Mountain forests. BioScience 54: 661-676.

Shi, T. and S. Horvath. 2006. Unsupervised learning with random forest predictors. Journal of Computational and Graphical Statistics 15: 118-138.

Shugart, H.H., S. Saatchi, and F.G. Hall. 2010. Importance of structure and its measurement in quantifying function of forest ecosystems. Journal of Geophysical Research 115: 1-16.

Simonson, W.D., H.D. Allen, D.A. Coomes. 2014. Applications of airborne lidar for the assessment of animal species diversity. Methods in Ecology and Evolution 5: 719-729.

Smart, L.S., J.J. Swenson, N.L. Christensen, and J.O. Sexton. 2012. Three-dimensional characterization of pine forest type and red-cockaded woodpecker habitat by small-footprint, discrete-return lidar. Forest Ecology and Management 281: 100-110.

Spies, T.A. 1998. Forest structure: a key to the ecosystem. Northwest Science 72.

Swanson, M.E., J.F. Franklin, R.L. Beschta, C.M. Crisafulli, D.A. DellaSala, R.L. Hutto, D.B. Lindenmayer, and F.J. Swanson. 2011. The forgotten stage of forest succession: early-successional ecosystems on forest sites. Frontiers in Ecology and the Environment 9: 117-125

Valbuena, R., P. Packalen, L. Mehtatalo, A. Garcia-Abril, M. Maltamo. 2013. Characterizing forest structural types and shelterwood dynamics from Lorenz-based indicators predicted by airborne laser scanning. Canadian Journal of Forest Research 43: 1063-1074.

Valbuena, R., M. Maltamo, L. Mehtatalo, and P. Packalen. 2017. Key structural features of Boreal forests may be detected directly using L-moments from airborne lidar data. Remote Sensing of Environment 194: 437-446.

Vauhkonen, J. and J. Imponen. 2016. Unsupervised classification of airborne laser scanning data to locate potential wildlife habitats for forest management planning. Forestry 89: 350-363.

Vierling, K.T., L.A. Vierling, W.A. Gould, S. Martinuzzi, and R.M. Clawges. 2008. Lidar: shedding new light on habitat characterization and modeling. Frontiers in Ecology and the Environment 6: 90-98.

Vierra, A.J. and J.M. Garrett. 2005. Understanding interobserver agreement: the kappa statistic. Family Medicine 37: 360-363.

Wellek, S. 1996. A new approach to equivalence assessment in standard comparative bioavailability trials by means of the Mann-Whitney statistics. Biometrical Journal 38: 695-710.

Wellek, S. 2010. Testing statistical hypothesis of equivalence and noninferiority (2nd ed.). Boca Raton, FL: CRC Press, pp. 15-16 & 126-136.

White, D. and R. B. Gramacy. 2012. Maptree: mapping, pruning, and graphing tree models. R package version 1.4-7.

Whitehurst, A.S., A. Swatantran , J.B. Blair , M.A. Hofton and R. Dubayah. 2013. Characterization of canopy in forested ecosystems using full waveform lidar. Remote Sensing 5: 2014-2036.

Wickham, H. 2009. Ggplot2: elegant graphics for data analysis. Springer Science & Business Media.

Wulder, M. A., N. C. Coops, A. T. Hudak, F. Morsdorf, R. Nelson, G. Newnham, and M. Vastaranta. 2013. Status and prospects for LiDAR remote sensing of forested ecosystems. Canadian Journal of Remote Sensing 39: S1-S5.

Yu, X., J. Hyyppa, M. Holopainen, and M. Vastaranta. 2010. Comparison of area-based and individual tree-based methods for predicting plot-level forest attributes. Remote Sensing 2: 1481-1495.

Zhang, Z. X. Liu, J. Peterson and W. Wright. 2011. Cool temperate rainforest and adjacent forests classification using airborne LiDAR data. Area 43: 438-448.

Zhao, K., S. Popescu, and R. Nelson. 2009. Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. Remote Sensing of Environment 113: 182-196.

Zimble, D.A., D.L. Evans, G.C. Carlson, R.C. Parker, S.C. Grado, P.D. Gerard. 2003. Characterizing vertical forest structure using small-footprint airborne LiDAR. Remote Sensing of Environment 87: 171-182.

Chapter 3: Mapping forest canopy fuels over the western US with LiDAR-Landsat relationships

1) Introduction

Characterization of forest structure remains a priority for a variety of scientific research and land management objectives, as highlighted in Chapter 2. Forest management across the globe has integrated and now relies on these spatial data to inform policy and decision-making (Wulder et al. 2004). For example, the LANDFIRE project produced nationally-consistent and comprehensive forest characterizations for the US and has facilitated landscape-scale management including fuel and restoration treatment planning and assessment (Collins et al. 2011, Cochrane et al. 2012, Ryan and Opperman 2013, Drury et al. 2016), prescribed fire planning and implementation (Wiedinmyer and Hurteau 2010), and wildfire prediction, suppression, impact mitigation, rehabilitation, and assessment (Liang et al. 2008, Calkin et al. 2011, Ager et al. 2012, Ryan and Opperman 2013). For wildfire management in particular, LANDFIRE provides the spatial data for fire models that predict the spread and intensity of wildfires (Rollins 2009). Strategic and tactical decisions are increasingly becoming explicitly risk-based (Calkin et al. 2011, Noonan-Wright et al. 2011), and these fire predictions are essential for robust risk assessments. At least forty countries have national forest inventories with satellite imagery becoming the predominant data source for full coverage spatial data (McRoberts et al. 2010). In short, maps of forest attributes are essential to and in many ways have ushered in modern forest and fire management.

To create this broad spatial data, standard approaches utilize field and photo-based inventories which provide the explicit measurements of attributes (McRoberts et al. 2010). These are then related to satellite imagery thereby creating spatially-complete datasets (e.g. Tomppo et al. 2008). These assessments have risen in scale from local (e.g. Makela and Pekkarinen 2004) to regional (e.g. Ohmann et al. 2014) to global assessments (e.g. Hu et al. 2016) in parallel with the rise in computing performance and analytical sophistication. However, traditional field-based methodologies are labor intensive, often do not capture the range of variation on the landscape, are inconsistent in spatial distribution and methods of data collection, and cannot feasibly capture many sought-after three-dimensional (3D) attributes (Hawbaker et al. 2009, Maltamo et al. 2016). Light detection and ranging (LiDAR), a now pervasive active remote sensing technology, has provided large-area 3D datasets and facilitated the conceptualization of new 3D forest attributes (Lim et al. 2003). As shown in Chapter 2 (Moran et al. 2018), point cloud features and unsupervised classification frameworks can be used to describe complex forest properties consistently over a diversity of landscapes without intensive field campaigns.

The spatial and temporal scales of these LiDAR data are still limited for a variety of management needs though. LiDAR is commonly attained in preparation for landscape-scale projects such as the Collaborative Forest Landscape Restoration Program implemented by the US Forest Service, but for acute disturbances threatening lives, homes, infrastructure, drinking water, and natural resources, LiDAR data is often incomplete, out-of-date, or not available. For wildfires in particular, the age of mega-fires in is full maturity with unprecedented destruction and costs (Williams 2013). With climate change set to increase aridity across the western US, mega-fires are likely to become the norm throughout the 21st century (Abatzoglou and Williams 2016, Schoennagel et al. 2017), Accurate, up-to-date, and comprehensive datasets are thus needed for effective fire management and could prevent loss of life and property and improve decision-making.

As satellite imagery provides the necessary spatial and temporal coverage and LiDAR the accurate 3D characterization, many studies have harnessed their complementary strengths and fused these data to create comprehensive spatial datasets. LiDAR metrics are used directly as the response variable (e.g. canopy cover and height) or indirectly as the feature of a modeled response (e.g. biomass, basal area, and Lorey's height). Correlations of satellite imagery to basic attributes characterizing vegetation coverage fractions, such as forest canopy cover, are well-documented, especially from the series of Landsat satellites (e.g. Hansen et al. 2013); the ability of satellite imagery to characterize complex forest attributes such as canopy height and height variability (Hudak et al. 2002, Hyde et al. 2006, Pascual et al. 2010, Stojanova et al. 2010, Ahmed et al. 2015, Wilkes et al. 2015, Hansen et al. 2016), basal area (Frazier et al. 2014), stem volume (Huang et al. 2013), biomass (Lefsky et al. 2005, Margolis et al. 2015, Bell et al. 2018), and other measures of structural complexity (Zald et al. 2016, LaRue et al. 2018, Matasci et al. 2018) has come relatively recently. Though the potential for these types of characterizations using imagery has been exploited for several decades (Cohen and Spies 1992), the ubiquity of LiDAR datasets with large numbers of samples and accurate 3-D characterizations have enabled more robust assessments over many forest types. For example, Matasci et al. (2018) showed that six forest structural attributes can be mapped across the entirety of Canada using LiDAR predictor variables with reasonable levels of accuracy (R² 0.49-0.61).

Landsat indices derived Landsat 5 TM, Landsat ETM+, and Landsat 8 OLI have proven important for characterizing forest structure. Tasseled cap indices have been used since the availability of Landsat MSS to characterize vegetation (Kauth and Thomas 1976). Tasseled cap transformations reduce the six spectral Landsat bands to three indices while maintaining most of the information pertinent to vegetation characterization, and they form a discrete, orthogonal separation in spectral space. Each sensor's transformations have also been tweaked to account for differences among Landsat sensors for continuity for the

entirety of the spatial library (Huang et al. 2002, Baig et al. 2014). Many studies have found the tassel cap indices to be among the most important for predictive modeling of forest structure from LiDAR training data (Pascual et al. 2010, Zald et al. 2016, Matasci et al. 2018). Normalized difference vegetation index (Rouse et al. 1974) and normalized burn ratio (Key and Benson 2003) are also two popular indices for characterizing vegetation. Both have shown to be sensitive to wildfire-caused vegetation change (Escuin et al. 2008) and have shown importance to characterizing forest canopy fuels (Erdody and Moskal 2010).

The rise of machine learning techniques has also contributed to the increase in correlative and predictive power for remote sensing of forests (Lawrence and Moran 2015). Mousivand et al. (2014) showed that the sum of second-order interactions was larger than first-order canopy effects on spectral reflectance at the scales and wavelengths of Landsat imagery. Machine learning algorithms have the ability to characterize these complex, non-linear relationships. Indeed, the success of Matasci et al. (2018) for country-wide mapping would not have been possible without the random forest (RF) algorithm (Brieman 2001). Of note, their approach followed methods presented in Chapter 2, where a proximity or distance metric characterizing the similarity of samples was derived from RF models. Machine learning algorithms have been developing at a rapid pace and new models and processing environments are available, which may further increase the ability to characterize forest attributes and their relationship to satellite spectral data.

Gradient boosting machines (GBM) are one of these ensemble algorithms now applied in ecological sciences (Elith et al. 2008). They combine the advantages of tree-based algorithms such as RF, which relate a response to predictor variables using recursive binary splits, with a boosting approach that adaptively combines many simple regression trees (Friedman 2001, Hastie et al. 2001). The major difference of GBM from RF is the sequential or 'stagewise' addition of trees rather than averaging the predictions of a set of independent trees. For classification problems, boosting focuses subsequent trees on error from previous trees by recursively weighting observations that are incorrectly predicted and for regression problems, focuses subsequent trees on the residuals (variation not yet explained in the model) from the previous tree. In this way, each subsequent tree attempts to correct its predecessor. Each tree minimizes a loss function (e.g. deviance) to the extent possible; each step moving down a loss gradient (gradient descent).

The major drawback of GBM is the tendency to overfit (Elith et al. 2008, Natekin and Knoll 2013). Indeed, enough trees can be added to completely overfit a dataset. Model generality must therefore be prioritized. LiDAR datasets are also rather arbitrarily collected representing the diverse needs of contributing stakeholders. At best, cohesive collection

campaigns have been conducted at the state level and the 3D Elevation Program (3DEP) of the USGS (www.usgs.gov/core-science-systems/ngp/3dep) is now planning acquisitions cohesively for more systematic coverage. Thus, with the sporadic spatial extents and differing years of acquisition, many forest ecosystems may not have any direct training data available, and a predictive model must be able to accurately provide estimates in these situations. A host of regularization techniques for GBM to improve generality have been developed (Hepp et al. 2016). Another opportunity for GBM is the ability to rapidly update layers following disturbance such as wildfire, logging, and insect outbreaks. While imputation approaches such as those employed by Matasci et al. (2018,) may maintain the covariance structure of several response variables leading to improved accuracy overall, they will not be able to predict new response covariance assemblages that may be produced by disturbances not captured in the training data. GBM and similar approaches are expected to benefit systematic mapping efforts such as LANDFIRE which currently devotes enormous resources to updating fuels data as disturbances occur.

Generalized models capable of reasonable accuracy across many landscapes has been a research priority for remote sensing and forestry applications for decades (Lefski et al. 2002, Bouvier et al. 2015, Bell et al. 2018). The inability of models derived from localized datasets to apply to other landscapes stems from three primary and related issues: model overfitting (Bouvier et al. 2015), mis-specified predictor variables not fully characterizing the response, and spatial variance in the feature-response relationships, in this case the LiDAR-Landsat relationships (Matasci et al. 2018). Spatial variance in feature-response relationships is a well-known problem but differing acquisition years among LiDAR datasets adds the additional complication of potential temporal variance in the relationship as well.

Hyperparameter tuning is one regularization approach that can increase model generality, but can still lead to overfitting if not properly applied. Hyperparameter tuning is a Monte Carlo approach where model parameters are varied and accuracy assessments (usually on the validation dataset) determine the optimal set of parameter values for a final model (Bergstra and Bengio 2012). For GBM, important parameters to tune for generality are the proportion of samples evaluated in each tree (bagging), number of predictor variables assessed at each binary split, learn rate annealing, stopping tolerances, maximum runtimes, and the series of tree pruning parameters such as max tree depth, minimum accuracy improvement for each split, and minimum number of samples in each terminal node (Natekin and Knoll 2013).

Another important approach to reduce overfitting is sample weighting (Torgo et al. 2013, Branco et al. 2017). Forest landscapes can display sample distributions that are

heavily skewed or have high kurtosis in one or more canopy variables. A machine learning model trained on a landscape with these characteristics will be biased as the model attempts to minimize a single error metric such as RMSE. The most extreme example would be a model predicting only the sample mean value, which could be quite accurate in cases of extreme kurtosis or skewness in the sample distribution. Analogous to class balancing in classification tasks, several approaches have been developed to address this issue (Branco et al. 2017). The simplest and most common are pre-processing steps to change the original data distribution such as over- and under-sampling certain observations. This approach can reduce overall model performance when assessed by single error metrics, such as RMSE, but generality would improve leading to improvements in the ability to assess disturbance or other changes to canopy fuels in landscapes without many examples of such disturbances in the training data.

For large-area predictions, spatial variance in the LiDAR-Landsat relationships also lead to decreased model generality if there is insufficient variation in training data (Lefsky et al. 2002, Bell et al. 2015, Bell et al. 2018). One solution is to create many localized models from each individual LiDAR dataset and stitch the predictions together. Determining boundaries where each local model is most applicable would be difficult however. Conversely, a single global model can be used with either sufficient generality or including locational features to correct for spatial variance as in Zald et al. (2016) and Matasci et al. (2018). A few models trained on data stratified using biophysical rationale may represent a compromise between these two extremes.

LANDFIRE is one such dataset originally derived from field-based inventories that can be significantly improved with the addition of LiDAR training data (Peterson et al. 2015). As canopy information is directly and accurately measured with LiDAR, updates to canopy fuel estimations are a logical starting point. For operational fire modeling, canopy fuels dictate the threshold surface fire intensity necessary to create crown fire, control the wind reduction factor, the speed of active crown fire spread, and the distance and amount of spot fires, all of which in turn can induce non-linear increases in fire intensity and spread (Cruz et al. 2003). Accurate canopy fuel layers are thus important for the characterization of large and intense wildfires which account for the majority of fatalities, structure lost, and acres burned (Williams 2013).

Canopy cover and height can be directly measured by LiDAR while the other two canopy variables, canopy base height (CBH) and canopy bulk density (CBD), can be estimated indirectly using established equations (Peterson et al. 2015). Though not direct, these LiDAR-based conceptualizations of CBH and CBD capture the spatial variation in 3D canopy structure important for fire modeling. While different in their conception from field-

based estimates, the absolute values of these variables are not as important as representing the structural variability present across the landscape because fire behavior modelers will often need to calibrate fuels data to match observed fire behavior (e.g. reduce all CBH by half or double CBD) (Stratton et al. 2009, Cochrane et al. 2012).

In this study, canopy fuels are estimated from LiDAR datasets and then related to Landsat spectral data in order to create models with sufficient generality to predict over the entirety of the western US. Model predictions are then be compared to current LANDFIRE data to quantify improvements due to LiDAR training data and GBM modeling. The research has five objectives:

1) Describe relationships between Landsat-derived spectral indices and LiDAR-derived canopy cover, stand height, crown base height, and canopy bulk density
2) Characterize the variance of the LiDAR-Landsat relationships over a diversity of landscapes
3) Create predictive model(s) that can be applied to any area in the western US
4) Compare model predictions to current LANDFIRE products
5) Assess the selected model's ability to update canopy layers following wildfire disturbance

2) Methods

2.1 Data

LiDAR datasets were selected based on availability and to represent the diversity of conifer forests, climates, and disturbance regimes in the western contiguous US (Fig. 1). LiDAR acquisitions were ultimately grouped into thirteen landscapes for analysis based on proximity and vegetation similarities. All LiDAR data acquisition parameters followed at a minimum the requirements for the US Geological Survey's Quality Level 1 (Heidemann 2018). Important collection parameters include ≥ 3 returns/pulse, ≥ 8 returns/m^2, and a relative vertical accuracy ≤ 0.06 m RMSD. Landscape sizes and descriptions are in Figure 1. LiDAR data totaled 1,258,993 ha for training and validation and 265,225 ha for testing. Elevations range from 0 to 3599 m a.s.l. and precipitation normal range from less than 200 mm to over 4000 mm annually (PRISM 2019). At least one dataset is within each EPA Level I forest ecoregion of the western US (Fig. 1).
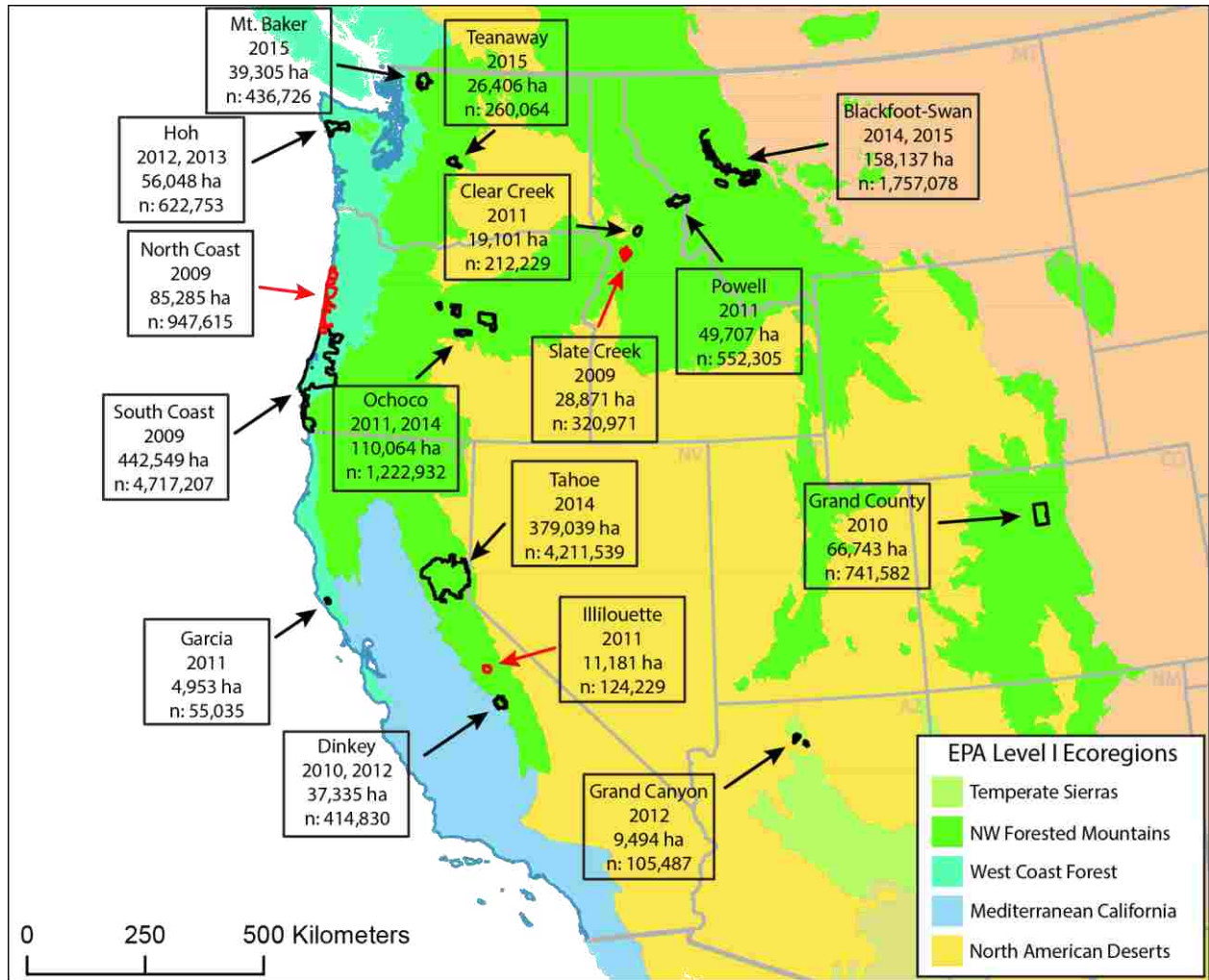
Figure 1. Map of western US and LiDAR datasets used in the study with year of acquisition, area of actual forest data used after filtering (see Methods), and number of samples in each dataset. Black perimeters show landscapes where training, validation, and testing data was used while red perimeters show landscapes used exclusively for testing (North Coast, Illilouette, and Slate Creek).

Table 1. Summary of canopy fuel response variables (LiDAR-sourced) and predictor variables used in the study.

Source	Variable Name	Description	Citation
LiDAR	Canopy Cover (CC) (%)	Percentage of first returns above 2 m	Hopkinson and Chasmer 2009; Smith et al. 2009; Peterson et al. 2015
	Canopy Height (CH) (m)	99th percentile return height	Peterson et al. 2015
	Canopy Base Height (CBH) (m)	Mean return height minus standard deviation of heights	Rowell et al. 2005; Peterson et al. 2015
	Canopy Bulk Density (CBD) (kg/m³)	$e\textasciicircum(-2.489 + 0.034\,(CC) - 0.357(SH1) - 0.601(SH2) - 1.107(PJ) - 0.001(CC*SH1) - 0.002(CC*SH2))$ If CH is 0-15m, SH1=0 and SH2=0 If CH is 15-30m, SH1=1 and SH2=0 If CH is 30-91m, SH1=0 and SH2=1 If EVT equals Pinyon or Juniper type, PJ =1 else PJ=0	Reeves et al. 2009
Landsat	Med NDVI	Median normalized difference vegetation index (NDVI) value	Rouse et al. 1974
	Max NDVI	Maximum NDVI value	Rouse et al. 1974
	Med NBR	Median normalized burn ratio (NBR)	Key and Benson 2002
	Max NBR	Maximum NBR	Key and Benson 2002
	Med Bright	Median Tasseled Cap Brightness	Crist 1985; Huang et al. 2002; Baig et al. 2014
	Max Bright	Maximum Tasseled Cap Brightness	Crist 1985; Huang et al. 2002; Baig et al. 2014
	Med Green	Median Tasseled Cap Greenness	Crist 1985; Huang et al. 2002; Baig et al. 2014
	Max Green	Maximum Tasseled Cap Greenness	Crist 1985; Huang et al. 2002; Baig et al. 2014
	Med Wet	Median Tasseled Cap Wetness	Crist 1985; Huang et al. 2002; Baig et al. 2014
	Max Wet	Maximum Tasseled Cap Wetness	Crist 1985; Huang et al. 2002; Baig et al. 2014
LANDFIRE	EVT	Existing vegetation type	Rollins et al. 2009
	FRG	Fire regime group	Rollins et al. 2007
	Slope (%)	Slope	
	Aspect (deg)	Aspect	
	Elev (m)	Elevation	

2.2 LiDAR Processing

LiDAR point clouds were processed through FUSION software (McGaughey 2015) to extract height above ground and calculate metrics to match the 30 m Landsat cell resolution. Bilinear interpolation aligned the Landsat and LiDAR rasters. Table 1 shows the formulation of the canopy fuel metrics from LiDAR. Canopy height (CH), canopy cover (CC), and canopy base height (CBH) are defined directly from the LiDAR metrics while canopy bulk density uses a combination of canopy cover and canopy height from the LiDAR data and existing vegetation type from LANDFIRE (Reeves et al. 2009). LiDAR data were filtered for a minimum canopy height of 2 m, canopy cover of 2%, and an existing LANDFIRE vegetation type of forest to describing a forest to ensure that non-forested pixels did not contaminate samples.

2.3 Landsat and LANDFIRE Data Processing

Landsat indices (Table 1) were calculated for the contiguous US for 2000-2016 using May to October imagery. Median and maximum values of each index for each year were calculated using US Geological Survey Landsat 5 TM, Landsat 7 ETM+, and Landsat 8 OLI surface reflectance products created using the Landsat Ecosystem Disturbance Adaptive Processing (LEDAPS) algorithm (Masek et al. 2006, Schmidt et al. 2013). Landsat TM and ETM+ were merged and constituted the data source for 2000-2014, and OLI data was used for 2014-2016. Equations for index calculations are within the citations in Table 1. Pixels were filtered for contamination by clouds, cloud shadows, adjacency to clouds, snow, and water. All Landsat data was processed within Google Earth Engine, exported, and then mosaicked into multi-band rasters organized by year.

Topography metrics of slope, aspect, and elevation, were taken from LANDFIRE (Rollins 2009). Fire regime groups (FRG) (Rollins et al. 2007) and existing vegetation type (EVT) were also extracted along with LANDFIRE's existing canopy fuel layers CC, CH, CBH, and CBD (Reeves et al. 2009) for comparison to model outputs.

2.4 Dataset Stratification, Sample Weighting, and Model Development

All subsequent data processing and model development was completed using R statistical software as a wrapper to the Apache Spark analytics engine. Spark is an open-source, parallel, scalable, and resilient Big Data processing environment (Zaharia et al. 2016). H2O machine learning algorithms (www.h2o.ai) were employed within R and SPARK using the R packages h2o (LeDell et al. 2019), sparklyr (Luraschi et al. 2019), and rsparkling (Hava et al. 2019). Data preprocessing, stratification, and weighting was

47

completed using the R packages raster (Hijmans et al. 2019), ggplot2 (Wickham et al. 2019), and dplyr (Wickham et al. 2019).

Three data stratification approaches were devised leading to three sets of models. Local models, models stratified by fire regime groups (FRG), and a global model taking data from all landscapes. For local models, training, validation, and test data were taken from within each individual landscape and a separate model derived for each. This represents a traditional modeling approach ubiquitous in remote sensing literature. Recognizing these local models are likely overfit to their landscapes, they represent the baseline accuracy or maximum potential extractable information that subsequent, more generalized models can be compared to. Next, I hypothesized that the nature of the spectral relationships may vary across forest types and separate models may better characterize these relationships. Fire regime groups produced by LANDFIRE represent an integration of existing and potential vegetation, climate, topography, and disturbance regimes (Rollins et al. 2007). For forested areas, four fire regime groups within the landscapes were present:

FRG 1: ≤35 yr return interval, low and mixed severity
FRG 3: 35-200 yr return interval, low and mixed severity
FRG 4: 35-200 yr return interval, replacement severity
FRG 5: > 200 yr return interval, any severity

FRG 5 represents a diverse category including multiple vegetation types across the US but in the western US and for the landscapes in this study, FRG 5 represents the western Cascade Mountains and coastal forests in the Mt. Baker, Hoh, North Coast, and South Coast landscapes. All pixels representing a particular FRG from all landscapes were binned and used to construct a model.

Finally, a global model containing all data from all landscapes was created as a parsimonious option requiring no stratification. However, with the significant differences in dataset sizes between landscapes (cf. Fig. 1), predictor-response relationships in large datasets may dominate the smaller datasets. Thus, landscape datasets were randomly sampled so that each landscape's contribution did not exceed 20% of the total sample size for the global model. This rule was also applied to the FRG models for the same purpose, but the proportion was adjusted to 30% for FRG 4 and FRG 5 because only a few landscapes contained enough data.

An additional sample weighting scheme for every model was applied to improve generality and the ability to characterize disturbance on any landscape. For each of the four canopy response variables (CC, CH, CBH, CBD), the range of values was assessed and then split into ten even-sized bins (e.g. CC: 1-10%, 11-20%, 21-30%, etc.). The bins were

ordered by sample size, and the samples within the bin with the highest count all received a weighting value of one (i.e. would be considered once in the model). Each other bin's samples were then multiplied by the necessary weighting value so that the sum of the sample weights were equal across all bins. This effectively gave an equal weighting to the full range of values, helping to prevent model bias due to skewed or leptokurtotic response distributions. Several classes for certain models had very few samples within a class which led to extreme weighting values to few samples (> 100,000x weight in one case). Therefore, the maximum weight for any sample was set to 100x to reduce model instability created by extreme weighting.

Once stratification and weighting were completed, model development was then possible. From each model's sample pool, 80% were selected for training, 10% for validation and internal error estimation during training, and 10% for final testing and error estimation. This form of testing is a largely accepted for error estimation in research, but does not represent a truly independent dataset. The testing data's acquisition parameters, especially the year of acquisition, and the geographic location and vegetation type are still related to the particular landscape from which they were derived. Thus, three separate LiDAR datasets without any data used in the training process were used as a final assessment of model performance.

Table 2 shows the GBM model parameters including those varied during hyperparameter tuning. For more explanation of each parameter see H2O.AI GBM documentation (http://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/gbm.html). A maximum of ten models were trained before selecting each final model, and the various parameter options were randomly selected for each model run. Final model selection used minimum RMSE as the selection metric.

Table 2. Model parameters. Those with multiple values were varied in the hyperparameter tuning.

GBM Model Parameters	Value(s)
ntrees	Up to 4000
learn_rate	0.1
learn_rate_annealing	0.01
max_runtime_secs	750
sample_rate	0.4, 0.6, 0.9, 1
col_sample_rate	0.6, 0.9, 1
col_sample_rate_per_tree	0.6, 0.9, 1
col_sample_rate_change_per_level	0.01, 0.9, 1.1
nbins	32, 64, 128, 256
min_split_improvement	0, 1e-4, 1e-6, 1e-8
max_depth	20, 30, 40
histogram_type	AUTO, UniformAdaptive, QuantilesGlobal
stopping_metric	RMSE
stopping_tolerance	0.01
Score_tree_interval	10
Stopping_rounds	3

2.5 Spectral Response and Accuracy Assessment

In order to assess the stationarity of the spectral features to the canopy fuel variables, partial dependence plots were calculated for the local, FRG, and global models. Partial dependence plots assess the marginal effect of a single feature on the response variable (Friedman 2001). The major assumption is that features are independent of each other. While the GBM model is resistant to multicollinearity, if a feature is highly correlated with another, then data combinations created from the feature distributions can be highly unlikely in reality and therefore bias partial dependence estimates. Thus, local, FRG, and global models were first trained on a reduced subset of features for creation of partial dependence plots. The reduced set included only the median spectral index values as the median and maximum indices were highly correlated for most features (r=0.7-0.9).

For accuracy assessment and local-FRG-global model comparisons, models were then trained on the full suite of features. Root mean squared error (RMSE), mean absolute error (MAE), and r-squared (R^2) were used as model evaluation metrics. For the independent datasets (Illilouette, North Coast, and Slate Creek), the nearest local models were used for comparison to the FRG and global model predictions. The independent dataset model outputs were also compared to existing LANDFIRE canopy metrics.

Model predictions should be able to account for disturbance effects on canopy fuels accurately, especially wildfire effects. Within the spatial extents of the LiDAR datasets, I searched for wildfires that occurred after the LiDAR acquisition date but before 2016 so post-fire imagery the following growing season could be attained. Spectral indices for one year before the fire and one year after the fire were calculated. Burn severity maps from the Monitoring Trends in Burn Severity (MTBS; mtbs.gov) were acquired and samples stratified by high, moderate, and unburned to low severity classes. The global model and local model pre-fire and post-fire predictions were compared to evaluate whether predictions followed post-fire expectations of change. CC, CH, CBH, and CBD were all hypothesized to decrease the most in the high severity classes with the unburned to low class having little change and the moderate severity class in-between.

3) Results

The partial dependence plots show complex and non-linear feature-response relationships. The spectral features exhibited relatively consistent response shapes but shifted in mean response values among landscapes (Fig. 2 and Fig. 3). For example, for the NDVI-CC relationship (top left, Fig. 2), the mean response for a median NDVI value of 0.6 is ~40% CC for the Ochoco landscape and ~80% for the Garcia landscape. CBH and CBD had more inconsistent response shapes between landscapes overall compared to CC and CH. For example for the TC Brightness-CBD relationship (third row, second column, Fig. 3), the Mt. Baker, Garcia, and Hoh landscapes all show an increase in CBD as TC Brightness increased while the other landscapes showed a mean decrease. Present in most feature-response relationships but especially noticeable for CBH and CBD, Mt. Baker, Clear Creek, Garcia, Hoh, and South Coast landscapes all had larger canopy fuel values across the distribution of feature values while the Ochoco and Grand County landscapes were consistently on the lower end. The partial dependence plots derived from the FRG models (Fig. 4 and Fig. 5) characterize this same trend with FRG 5 (primarily derived from the Mt. Baker, Hoh and South Coast landscapes), consistently having higher canopy fuel values and FRG 4 (present in many upper elevation, inland landscapes) showing lower canopy fuel values. FRG 1 and FRG 3 follow similar trends to each other and the global dataset overall.

Figure 2. Partial dependence plots for the local landscape and global gradient boosting machine (GBM) models for the canopy cover and canopy height response variables.

Figure 3. Partial dependence plots for the local landscape and global gradient boosting machine (GBM) models for the canopy base height and canopy bulk density response variables.
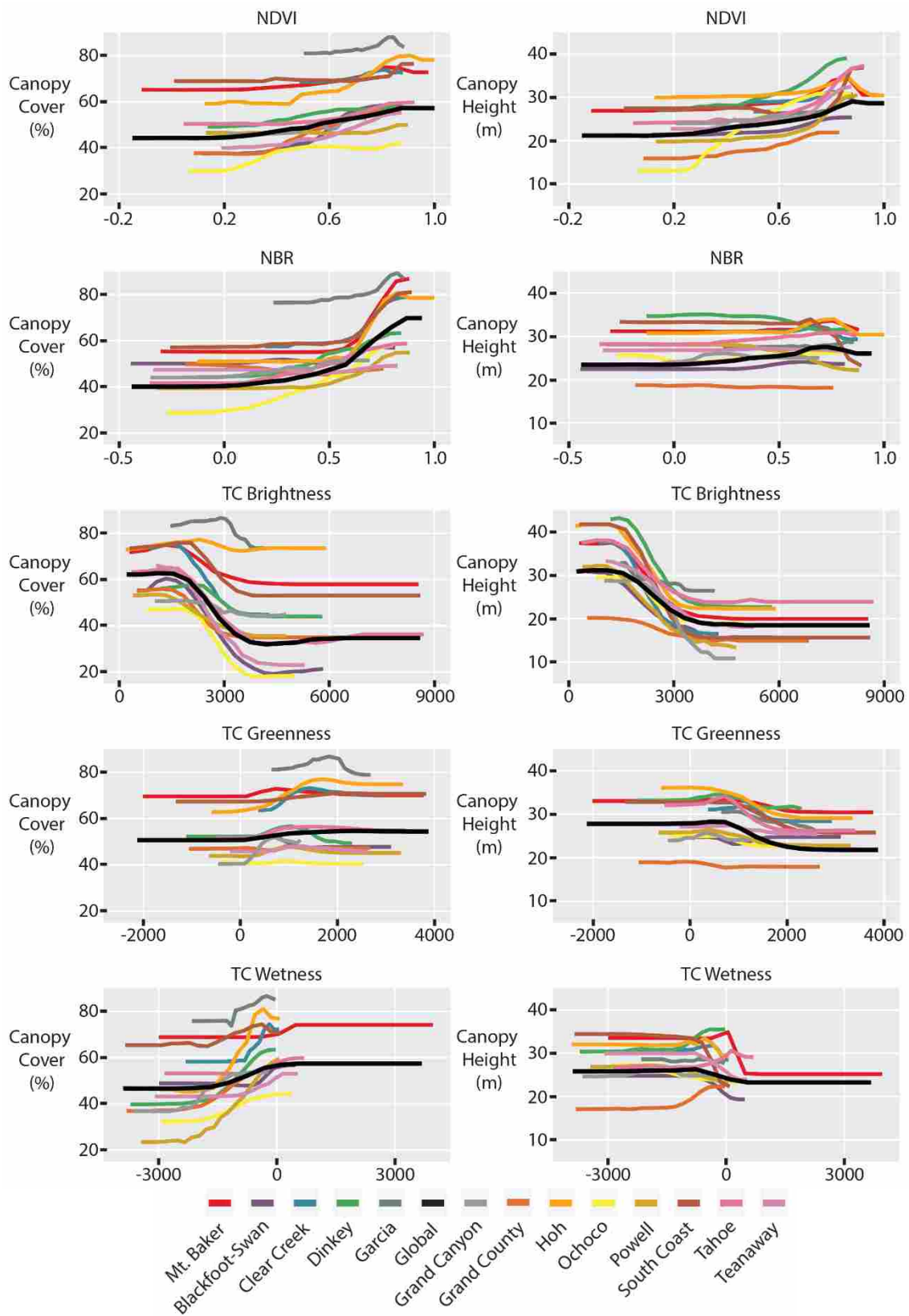
53

Figure 4. Partial dependence plots for the fire regime group (FRG) and global gradient boosting machine (GBM) models for the canopy cover and canopy height response variables.
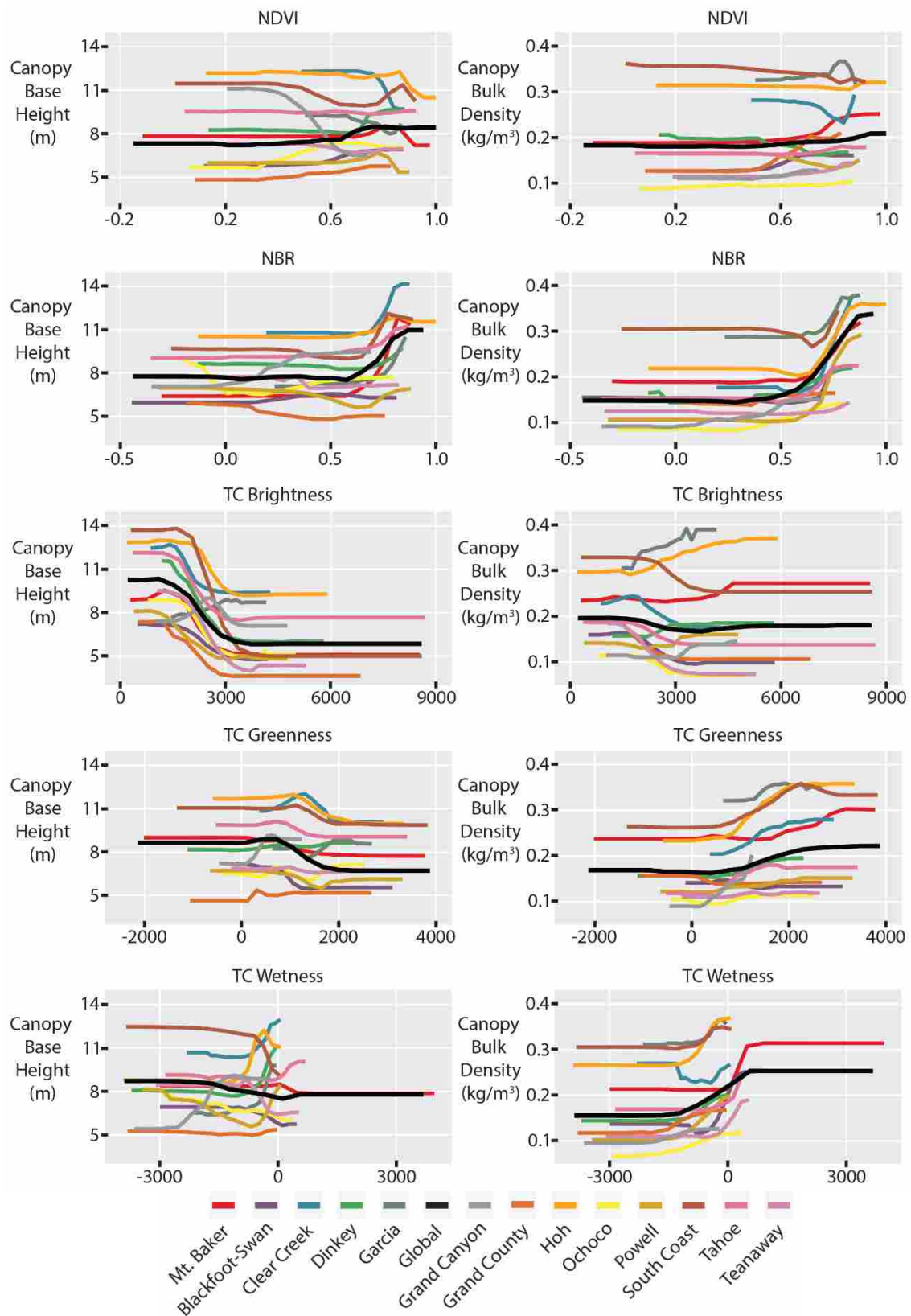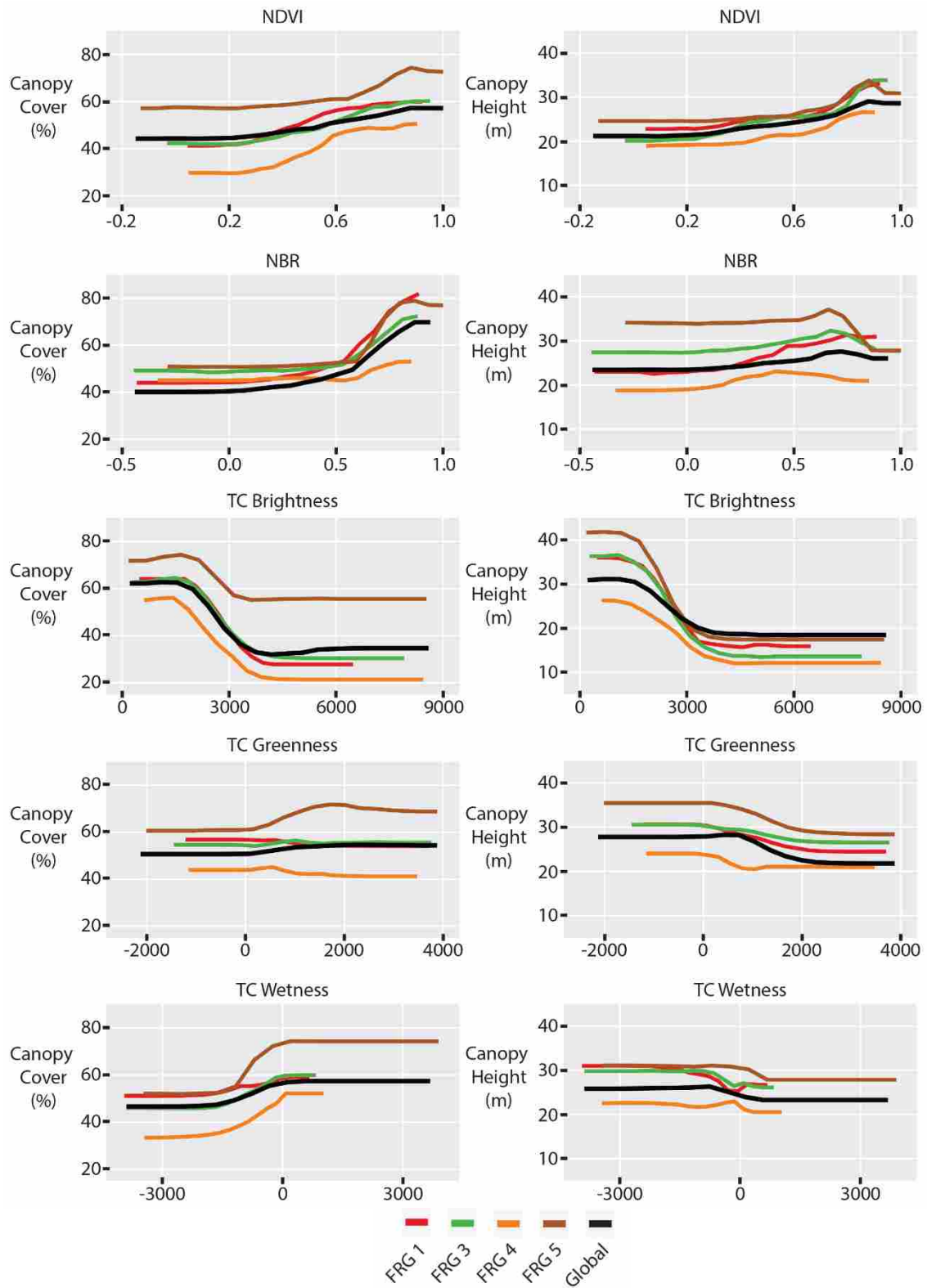
54
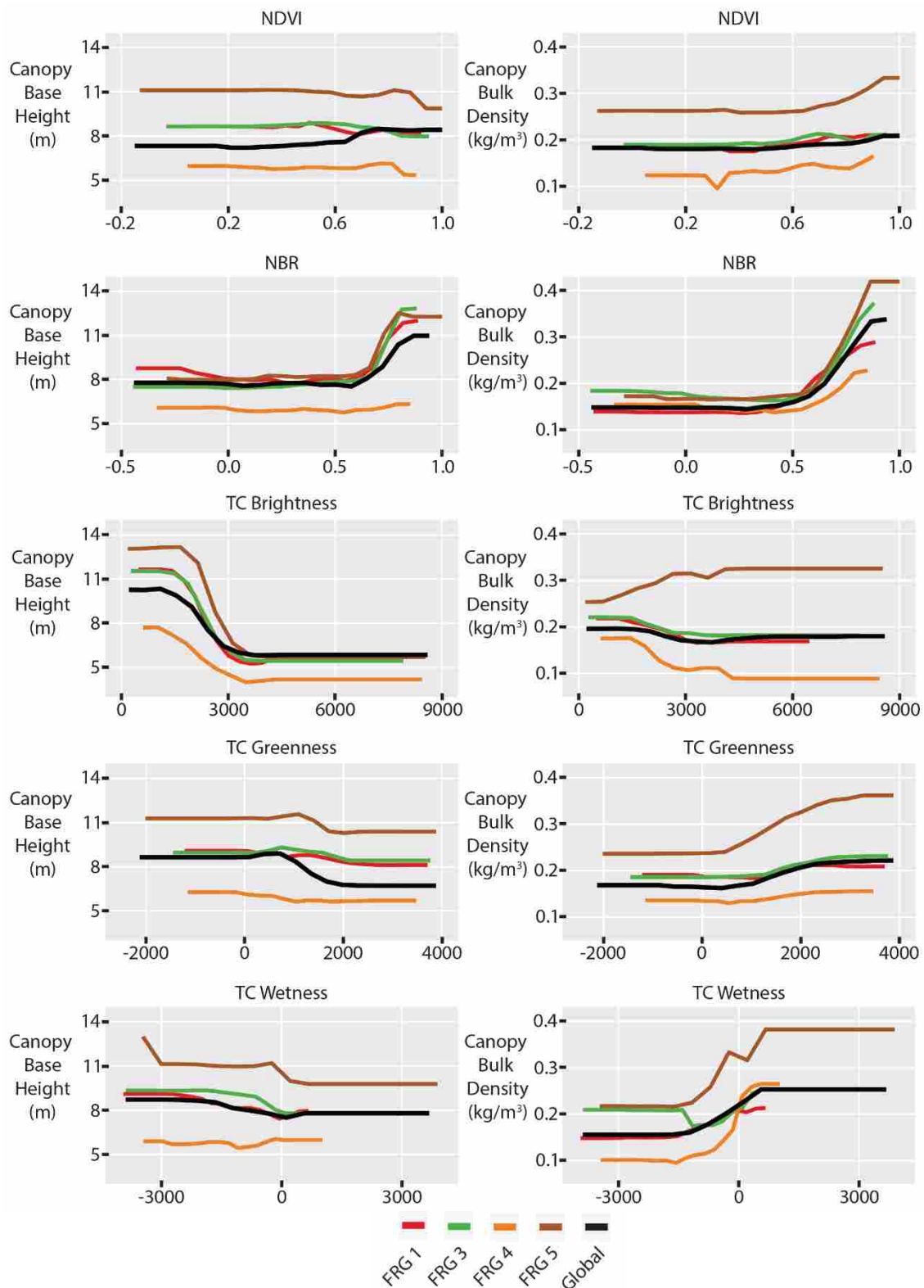
Figure 5. Partial dependence plots for the fire regime group (FRG) and global gradient boosting machine (GBM) models for the canopy base height and canopy bulk density response variables.

The global model compared favorably to the local models in the error assessments (Table 3). The complete breakdown of each landscape including FRG model accuracies is shown in Table 4. The following mean change in error metrics are averages of local-global comparisons with equal weighting given to each landscape regardless of size. For CC, the use of the global model increased error by 0.08 % and 0.11 % for RMSE and MAE, respectively, and increased R^2 by 0.004. R^2 decreased slightly with the global model for all landscapes, except for the Garcia and Hoh landscapes (Table 4). For CH, use of the global model increased RMSE by 0.4 m, MAE by 0.32 m, and decreased R^2 by 0.041. For CBH, the global model increased RMSE by 0.17 m, MAE by 0.15 m, and decreased R^2 by 0.053. For CBD, the global model increased RMSE and MAE each by 0.002 kg/m^3 and decreased R^2 by 0.02.

Table 3. Comparison of local models to the global model. Each landscape weighted equally in calculation of the mean accuracy metrics (i.e. not area-weighted).

Accuracy	CC (%)	CH (m)	CBH (m)	CBD (kg/m^3)
Local RMSE	10.02	4.91	2.33	0.057
Global RMSE	10.10	5.31	2.50	0.059
Local MAE	7.42	3.67	1.73	0.041
Global MAE	7.53	3.99	1.88	0.043
Local R^2	0.725	0.672	0.600	0.622
Global R^2	0.729	0.631	0.547	0.602

The error comparisons of individual landscapes to the global model predictions largely follow this same trend of slight increases in error with global model use (Table 4). The independent test landscapes flip this trend with the global modeling performing better in 9 of 12 (75%) landscape-response variable combinations (Table 5). The starkest difference is in the Slate Creek landscape for CBD where the R^2 using the nearest local model (Clear Creek) was -0.889 and use of the global model increased the R^2 to 0.479. The South Coast model performed better than the global model in the North Coast landscape for CH, CBH, and CBD. The FRG models also performed better than the nearest local models in most comparisons, on par or near to global model performance. The performance of the global models on the test landscapes is further highlighted in the predicted versus observed graphs (Figs. 6, 7, and 8).

The predicted versus observed graphs also show the increase in performance compared to the existing LANDFIRE data. The global model reduced RMSE on average by 11.3 % for CC, 5.45 m for CH, 5.78 m for CBH, and 0.062 kg/m^3 for CBD compared to LANDFIRE results. The LANDFIRE data had little ability to characterize canopy fuels in general with negative R^2 values for 8 of 12 (66.7%) LANDFIRE-response variable combinations. For those with positive values, LANDFIRE had a mean R^2 of 0.183. The global model performed worse on the test landscapes compared to the training landscapes,

but produced a mean R² of 0.439 overall compared to a mean R² of -1.375 for LANDFIRE. The global model performed comparably to the training landscapes for the Illilouette and North Coast landscapes except for Illilouette-CBH which had a -0.094 R². For Slate Creek, R² values were noticeably worse than training landscapes with a mean of 0.372.

Table 4. Error assessment of local, global, and fire regime group (FRG) models for landscapes used in model training. 80% of landscape data used for training, 10% for validation, and 10% for testing. N refers to the number of samples used for the performance metrics.

Landscape	Model	Canopy Fuel Variable	N	RMSE	MAE	R²
Mt. Baker	Local	CC	43636	9.34%	6.63%	0.862
Mt. Baker	Global	CC	43636	9.84%	7.27%	0.846
Mt. Baker	FRG 5	CC	42840	9.79%	7.16%	0.844
Mt. Baker	Local	CH	43636	5.36 m	3.86 m	0.83
Mt. Baker	Global	CH	43636	5.96 m	4.33 m	0.747
Mt. Baker	FRG 5	CH	42840	5.44 m	3.97 m	0.822
Mt. Baker	Local	CBH	43636	2.16 m	1.57 m	0.79
Mt. Baker	Global	CBH	43636	2.35 m	1.74 m	0.747
Mt. Baker	FRG 5	CBH	42840	2.29 m	1.68 m	0.759
Mt. Baker	Local	CBD	43636	0.057 kg/m^3	0.040 kg/m^3	0.8
Mt. Baker	Global	CBD	43636	0.063 kg/m^3	0.045 kg/m^3	0.749
Mt. Baker	FRG 5	CBD	42840	0.060 kg/m^3	0.043 kg/m^3	0.77
Blackfoot-Swan	Local	CC	175639	8.11%	5.93%	0.839
Blackfoot-Swan	Global	CC	175639	8.55%	6.36%	0.822
Blackfoot-Swan	FRG 1	CC	85805	8.19%	6.08%	0.819
Blackfoot-Swan	FRG 3	CC	45767	8.69%	6.35%	0.84
Blackfoot-Swan	FRG 4	CC	41856	8.50%	6.17%	0.82
Blackfoot-Swan	Local	CH	175639	3.37 m	2.47 m	0.757
Blackfoot-Swan	Global	CH	175639	3.84 m	2.85 m	0.686
Blackfoot-Swan	FRG 1	CH	85805	3.65 m	2.72 m	0.662
Blackfoot-Swan	FRG 3	CH	45767	3.89 m	2.87 m	0.732
Blackfoot-Swan	FRG 4	CH	41856	3.70 m	2.74 m	0.726
Blackfoot-Swan	Local	CBH	175639	1.54 m	1.13 m	0.645
Blackfoot-Swan	Global	CBH	175639	1.68 m	1.25 m	0.586
Blackfoot-Swan	FRG 1	CBH	85805	1.64 m	1.23 m	0.588
Blackfoot-Swan	FRG 3	CBH	45767	1.73 m	1.28 m	0.597
Blackfoot-Swan	FRG 4	CBH	41856	1.51 m	1.11 m	0.655
Blackfoot-Swan	Local	CBD	175639	0.049 kg/m^3	0.032 kg/m^3	0.723
Blackfoot-Swan	Global	CBD	175639	0.047 kg/m^3	0.034 kg/m^3	0.712
Blackfoot-Swan	FRG 1	CBD	85805	0.045 kg/m^3	0.032 kg/m^3	0.702
Blackfoot-Swan	FRG 3	CBD	45767	0.048 kg/m^3	0.035 kg/m^3	0.736
Blackfoot-Swan	FRG 4	CBD	41856	0.047 kg/m^3	0.034 kg/m^3	0.712
Clear Creek	Local	CC	21060	10.24%	7.71%	0.72
Clear Creek	Global	CC	21060	10.44%	8.00%	0.718

Clear Creek	FRG 3	CC	17122	9.91%	7.35%	0.732
Clear Creek	Local	CH	21060	4.84 m	3.56 m	0.788
Clear Creek	Global	CH	21060	5.43 m	4.03 m	0.734
Clear Creek	FRG 3	CH	17122	5.17 m	3.80 m	0.768
Clear Creek	Local	CBH	21060	2.67 m	1.93 m	0.628
Clear Creek	Global	CBH	21060	2.85 m	2.11 m	0.586
Clear Creek	FRG 3	CBH	17122	2.87 m	2.15 m	0.589
Clear Creek	Local	CBD	21060	0.065 kg/m^3	0.049 kg/m^3	0.631
Clear Creek	Global	CBD	21060	0.069 kg/m^3	0.052 kg/m^3	0.598
Clear Creek	FRG 3	CBD	17122	0.069 kg/m^3	0.053 kg/m^3	0.577
Dinkey	Local	CC	41443	10.42%	7.91%	0.748
Dinkey	Global	CC	41443	10.60%	8.19%	0.739
Dinkey	FRG 1	CC	33790	10.43%	7.99%	0.73
Dinkey	FRG 3	CC	6989	10.17%	7.75%	0.765
Dinkey	Local	CH	41443	6.60 m	5.05 m	0.702
Dinkey	Global	CH	6.88	6.88 m	5.31 m	0.675
Dinkey	FRG 1	CH	33790	7.31 m	5.67 m	0.64
Dinkey	FRG 3	CH	6989	6.80 m	5.23 m	0.563
Dinkey	Local	CBH	41443	2.60 m	1.95 m	0.551
Dinkey	Global	CBH	41443	2.73 m	2.07 m	0.508
Dinkey	FRG 1	CBH	33790	2.75 m	2.07 m	0.51
Dinkey	FRG 3	CBH	6989	2.86 m	2.27 m	0.355
Dinkey	Local	CBD	41443	0.057 kg/m^3	0.038 kg/m^3	0.666
Dinkey	Global	CBD	41443	0.057 kg/m^3	0.039 kg/m^3	0.667
Dinkey	FRG 1	CBD	33790	0.059 kg/m^3	0.041 kg/m^3	0.659
Dinkey	FRG 3	CBD	6989	0.042 kg/m^3	0.026 kg/m^3	0.667
Garcia	Local	CC	5522	8.99%	6.16%	0.216
Garcia	Global	CC	5522	8.56%	6.24%	0.269
Garcia	FRG 1	CC	5493	7.98%	5.30%	0.363
Garcia	Local	CH	5522	4.72 m	3.57 m	0.152
Garcia	Global	CH	41443	5.02 m	3.81 m	0.081
Garcia	FRG 1	CH	5493	5.13 m	3.92 m	0.041
Garcia	Local	CBH	5522	2.29 m	1.78 m	0.423
Garcia	Global	CBH	5522	2.40 m	1.88 m	0.368
Garcia	FRG 1	CBH	5493	2.40 m	1.88 m	0.365
Garcia	Local	CBD	5522	0.081 kg/m^3	0.064 kg/m^3	0.284
Garcia	Global	CBD	41443	0.080 kg/m^3	0.064 kg/m^3	0.305
Garcia	FRG 1	CBD	5493	0.079 kg/m^3	0.062 kg/m^3	0.325
Grand Canyon	Local	CC	10548	7.09%	5.30%	0.738
Grand Canyon	Global	CC	10548	7.32%	5.51%	0.717
Grand Canyon	FRG 1	CC	7629	7.31%	5.50%	0.716
Grand Canyon	FRG 4	CC	2671	6.82%	5.05%	0.725
Grand Canyon	Local	CH	10548	3.09 m	2.32 m	0.663
Grand Canyon	Global	CH	10548	3.13 m	2.37 m	0.673
Grand Canyon	FRG 1	CH	7629	3.23 m	2.45 m	0.674
Grand Canyon	FRG 4	CH	2671	3.13 m	2.38 m	0.487

Grand Canyon	Local	CBH	10548	2.06 m	1.55 m	0.822
Grand Canyon	Global	CBH	10548	2.21 m	1.69 m	0.792
Grand Canyon	FRG 1	CBH	7629	2.37 m	1.82 m	0.759
Grand Canyon	FRG 4	CBH	2671	1.69 m	1.30 m	0.668
Grand Canyon	Local	CBD	10548	0.028 kg/m^3	0.021 kg/m^3	0.556
Grand Canyon	Global	CBD	10548	0.030 kg/m^3	0.021 kg/m^3	0.506
Grand Canyon	FRG 1	CBD	7629	0.030 kg/m^3	0.022 kg/m^3	0.509
Grand Canyon	FRG 4	CBD	2671	0.022 kg/m^3	0.017 kg/m^3	0.654
Grand County	Local	CC	74132	9.87%	7.20%	0.76
Grand County	Global	CC	74132	10.03%	7.43%	0.753
Grand County	FRG 1	CC	14605	12.08%	9.01%	0.696
Grand County	FRG 4	CC	55336	9.49%	6.95%	0.745
Grand County	Local	CH	74132	3.02 m	2.25 m	0.619
Grand County	Global	CH	74132	3.22 m	2.39 m	0.566
Grand County	FRG 1	CH	14605	3.73 m	2.85 m	0.421
Grand County	FRG 4	CH	55336	3.01 m	2.23 m	0.581
Grand County	Local	CBH	74132	1.33 m	0.98 m	0.552
Grand County	Global	CBH	74132	1.48 m	1.08 m	0.448
Grand County	FRG 1	CBH	14605	1.68 m	1.26 m	0.387
Grand County	FRG 4	CBH	55336	1.30 m	0.96 m	0.531
Grand County	Local	CBD	74132	0.043 kg/m^3	0.030 kg/m^3	0.599
Grand County	Global	CBD	74132	0.044 kg/m^3	0.031 kg/m^3	0.587
Grand County	FRG 1	CBD	14605	0.049 kg/m^3	0.033 kg/m^3	0.545
Grand County	FRG 4	CBD	55336	0.042 kg/m^3	0.031 kg/m^3	0.585
Hoh	Local	CC	62288	12.93%	9.71%	0.61
Hoh	Global	CC	62288	11.23%	8.03%	0.712
Hoh	FRG 5	CC	62016	11.50%	8.35%	0.697
Hoh	Local	CH	62288	5.75 m	4.13 m	0.871
Hoh	Global	CH	62288	6.42 m	4.68 m	0.841
Hoh	FRG 5	CH	62016	5.89 m	4.27 m	0.865
Hoh	Local	CBH	62288	3.28 m	2.45 m	0.67
Hoh	Global	CBH	62288	3.78 m	2.87 m	0.564
Hoh	FRG 5	CBH	62016	3.73 m	2.84 m	0.575
Hoh	Local	CBD	62288	0.080 kg/m^3	0.059 kg/m^3	0.644
Hoh	Global	CBD	62288	0.084 kg/m^3	0.062 kg/m^3	0.607
Hoh	FRG 5	CBD	62016	0.083 kg/m^3	0.062 kg/m^3	0.617
Ochoco	Local	CC	122339	8.52%	6.46%	0.79
Ochoco	Global	CC	122339	8.60%	6.60%	0.787
Ochoco	FRG 1	CC	93212	8.53%	6.49%	0.782
Ochoco	FRG 3	CC	22380	8.53%	6.50%	0.778
Ochoco	Local	CH	122339	4.86 m	3.73 m	0.646
Ochoco	Global	CH	122339	5.08 m	3.93 m	0.615
Ochoco	FRG 1	CH	93212	5.13 m	4.00 m	0.564
Ochoco	FRG 3	CH	22380	4.68 m	3.53 m	0.704
Ochoco	Local	CBH	122339	1.91 m	1.41 m	0.501
Ochoco	Global	CBH	122339	1.95 m	1.44 m	0.47

Ochoco	FRG 1	CBH	93212	2.01 m	1.50 m	0.448
Ochoco	FRG 3	CBH	22380	1.78 m	1.27 m	0.53
Ochoco	Local	CBD	122339	0.031 kg/m^3	0.022 kg/m^3	0.585
Ochoco	Global	CBD	122339	0.030 kg/m^3	0.021 kg/m^3	0.608
Ochoco	FRG 1	CBD	93212	0.031 kg/m^3	0.022 kg/m^3	0.598
Ochoco	FRG 3	CBD	22380	0.028 kg/m^3	0.019 kg/m^3	0.555
Powell	Local	CC	55239	8.81%	6.52%	0.859
Powell	Global	CC	55239	9.46%	7.17%	0.837
Powell	FRG 3	CC	24599	10.04%	7.54%	0.83
Powell	FRG 4	CC	28985	8.10%	6.03%	0.835
Powell	Local	CH	55239	4.44 m	3.33 m	0.752
Powell	Global	CH	55239	4.76 m	3.58 m	0.718
Powell	FRG 3	CH	24599	4.90 m	3.67 m	0.75
Powell	FRG 4	CH	28985	4.06 m	3.05 m	0.739
Powell	Local	CBH	55239	2.13 m	1.54 m	0.533
Powell	Global	CBH	55239	2.10 m	1.49 m	0.545
Powell	FRG 3	CBH	24599	2.31 m	1.64 m	0.588
Powell	FRG 4	CBH	28985	1.76 m	1.28 m	0.54
Powell	Local	CBD	55239	0.054 kg/m^3	0.037 kg/m^3	0.59
Powell	Global	CBD	55239	0.051 kg/m^3	0.035 kg/m^3	0.635
Powell	FRG 3	CBD	24599	0.062 kg/m^3	0.045 kg/m^3	0.603
Powell	FRG 4	CBD	28985	0.036 kg/m^3	0.024 kg/m^3	0.61
Southern Coast	Local	CC	491731	16.35%	12.32%	0.613
Southern Coast	Global	CC	491731	15.75%	11.19%	0.642
Southern Coast	FRG 1	CC	125537	15.97%	11.11%	0.684
Southern Coast	FRG 3	CC	83831	13.30%	8.60%	0.731
Southern Coast	FRG 5	CC	281017	16.81%	12.09%	0.554
Southern Coast	Local	CH	491731	7.31 m	5.30 m	0.78
Southern Coast	Global	CH	491731	8.39 m	6.19 m	0.709
Southern Coast	FRG 1	CH	125537	7.37 m	5.51 m	0.634
Southern Coast	FRG 3	CH	83831	7.90 m	5.64 m	0.721
Southern Coast	FRG 5	CH	281017	8.47 m	6.27 m	0.738
Southern Coast	Local	CBH	491731	4.00 m	3.00 m	0.575
Southern Coast	Global	CBH	491731	4.41 m	3.36 m	0.485
Southern Coast	FRG 1	CBH	125537	3.33 m	2.50 m	0.555
Southern Coast	FRG 3	CBH	83831	4.37 m	3.31 m	0.54
Southern Coast	FRG 5	CBH	281017	4.70 m	3.60m	0.448
Southern Coast	Local	CBD	491731	0.107 kg/m^3	0.077 kg/m^3	0.552
Southern Coast	Global	CBD	491731	0.120 kg/m^3	0.087 kg/m^3	0.431
Southern Coast	FRG 1	CBD	125537	0.123 kg/m^3	0.090 kg/m^3	0.478
Southern Coast	FRG 3	CBD	83831	0.110 kg/m^3	0.079 kg/m^3	0.52
Southern Coast	FRG 5	CBD	281017	0.118 kg/m^3	0.084 kg/m^3	0.431
Tahoe	Local	CC	420960	9.64%	7.09%	0.874
Tahoe	Global	CC	420960	10.51%	7.98%	0.85
Tahoe	FRG 1	CC	337658	10.35%	7.64%	0.853
Tahoe	FRG 3	CC	76895	9.98%	7.55%	0.818

Tahoe	Local	CH	420960	5.96 m	4.54 m	0.667
Tahoe	Global	CH	420960	6.35 m	4.90 m	0.622
Tahoe	FRG 1	CH	337658	6.67 m	5.18 m	0.591
Tahoe	FRG 3	CH	76895	5.82 m	4.44 m	0.63
Tahoe	Local	CBH	420960	2.28 m	1.69 m	0.587
Tahoe	Global	CBH	420960	2.41 m	1.79 m	0.541
Tahoe	FRG 1	CBH	337658	2.51 m	1.86 m	0.53
Tahoe	FRG 3	CBH	76895	2.15 m	1.63 m	0.484
Tahoe	Local	CBD	420960	0.053 kg/m^3	0.036 kg/m^3	0.789
Tahoe	Global	CBD	420960	0.054 kg/m^3	0.037 kg/m^3	0.775
Tahoe	FRG 1	CBD	337658	0.058 kg/m^3	0.040 kg/m^3	0.761
Tahoe	FRG 3	CBD	76895	0.038 kg/m^3	0.025 kg/m^3	0.75
Teanaway	Local	CC	25817	10.01%	7.48%	0.8
Teanaway	Global	CC	25817	10.41%	7.91%	0.785
Teanaway	FRG 1	CC	6102	10.29%	7.78%	0.702
Teanaway	FRG 3	CC	19181	10.04%	7.54%	0.809
Teanaway	Local	CH	25817	4.51 m	3.54 m	0.509
Teanaway	Global	CH	25817	4.52 m	3.51 m	0.533
Teanaway	FRG 1	CH	6102	4.32 m	3.32 m	0.419
Teanaway	FRG 3	CH	19181	4.23 m	3.26 m	0.611
Teanaway	Local	CBH	25817	2.03 m	1.53 m	0.525
Teanaway	Global	CBH	25817	2.14 m	1.62 m	0.47
Teanaway	FRG 1	CBH	6102	2.16 m	1.64 m	0.451
Teanaway	FRG 3	CBH	19181	2.04 m	1.55 m	0.521
Teanaway	Local	CBD	25817	0.042 kg/m^3	0.030 kg/m^3	0.667
Teanaway	Global	CBD	25817	0.043 kg/m^3	0.031 kg/m^3	0.645
Teanaway	FRG 1	CBD	6102	0.038 kg/m^3	0.026 kg/m^3	0.501
Teanaway	FRG 3	CBD	19181	0.044 kg/m^3	0.032 kg/m^3	0.662

Table 5. Accuracy assessment of local, global, and fire regime group (FRG) models for test landscapes. All data within the landscape's datasets used for model assessment. N refers to the number of samples used for the performance metrics.

Landscape	Model	Canopy Fuel Variable	N	RMSE	MAE	R²
Illilouette	Dinkey	CC	124229	10.95%	8.53%	0.680
Illilouette	Global	CC	124229	10.68%	8.31%	0.696
Illilouette	FRG 1	CC	59941	11.51%	9.05%	0.684
Illilouette	FRG 3	CC	59606	9.56%	7.36%	0.717
Illilouette	Dinkey	CH	124229	8.06 m	6.28 m	0.300
Illilouette	Global	CH	124229	7.84 m	6.18 m	0.338
Illilouette	FRG 1	CH	59941	8.81 m	6.90 m	0.280
Illilouette	FRG 3	CH	59606	7.40 m	5.96 m	0.191
Illilouette	Dinkey	CBH	124229	4.25 m	3.16 m	-0.168
Illilouette	Global	CBH	124229	4.11 m	3.03 m	-0.094
Illilouette	FRG 1	CBH	59941	4.77 m	3.49 m	-0.120
Illilouette	FRG 3	CBH	59606	3.35 m	2.55 m	-0.094
Illilouette	Dinkey	CBD	124229	0.035 kg/m^3	0.024 kg/m^3	0.417
Illilouette	Global	CBD	124229	0.030 kg/m^3	0.020 kg/m^3	0.578
Illilouette	FRG 1	CBD	59941	0.033 kg/m^3	0.022 kg/m^3	0.540
Illilouette	FRG 3	CBD	59606	0.029 kg/m^3	0.019 kg/m^3	0.529
North Coast	South Coast	CC	947615	15.93%	12.47%	0.553
North Coast	Global	CC	947615	14.22%	10.76%	0.644
North Coast	FRG 3	CC	49945	14.05%	9.73%	0.692
North Coast	FRG 5	CC	895207	14.33%	10.24%	0.631
North Coast	South Coast	CH	947615	7.83 m	5.98 m	0.702
North Coast	Global	CH	947615	8.15 m	6.27 m	0.677
North Coast	FRG 3	CH	49945	9.05 m	6.80 m	0.565
North Coast	FRG 5	CH	895207	10.04 m	7.61 m	0.510
North Coast	South Coast	CBH	947615	4.65 m	3.59 m	0.430
North Coast	Global	CBH	947615	4.66 m	3.64 m	0.428
North Coast	FRG 3	CBH	49945	4.37 m	3.44 m	0.433
North Coast	FRG 5	CBH	895207	4.94 m	3.82 m	0.359
North Coast	South Coast	CBD	947615	0.116 kg/m^3	0.085 kg/m^3	0.511
North Coast	Global	CBD	947615	0.121 kg/m^3	0.090 kg/m^3	0.465
North Coast	FRG 3	CBD	49945	0.125 kg/m^3	0.092 kg/m^3	0.510
North Coast	FRG 5	CBD	895207	0.126 kg/m^3	0.095 kg/m^3	0.417
Slate Creek	Clear	CC	320971	25.29%	22.28%	-0.446
Slate Creek	Global	CC	320971	14.96%	12.18%	0.494
Slate Creek	FRG 1	CC	55873	13.90%	10.74%	0.633
Slate Creek	FRG 3	CC	190294	14.62%	11.68%	0.561
Slate Creek	FRG 4	CC	73862	15.49%	13.02%	-0.057
Slate Creek	Clear	CH	320971	7.05 m	5.52 m	0.249
Slate Creek	Global	CH	320971	6.36 m	4.90 m	0.390
Slate Creek	FRG 1	CH	55873	7.23 m	5.72 m	0.397
Slate Creek	FRG 3	CH	190294	6.62 m	5.13 m	0.345

Slate Creek	FRG 4	CH	73862	5.91 m	4.65 m	-0.099
Slate Creek	Clear	CBH	320971	3.84 m	2.84 m	-0.004
Slate Creek	Global	CBH	320971	3.58 m	2.69 m	0.125
Slate Creek	FRG 1	CBH	55873	4.33 m	3.36 m	0.142
Slate Creek	FRG 3	CBH	190294	3.72 m	2.81 m	0.070
Slate Creek	FRG 4	CBH	73862	2.75 m	2.23 m	-0.411
Slate Creek	Clear	CBD	320971	0.114 kg/m^3	0.080 kg/m^3	-0.889
Slate Creek	Global	CBD	320971	0.060 kg/m^3	0.045 kg/m^3	0.479
Slate Creek	FRG 1	CBD	55873	0.060 kg/m^3	0.045 kg/m^3	0.503
Slate Creek	FRG 3	CBD	190294	0.063 kg/m^3	0.047 kg/m^3	0.473
Slate Creek	FRG 4	CBD	73862	0.054 kg/m^3	0.042 kg/m^3	0.198
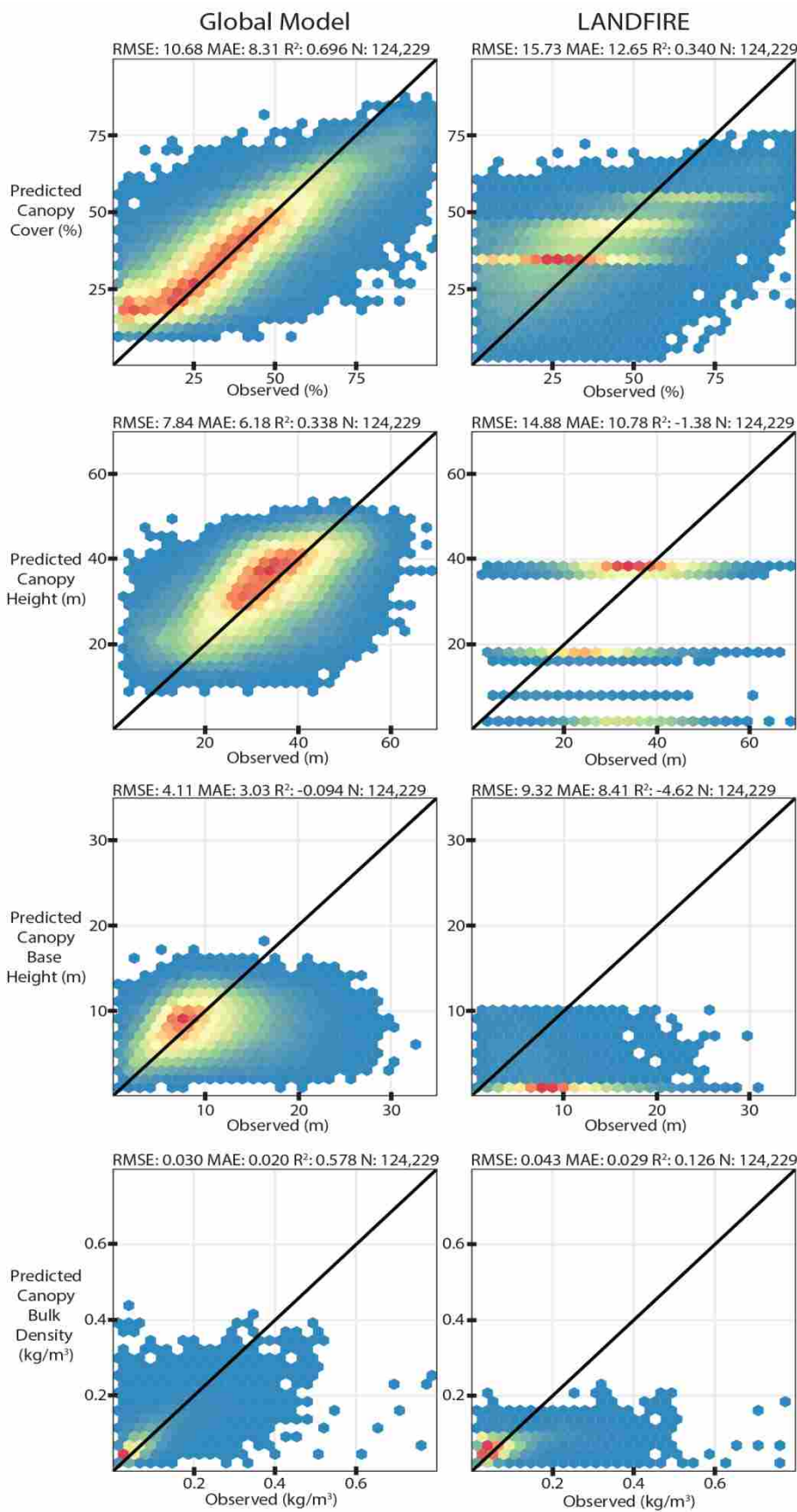
Figure 6. Predicted versus observed plots for the Illilouette test landscape using the global GBM model (left column) and existing LANDFIRE layers (right column). Point density indicated with blue (low) to red (high) gradient.
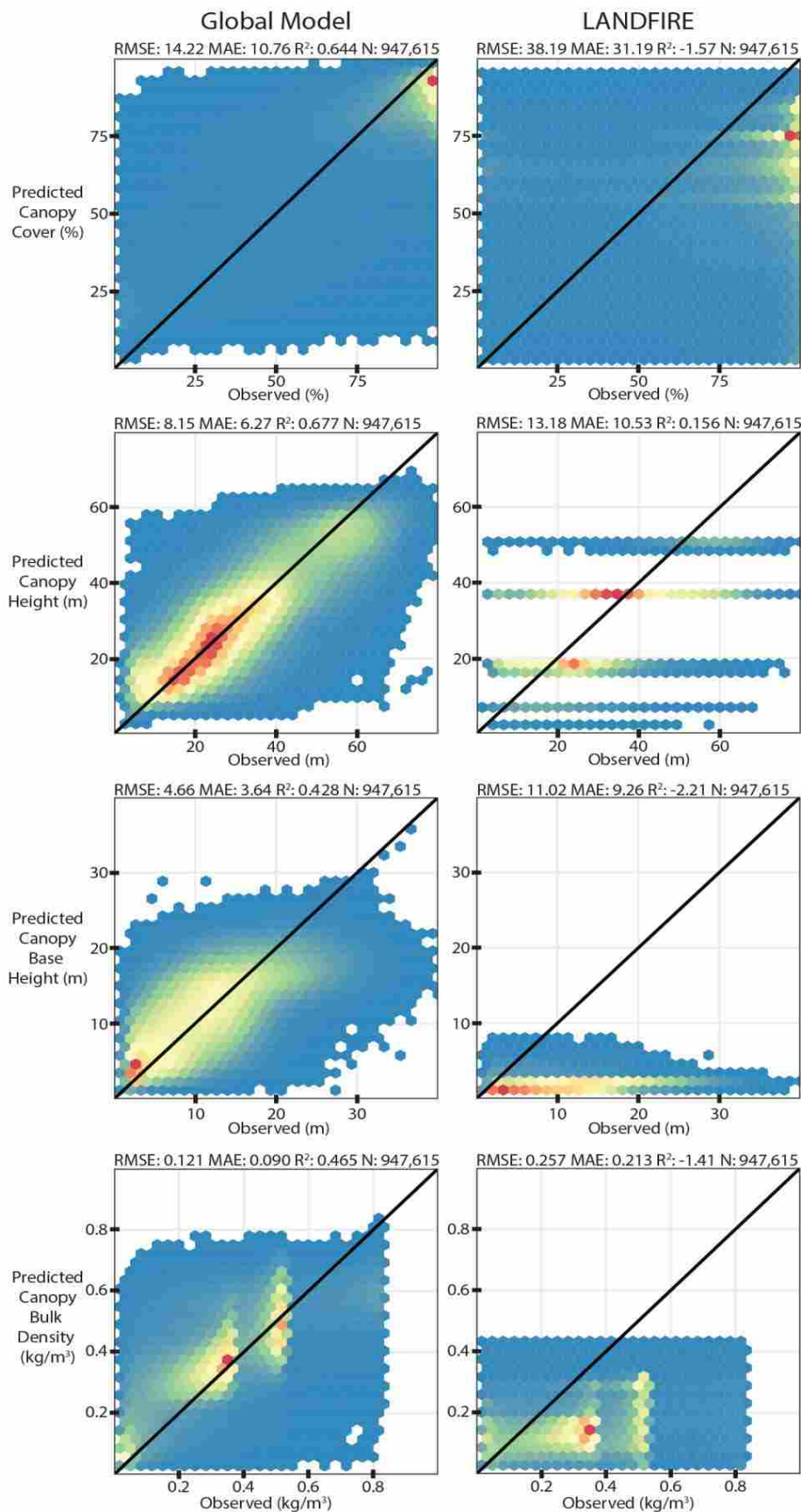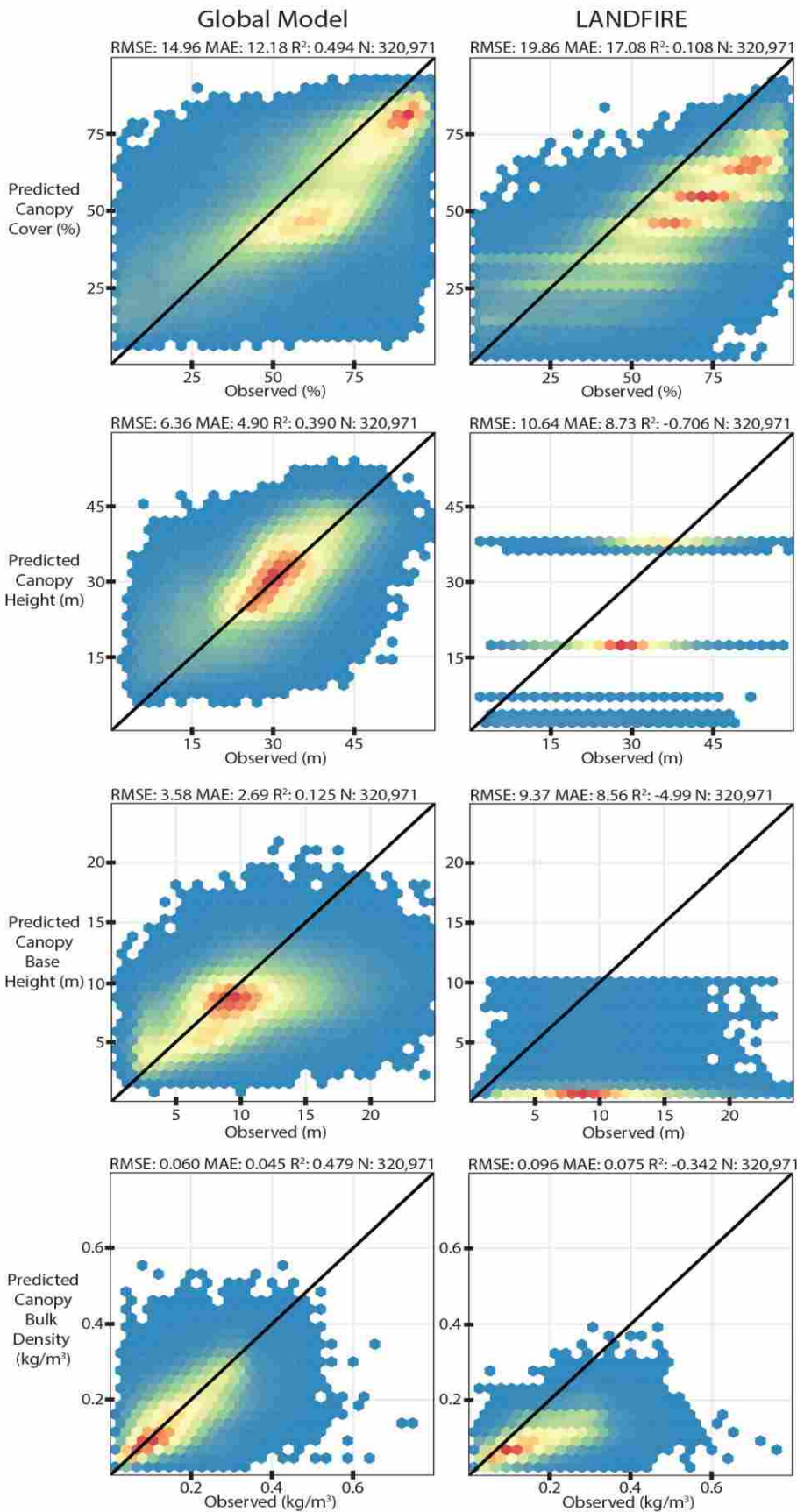
Figure 7. Predicted versus observed plots for the North Coast test landscape using the global GBM model (left column) and existing LANDFIRE layers (right column). Point density indicated with blue (low) to red (high) gradient.

Figure 8. Predicted versus observed plots for the Slate Creek test landscape using the global GBM model (left column) and existing LANDFIRE layers (right column). Point density indicated with blue (low) to red (high) gradient.

66

For comparisons between response variables using the global model, CBH proved the most difficult to characterize with mean R^2 of 0.547 for the training landscapes and 0.153 for the test landscapes. CC had the best variance explained with a mean R^2 of 0.73 for the training landscapes and 0.611 for the test landscapes. CBD had a mean R^2 of 0.602 for the training landscapes and 0.507 for the test landscapes. CH had a mean R^2 of 0.631 for the training landscapes and 0.385 for the test landscapes.

The Ochoco landscape contained the Corner Creek Fire mostly within its extent with an ignition date of June 29, 2015. Pre-fire spectral indices were calculated using 2014 imagery and post-fire used 2016 imagery. Model predictions for both years and from the global and Ochoco landscape models were compared (Figs. 9, 10, 11, and 12). High severity pixels showed the largest changes in canopy predictions with a mean percent decrease of 51.2% for CC (55.6% to 27.13%), 39.4% for CH (28.0 m to 16.9 m), 40.3% for CBH (8.3 m to 5.0 m), and 55.0% for CBD (0.135 to 0.061 kg/m^3) using the global model. The local Ochoco model showed similar results but the percent changes were larger for CC (60.0% decrease), smaller for CH (14.9% decrease), larger for CBH (46.2% decrease), and smaller for CBD (52.2% decrease).

Moderate burn severity pixels followed expectations with mean decreases for each canopy variable but at reduced percentages compared to high severity pixels. For the global model, CC decreased by 31.7%, CH decreased by 20.0%, CBH decreased by 26.3%, CBD decreased by 32.5%. For the Ochoco local model, CC decreased by 38.0%, CH decreased by 1.2%, CBH decreased by 29.5%, and CBD decreased by 32.8%.

For unburned to low severity pixels, little change was seen in the canopy fuel variables. For the global model, CC had a mean percentage decrease of 5.8%, CH decreased by 2.3%, CBH decreased by 5.3%, and CBD decreased by 2.5%. For the local Ochoco model, CC decreased by 7.9%, CH increased by 3.8%, CBH decreased by 4.5%, and CBD decreased by 5.1%.

Variable importance calculations show the relative influence of each feature by assessing whether the feature was used to split during the tree building process and how much the squared error reduced as a result of splitting on that feature over all trees. Fig. 13 shows the variable importance for the four canopy fuel variables using the global model. The spectral predictors Med NBR, Med Bright, and Med Green each had the highest importance for at least one variable. Med NBR had the highest importance for CC and CBH and was also of secondary importance for CH and CBD. Med Green had low importance for CC, CH, and CBH but was the highest for CBD; an interesting result considering CBD is calculated using CC and CH, and Med Green had low variable importance for these two variables. Latitude, elevation, and aspect had moderate importance in all responses except

for CBD. The median spectral indices had higher importance than all their maximum counterparts. Tasseled cap wetness had the least importance with both the maximum and minimum indices in the bottom four of importance. Max NDVI and Max NBR were the other two predictors with the lowest importance overall.
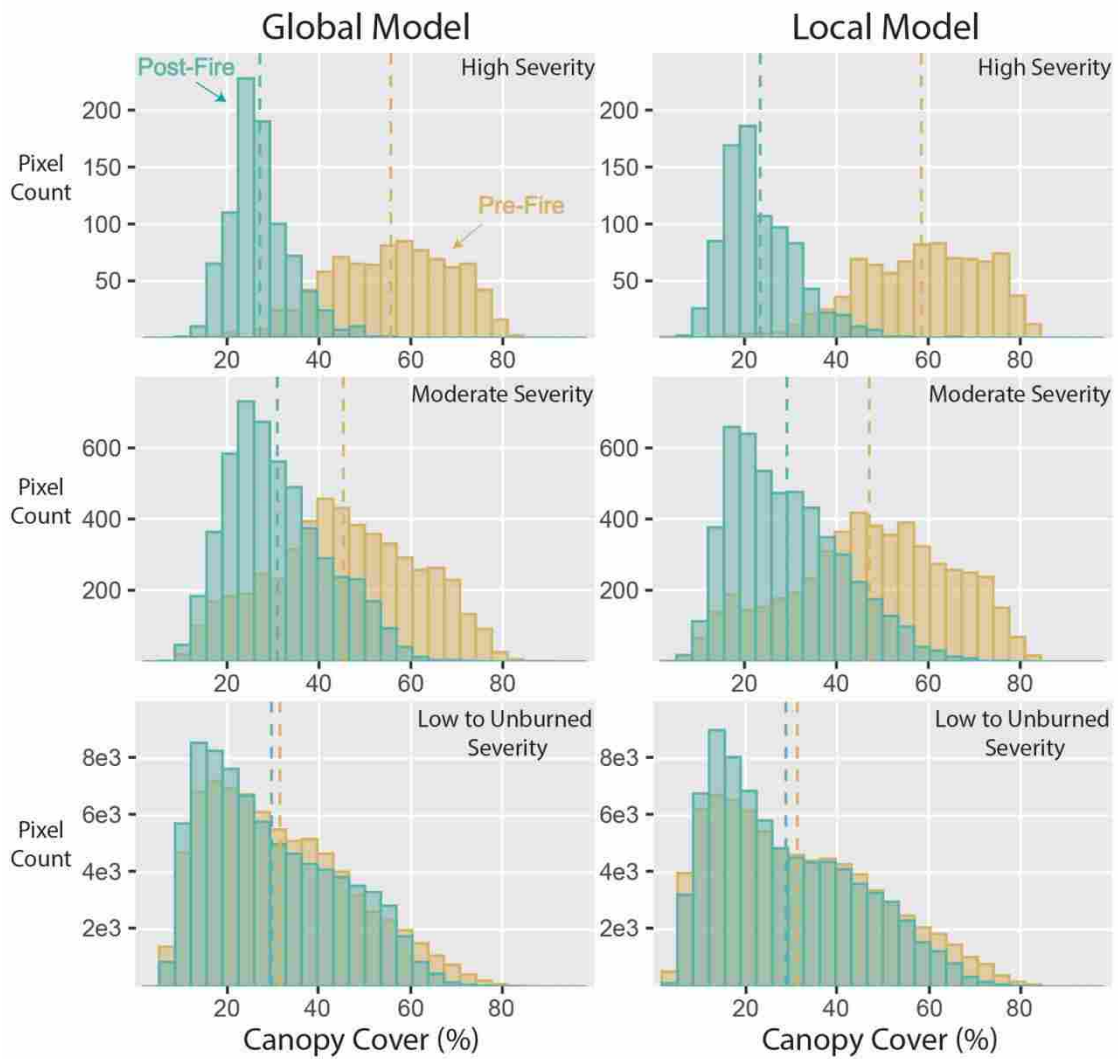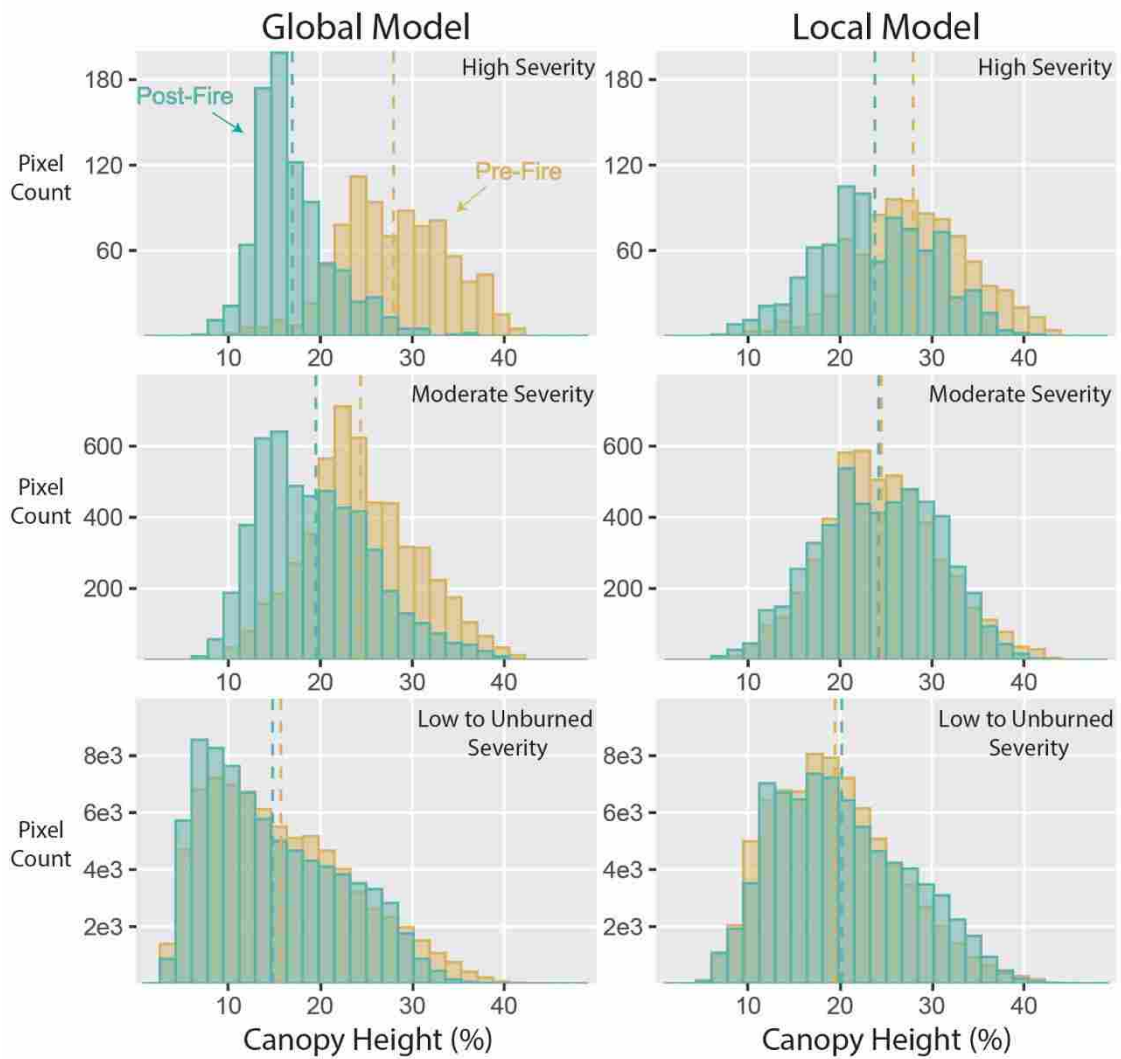

Figure 9. Comparison of pre- and post-fire, global and local (Ochoco) model predictions for canopy cover in high (top row), moderate (middle row), and low to unburned (bottom row) fire severity classes for the Corner Creek Fire. Vertical dotted lines depict the mean prediction value.
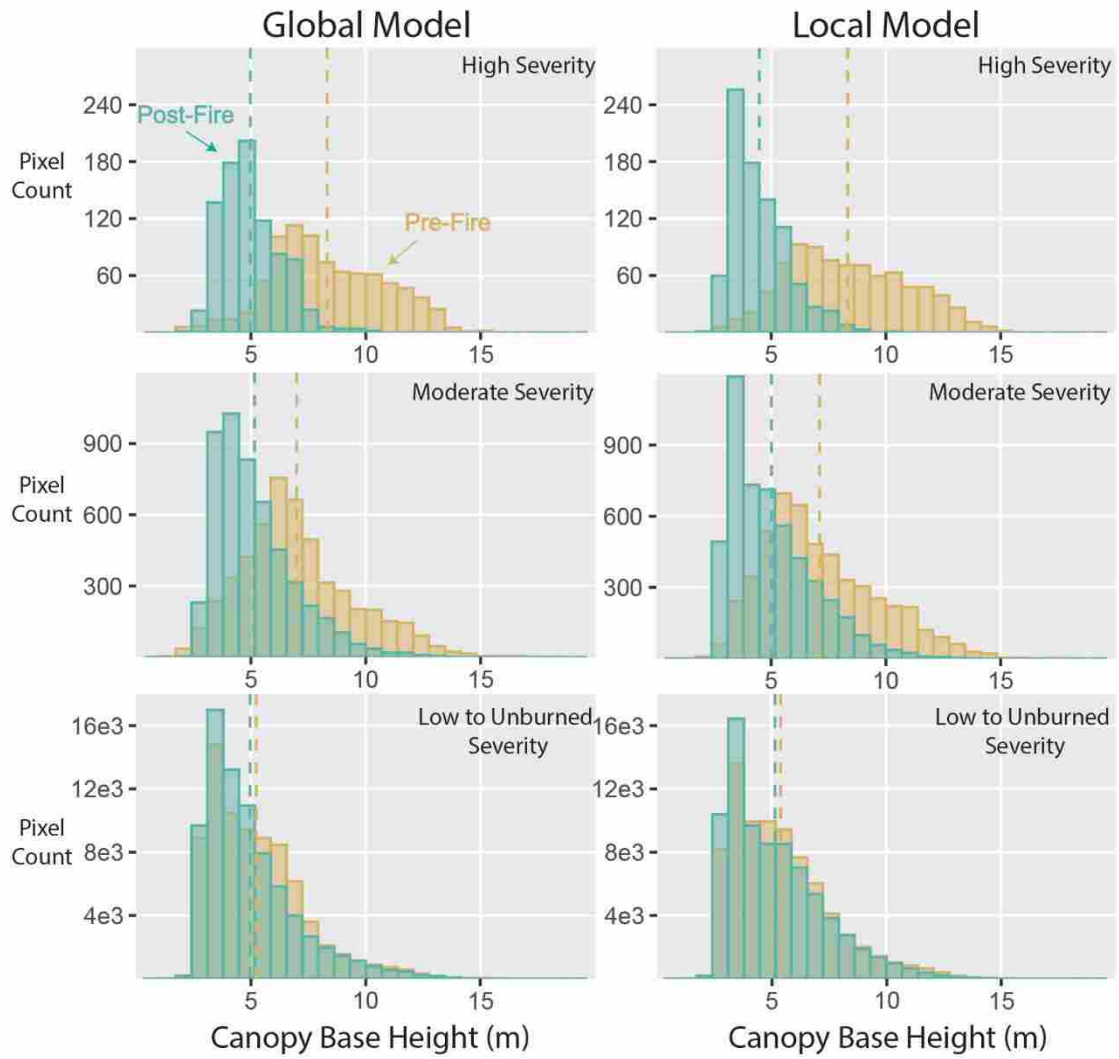
Figure 10. Comparison of pre- and post-fire, global and local (Ochoco) model predictions for canopy height in high (top row), moderate (middle row), and low to unburned (bottom row) fire severity classes for the Corner Creek Fire. Vertical dotted lines depict the mean prediction value.
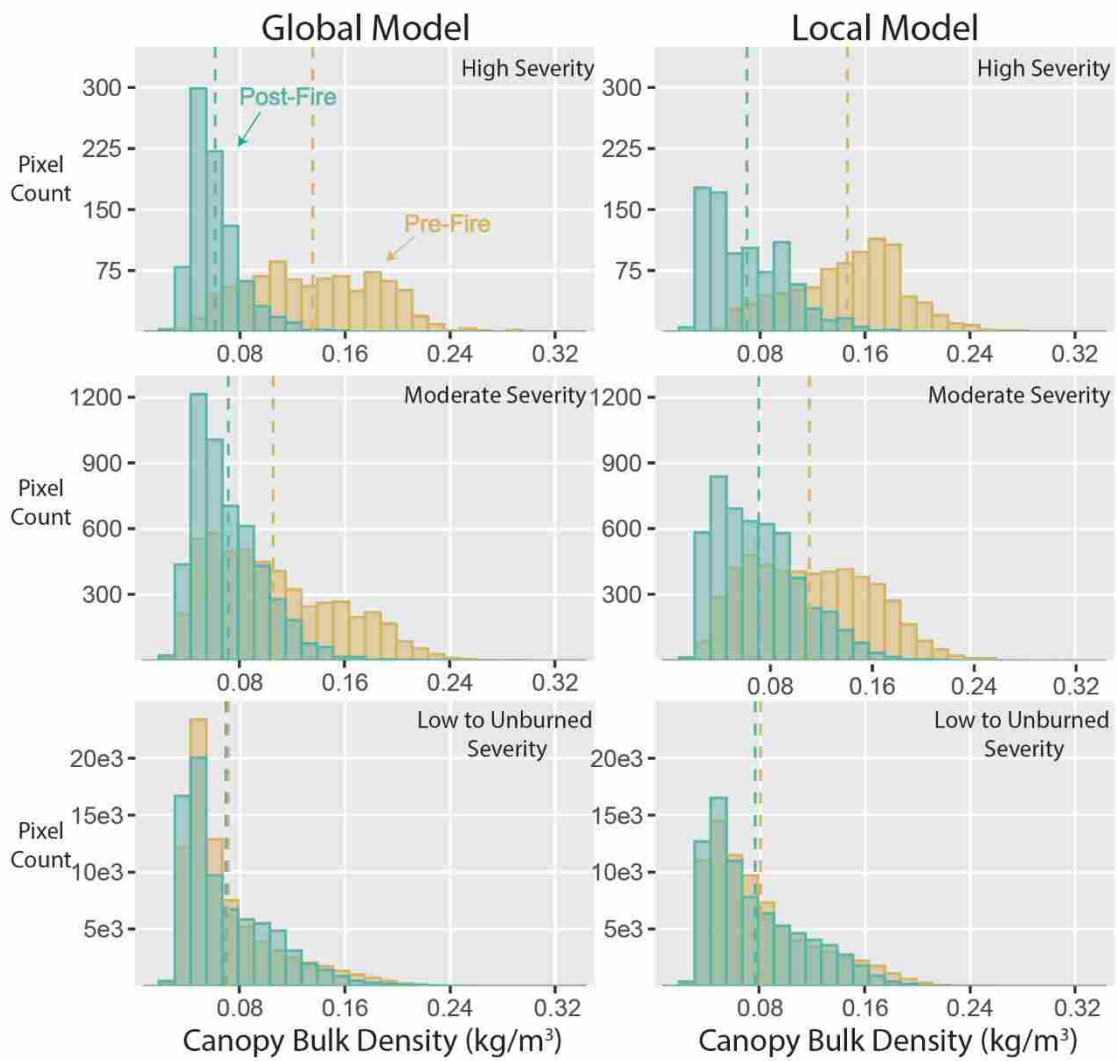
Figure 11. Comparison of pre- and post-fire, global and local (Ochoco) model predictions for canopy base height in high (top row), moderate (middle row), and low to unburned (bottom row) fire severity classes for the Corner Creek Fire. Vertical dotted lines depict the mean prediction value.

Figure 12. Comparison of pre- and post-fire, global and local (Ochoco) model predictions for canopy bulk density in high (top row), moderate (middle row), and low to unburned (bottom row) fire severity classes for the Corner Creek Fire. Vertical dotted lines depict the mean prediction value.
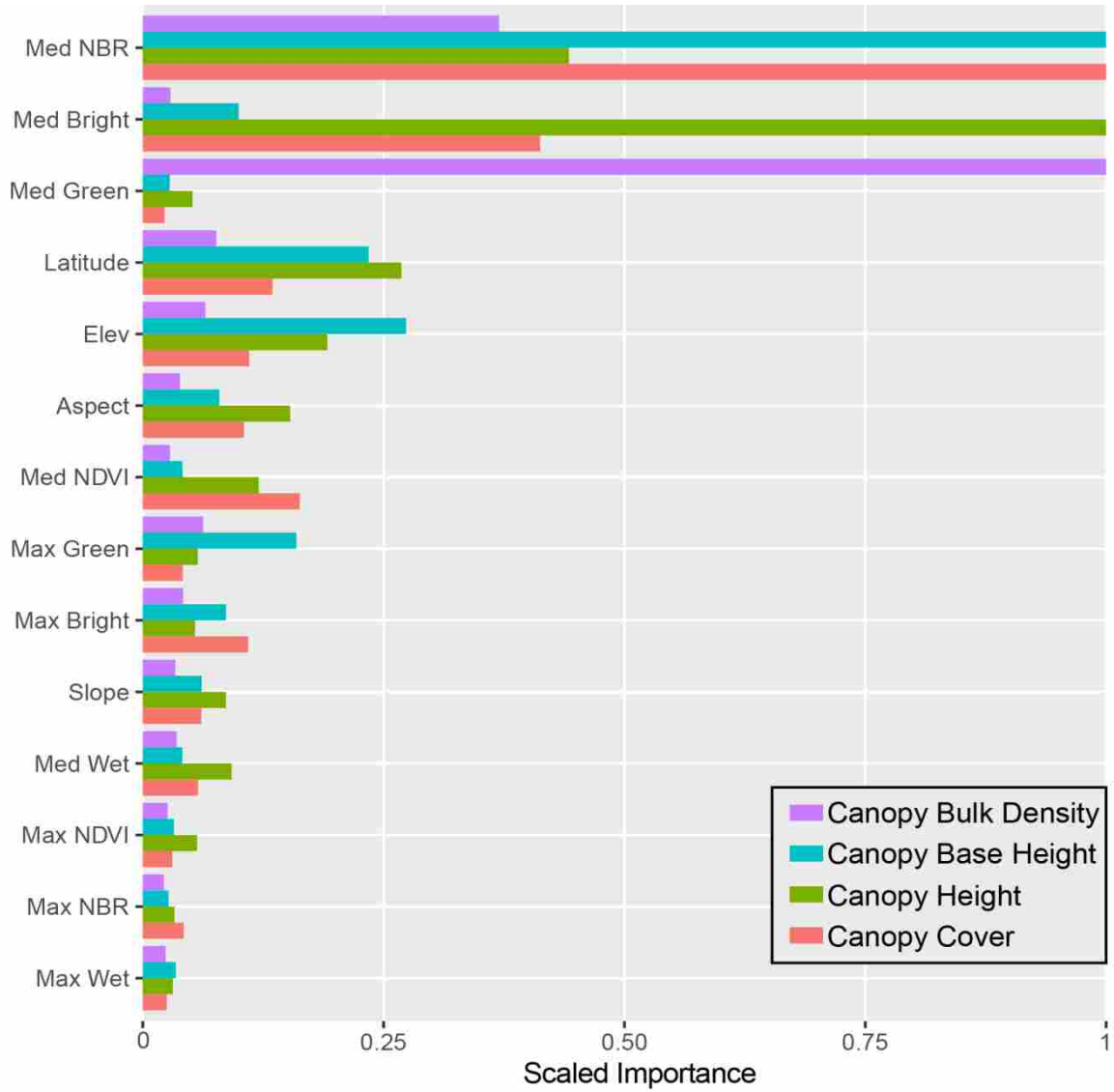
Figure 13. Variable importance for each canopy fuel response variable derived from the global gradient boosting maching (GBM) model using training landscape data.

4) Discussion

While the local landscapes show less error and increased variance explained than the FRG and global models in the training landscapes, the test landscapes expose evidence of overfitting and reduced model performance. In addition to the parsimony provided, the results support the use of the global model for predicting canopy fuels over the western US. The FRG models have nearly equivalent performance to the global models but the splitting of data and use of multiple models is unnecessary given their near equivalency. GBM's boosting approach is likely able to differentiate the areas of spatial variance in the feature-response relationships observed in the partial dependence plots. In effect, a single GBM can create local models within the ensemble process, as samples with different feature-response relationships would show as residuals that GBM then trains on in subsequent trees. The use of latitude and its moderate importance for 3 of 4 of the canopy fuel variables supports this assertion as spatial variance in spectral-canopy fuel relationship can be captured within this locational metric. While longitude was considered as a potential feature, not enough data was present to confidently characterize the longitudinal precipitation gradients across the western US. Without adequate coverage, use of longitude could potentially lead to overfitting and a lack of generality in the model.

4.1 Comparisons

Comparisons to LANDFIRE show increased performance in every case. LiDAR's ability to provide added value is unquestionable. However, the conception of these LiDAR-based metrics is different than LANDFIRE's. Indeed, ambiguity in the formulation of CBH and CBD largely stems from the inability to quickly and consistently measure them in the field (Keane et al. 2005). LANDFIRE's definition of CBD (Table 1) allows LiDAR CC and CH metrics to be inserted directly into the equation for near equivalency. However, this equation was only one of two equations used to calculate CBD in Reeves et al. (2009), but discriminating between which equation was applied at the pixel level is not possible and could explain part of the mismatch in the comparisons here. CBH's LiDAR-based definition (mean height minus standard deviation, Rowell et al. 2005)) is not equivalent to LANDFIRE's formulation, which is the lowest vertical height at which the vertical distribution of CBD is ≥ 0.012 kg/m^3 (Reinhardt et al. 2006), though there is conceptual correspondence. As Peterson et al. (2015) notes, the LiDAR-derived definition of CBH over-predicts compared to Reinhardt et al. (2006) but use of this simplified, LiDAR-based metric allows improved characterization of variation over a landscape. A bias correction is necessary as over-prediction of CBH leads to under-prediction of crown fire in operational fire models (Cruz and Alexander 2010). The best interpretation for application is that these

73

remotely-sensed canopy fuel layers relate to and covary in the same fashion as but do not exactly replicate the canopy fuel variables conceived in the fire models.

Direct comparisons of findings to other research are difficult given the breadth of the study area and the use of LiDAR-derived canopy fuel-specific response variables. Matasci et al. (2018) used a similar approach, predicted over the entirety of the Canadian boreal shield, and used similar response canopy variables. Table 6 summarizes the comparison. The results from the training and test landscapes are shown, as Matasci et al. (2018)'s R^2 estimation is more comparable to the training landscapes than the test landscapes (though not as independent of a test dataset). Hansen et al. (2016) used space-borne LiDAR and Landsat 7 and 8 data to predict tree height across Sub-Saharan Africa and reported MAE of 2.45 m. They report a MAE of 4.65 m for tree heights > 20 m though, which is more consistent in error and observed tree heights presented here (MAE 3.99 m for training landscapes and 5.78 m for test landscapes). Wilkes et al. (2015) also map canopy heights over a broad-area in Australia and report an RMSE of 5.6 m. Ahmed et al. (2015) focus on a 2,600 ha area in British Columbia and report an R^2 of 0.67 for CC and 0.82 for CH. Stojanova et al. (2010) use LiDAR and Landsat data over Slovenian forests and report RMSE of ~14.7% for CC and ~2.1 m for CH. Hyde et al. (2006) uses a dataset near the Dinkey landscape and reports similar results (Dinkey results in parentheses): for

Table 6. Comparison of R^2 estimates between global model results and those of Matasci et al. (2018). Results are shown for both training (top) and test landscapes (bottom; right column).

Matasci et al. (2018) Results			Global Model Results		
Variable	N	R^2	Variable	N	R^2
Mean Height	20,182	0.49	CBH*	1,550,354	0.547
SD Height	20,182	0.377		1,392,815	0.153
Height (95th Percentile)	20,182	0.495	Height (99th Percentile)	1,550,354	0.631
				1,392,815	0.478
Cover	20,182	0.612	Cover	1,550,354	0.729
				1,392,815	0.611

*CBH is calculated by subtracting SD Height from Mean Height.

max.height, R^2 of 0.712 (0.675) and RMSE of 9.6 m (6.88 m); for mean height, R^2 of 0.603 and RMSE of 7.5 m; for SD of Heights, R^2 of 0.517 and RMSE 3.7 m (Dinkey CBH R^2 is 0.508 and RMSE 2.73 m). Pascual et al. (2010) report R^2 of 0.62 for mean height and 0.66 for height coefficient of variation (CV) using Landsat imagery in Spain. Erdody and Moskal (2010) use field-based estimates of canopy fuels and relate them to Landsat spectral

indices in south-central WA and report R² values of 0.415 for canopy height (RMSE 5.52 m), 0.309 for CBH (RMSE 3.6 m), and 0.602 for CBD (RMSE 0.041 kg/m^3).

While these studies tend to produce better results compared to the test landscapes, the test data from the training landscapes is the more accurate comparison as none of these studies use entirely independent LiDAR datasets from different landscapes. While differences in LiDAR acquisition parameters, especially point density, can be small for height and cover (Jakubowski et al. 2013), these differences likely contribute a source of error especially for CBH, which requires penetration into the canopy. These studies also did not have variable years of LiDAR acquisition and annual differences can have large effects on spectral features driven primarily by precipitation timing and amount, temperature fluctuations, and variable cloud and cloud shadow cover leading to inconsistencies in measurement of median and maximum spectral indices (Ichii et al. 2002, Hermosilla et al. 2015). Nonetheless, the results here are either competitive with or exceed the performance of other research of canopy structure based on Landsat TM, ETM+, and OLI imagery.

4.2 Relevance of predictors and predictor-response spatiotemporal variance

Beyond accuracy comparisons, a generalized and robust model must be able to characterize change due to disturbance. While topographic and locational features aid in the modeling process, they are static through time and a model overfit to these feature will not be responsive to temporal change. For example, despite Matasci et al. (2018)'s relatively high accuracy in prediction across Canada, four of the top five predictors in terms of variable importance are either topographic or locational in nature (elevation, latitude, longitude, and slope). This implies that their model may produce similar predictions over time regardless of changes in vegetation characteristics. In my study, elevation, aspect, and latitude show moderate variable importance, but three spectral features are much more important. For CC and CH, Med NBR and Med Bright have the vast majority of the importance. For CBH, Med NBR is the most important followed by latitude and elevation, and CBD has Med Green and Med NBR predominately important. This implies that the model is able to recognize and characterize change over time. This assertion is supported by findings from the Corner Creek Fire on the Ochoco landscape which shows large reductions in canopy fuels in high burn severity pixels, moderate change in moderate burn severity pixels, and little change in low to unburned pixels. Although validation data is not available for this assessment, the changes predicted are logically intuitive. Despite the large reductions in fuels shown from the Corner Creek Fire, however, the model is likely underestimating the amount of change. As noted in multiple modeling studies at these scales (e.g. Hansen et al. 2016), and seen in the predicted vs. observed plots here (Figs. 6, 7, 8), predictions tend to over-estimate at the low end of the ranges and tend to under-

estimate at the high end of the ranges. Indeed, the models were only trained with forested samples and thus have no ability to predict type changes to grass- or shrub-dominated areas. A separate step in a complete algorithm could determine forest or non-forest ecosystem types before applying canopy fuel models as done in LANDFIRE (Rollins et al. 2009).

The consistent distributions between multiple years in the unburned to low severity pixels also imply that global and local models provide consistent canopy estimations through time despite year to year variability in the spectral indices. Though the variable year of LiDAR acquisition and corresponding Landsat indices may be a source of error in the global model when predicting on a particular landscape in a particular year, it also likely improves generality and the ability to predict over time. Additional assessment is necessary to characterize the model's ability to predict over time, but the results here are encouraging.

4.3 Imputation vs. Regression

Imputation approaches, such as in Zald et al. (2016) and Matasci et al. (2018), ensure realistic covariance between several response variables but also limit predictions to observed covariance structures. Given the large amount of sample data used in this approach (N= 16,702,582), imputation is a viable alternative approach. However, the flexibility to model new structural assemblages and maintain performance in significantly different forest ecosystems is a priority. In imputation approaches, a similarity or distance metric condenses information from all predictor variables to one value, which is used to match a sample for prediction to an existing sample. The observed sample's response values are then applied as the prediction. Any confusion in the sample matching would potentially lead to errors in all response variables. Unique or rare covariance structures could be ignored as well. For this study, post-modeling rules could effectively ensure realistic covariance structures as well as imputation approaches. For example, CBH is the worst performing model here, and updates to achieve realistic values could be based on a better performing model such as CH (e.g., force CBH to a maximum proportion of CH or relate it to CBD, which is more in line with the Reinhardt et al. (2006) CBH formulation).

4.4 Data Processing Architecture

The combination of Spark, R, and H2O software provides an open-source, parallel, scalable, and efficient data processing and model-building environment. Though processing and analysis was completed on a modest server (48 processing cores, 256 GB of RAM), the entirety of the workflow could run on cloud services and expanded or contracted depending on processing needs. Each individual model was restricted to 750 seconds of modeling time

and most would be stopped before reaching this limit by the error tolerance metric. R is a popular statistical and data processing software, but its object-oriented design holds data in memory which limits users to available RAM on a system. A number packages are available to hold files on disk including the raster R package (Hijmans et al. 2019), but the Spark processing environment features lazy evaluation and transformations that do not necessarily require the full dataset to be loaded into memory. Data can be manipulated within Spark using an R interface with the dplyr (Wickham et al. 2019) and sparklyr (Luraschi) packages. Ultimately, the final dataset does need to be held in-memory for the H20 machine learning modeling process, but H20 can use the entire cluster to hold the dataset and additional Spark nodes can be added to scale to project needs.

Google Earth Engine (GEE) offers a cloud-based processing platform and all Landsat imagery was processed to the spectral indices. The processing capabilities are vast, but certain limitations are purposely applied to reduce the potential data download sizes a single user can attempt. This required tiling and manual downloading of data which required significant amounts of time. A potential alternative option would be to complete the entire workflow within GEE and output the final fuels prediction maps on an annual basis. However, GEE's machine learning capabilities are currently limited and do not have the significant parameterization possible in the H2O libraries including hyperparameter tuning. The LiDAR datasets would also have to be uploaded and stored for GEE's access. However, as GEE increases in sophistication, this modeling process could likely perform entirely in the GEE environment at some point in the future.

4.5 Potential improvements and future work

Though Landsat TM and ETM+ sensors are practically equivalent, the new Landsat 8 OLI sensor requires transformations for continuity (Roy et al. 2016). While the tasseled cap transformations were designed to maintain continuity among sensors (Baig et al. 2014), NDVI and NBR do not have such corrections considered. While Roy et al. (2016) show small but significant difference in NDVI between ETM+ and OLI sensors, no such assessment has been performed for NBR, but differences are expected. Correcting for differences in these two normalized metrics will likely improve temporal continuity and minimize differences in the datasets and model predictions due to varying years of acquisition and likely improve the ability of the model to predict change over time.

The tasseled cap transformations themselves were designed for top-of-atmosphere (TOA) reflectance but applied to the surface reflectance products here. Inconsistencies in the literature are present with multiple approaches used but in general, the application of TOA transformations has been successful with surface reflectance products. Indeed, the scenes used to develop the transformations were specifically chosen to have little

atmospheric contamination (Crist 1985, Huang et al. 2002, Baig et al. 2014). DeVries et al. (2016) argue for the use of one set of coefficients for use with surface reflectance products from multiple sensors and apply those of Crist (1985), which were designed for Landsat-5 TM data. Kennedy et al. (2010) use the same logic but first applied a scene normalization algorithm to make the spectral space relatively consistent. As Baig et al. (2014) and Huang et al. (2002) designed the Landsat-7 ETM+ and Landsat-8 OLI transformations for continuity and no scene normalization process was applied, the transformations specific to each sensor are used here (Table 1). Following recommendations of Crist (1985), I compared the resulting tasseled cap transformations against each other to evaluate whether orthogonal data is achieved. Figure 14 shows the first two tasseled cap indices (median Brightness and Greenness) for the South Coast (2009 acquisition year using Landsat TM and ETM+ imagery) and Tahoe landscapes (2015 acquisition year using Landsat OLI imagery), and the characteristic tassel cap shapes resembling that of Crist (1985) are present. Nonetheless, increased temporal continuity for spectral indices could improve temporal stability of model predictions especially if features characterizing multi-year spectral data are added as predictor variables.
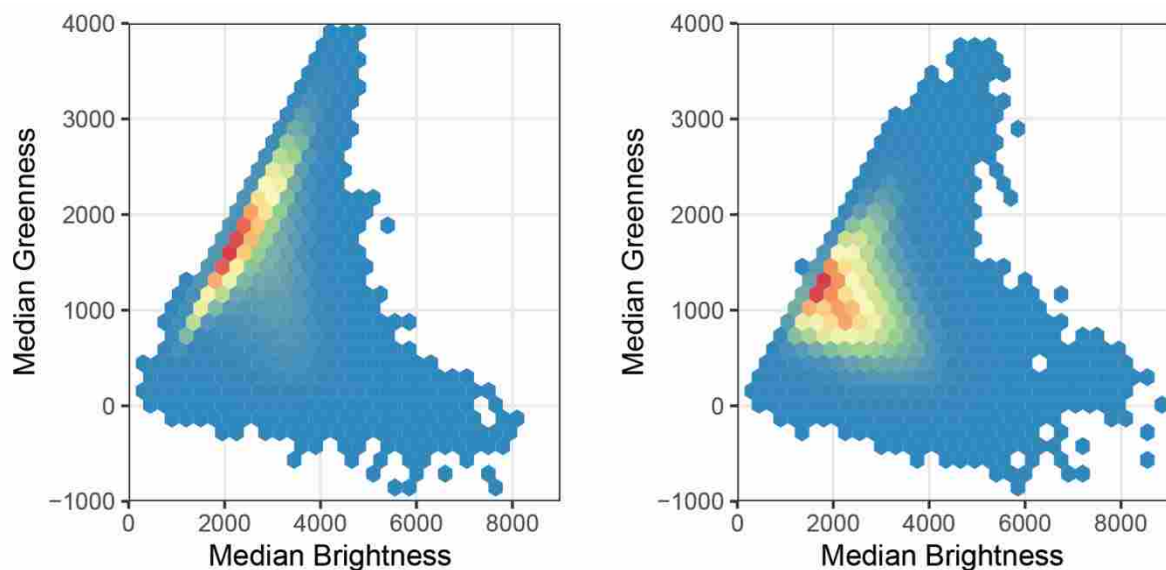

Figure 14. Comparison of relationships between Med Bright and Med Greenness for Landsat TM/ETM+ (right) and Landsat OLI (left). Data taken from the South Coast landscape with a 2009 acquisition year for TM/ETM+ and the Tahoe landscape with a 2015 acquisition year for OLI.

Addition of time series metrics derived from multiple years of Landsat data may also improve model predictions (use of time series data reviewed by Banskota et al. 2014). The primary benefit in this case would be the ability to identify previous disturbance timing and severity and potentially help stabilize the spectral indices' year-to-year variability

caused by cloud and shadow contamination or precipitation variability. Although Zald et al. (2016) and Matasci et al. (2018) employ these sophisticated metrics, their variable importance was minimal in the final assessment. A major downside is the need to re-calculate the temporal trends annually, which can be computationally expensive. A custom script could update features (i.e. add one year to time since disturbance each annual update) and only analyze that year's data for change. The LandTrendr algorithm (Kennedy et al. 2010) segments Landsat time series trends and can create features describing them. This algorithm is integrated in Google Earth Engine and could be run simultaneously when the current-year Landsat spectral indices are calculated and integrated into the workflow.

Additional complex topographic, climatological, and energy balance features could also improve model performance if properly implemented. Examples include: topographic wetness index (Sorensen et al. 2006), solar radiation index, climatic water deficit (Lutz et al. 2010), and cold air pooling (Lundquist et al. 2008). However, these features describe the environmental template on which vegetation and disturbance act upon and their utility for characterizing the current state of canopy fuels is limited and may lead to overfitting.

Obviously, additional LiDAR datasets will improve the model's ability to predict in new areas. In this study, only one dataset is present for CO, one for AZ, none for NM, UT, WY, SD, and NV. The method's focus on model generality and processing architecture can handle significantly more data with ease and will no doubt improve with dataset additions. Many more datasets have been collected and processed (20+) since this research project began and will be included in a future map of canopy fuels for the western US.

Finally, a necessary subject of future research is to develop more consistent methods to predict surface fuel models to accompany the improved canopy fuels data produced in this study, because changes in surface fuels generally have a stronger effect on fire behavior than canopy fuels. Application of rule sets developed as part of the LANDFIRE Total Fuel Change Tool may provide a starting point for predicting surface fuels from improved (and continuous rather than categorical) estimates of canopy properties.

5) Conclusions

LiDAR-Landsat fusion is a capable replacement for existing LANDFIRE canopy fuel mapping protocols, is more easily implemented, and produces better results. A single GBM global model offers a parsimonious solution with small decreases in performance compared to the use of many local models and does not require logic to determine where each local model is most applicable. Local models' partial dependence plots show spatiotemporal variability in predictor-response relationships but the response shapes are

relatively consistent across the predictor's range of values. The global model is able to account for these differences and shows increased generality by outperforming local models on independent datasets. The global model is also able to logically update canopy fuels after wildfire disturbance with similar performance compared to the local model. The increased accuracy and better representation of canopy fuel variability over broad areas will increase the ability to predict fire growth and intensity; thereby enhancing land management decision-making for pre-fire, during-fire, and post-fire activities.

Literature Cited

Abatzoglou, J.T., and A.P. Williams. Impact of anthropogenic climate change on wildfire across western US forests. Proceedings of the National Academy of Sciences 113: 11770-11775.

Ager, A.A., N.M. Vaillant, M.A. Finney, and H.K. Preisler. 2012. Analyzing wildfire exposure and source-sink relationships on a fire prone forest landscape. Forest Ecology and Management 267: 271-283.

Ahmed, O.S.; Franklin, S.E.; Wulder, M.A.; White, J.C. Characterizing stand-level forest canopy cover and height using Landsat time series, samples of airborne LiDAR, and the random forest algorithm. ISPRS Journal of Photogrammetry and Remote Sensing, 2015, 101, 89-101.

Banskota, A., N. Kayastha, M.J. Falkowski, M.A. Wulder, R.E. Froese, and J.C. White. 2014. Forest monitoring using Landsat time series data: a review. Canadian Journal of Remote Sensing 40: 362-384.

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

Bell, D.M. M.J. Gregory, J.L. Ohmann. 2015. Imputed forest structure uncertainty varies across elevational and longitudinal gradients in the western Cascade Mountains, Oregon, USA. Forest Ecology and Management 358: 154-164.

Bell, D.M., M.J. Gregory, V. Kane, J. Kane, R.E. Kennedy, H.M. Roberts, and Z. Yang. 2018. Multiscale divergence between Landsat and lidar-based biomass mapping is related to regional variation in canopy cover and composition. Carbon Balance and Management 13:15.

Bergstra, J. and Y. Bengio. 2012. Random search for hyper-parameter optimization. Journal of Machine Learning Research 13: 281-305.

Bouvier, M., S. Durrieu, R.A. Fournier, and J. Renaud. 2015. Generalizing predictive models of forestry inventory attributes using an area-based approach with airborne LiDAR data. Remote Sensing of Environment 156: 322-334.

Branco, P., L. Torgo, R.P. Ribiero. 2017. SMOGN: a pre-processing approach for imbalanced regression. Proceedings of Machine Learning Research 1: 1-15.

Breiman, L. 2001. Random forests. Machine Learning 45: 5-32.

Calkin, D.E., M. P. Thompson, M.A. Finney, and K.D. Hyde. 2011. A real-time risk assessment tool supporting wildland fire decisionmaking. Journal of Forestry 274-280.

Crist, E.P. 1985. A TM Tesseled Cap equivalent transformation for reflectance factor data. Remote Sensing of Environment 17: 301-306.

Cruz, M.G., and M.E. Alexander. 2010. Assessing crown fire potential in coniferous forests of western North America: a critique of current approaches and recent simulation studies. International Journal of Wildland Fire 19: 377-398.

Cochrane, M.A., C.J. Moran, M.C. Wimberly, A.D. Baer, M.A. Finney, K.L. Beckendorf, J. Eidenshink, and Z. Zhu. 2012. Estimation of wildfire size and risk changes due to fuels treatments. International Journal of Wildland Fire 21: 357-367.

Cohen, W.B., and T.A. Spies. 1992. Estimating structural attributes of douglas-fir/western hemlock forest stands from Landsat and SPOT imagery. Remote Sensing of Environment 41:1-17.

Collins, B.M., S.L. Stephens, G.B. Roller, and J.J. Battles. 2011. Simulating fire and forest dynamics for a landscape fuel treatment project in the Sierra Nevada. Forest Science 57: 77-88.

Cruz, M.G., M.E. Alexander, and R.H. Wakimoto. 2003. Assessing canopy fuel stratum characteristics in crown fire prone fuel types of western North America. International Journal of Wildland Fire 12: 39-50.

Drury, S.A., H.M. Rauscher, E.M. Banwell, S. Huang, and T.L. Lavezzo. 2016. The interagency fuels treatment decision support system: functionality for fuels treatment planning. Fire Ecology 12: 103-123.

DeVries, B., A.K. Pratihast, J. Verbesselt, L. Kooistra, and M. Herold. 2016. Characterizing forest change using community-based monitoring data and Landsat time series. PLOS One 11: e0147121

Elith, Jane, John R Leathwick, and Trevor Hastie. "A Working Guide to Boosted Regression Trees." Journal of Animal Ecology 77.4 (2008): 802-813

Erdody, T.L. and L.M. Moskal. 2010. Fusion of LiDAR and imagery for estimating forest canopy fuels. Remote Sensing of Environment 114: 725-737.

Escuin, S., R. Navarro, and P. Fernandez. 2007. Fire severity assessment by using NBR (normalized burn ratio) and NDVI (normalized difference vegetation index) derived from Landsat TM/ETM images. International Journal of Remote Sensing 29: 1053-1073.

Frazier, R.J., N.C. Coops, M.A. Wulder, and R. Kennedy. 2014. Characterization of aboveground biomass in an unmanaged boreal forest using Landsat temporal segmentation metrics. ISPRS Journal of Photogrammetry and Remote Sensing 92: 137-146.

Friedman, J.H. 2001. Greedy function approximation: a gradient boosting machine. The Annals of Statistics 29: 1189-1232.

Hansen, M.C., P.V. Potapov, R. Moore, M. Hancher, S.A. Turubanova, A. Tyukavina, D. Thau, S.V. Stehman, S.J. Goetz, T.R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C.O. Justice, and J.R.G. Townshend. 2013. High-resolution global maps of 21-century forest cover change. Science 342: 850-853.

Hansen, M.C.; Potapov, P.V.; Goetz, S.J.; Turubanova, S.; Tyukavina, A.; Krylov, A.; Kommareddy, A.; Egorov, A. Mapping tree height distributions in Sub-Saharan Africa using Landsat 7 and 8 data. Remote Sensing of Environment 2016, 185, 221-232.

Hastie, Trevor, Robert Tibshirani, and J Jerome H Friedman. The Elements of Statistical Learning. Vol.1. N.p., page 339: Springer New York, 2001.

Hava, J. et al. rsparkling: R interface for h2o sparkling water. 2019. R package version 0.2.18. Available at https://cran.r-project.org/web/packages/rsparkling/index.html (accessed 10 Aug 2019).

Hawbaker, T.J., N.S. Keuler, A.A. Lesak, T Gobakken, K. Contrucci, and V.C. Radeloff. 2009. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. Journal of Geophysical Research 114: G00E04.

Hepp, T., M. Schmid, O. Gefeller, E. Waldmann, and A. Mayr. 2016. Approaches to regularized regression – a comparison between gradient boosting and the lasso. Methods Inf Med 55: 422-430.

Heidemann, H. K., 2018, Lidar base specification (ver. 1.3, February 2018): U.S. Geological Survey Techniques and Methods, book 11, chap. B4, 101 p., https://doi.org/10.3133/tm11b4

Hermosilla, T., M.A. Wulder, J.C. White, N.C. Coops, and G.W. Hobart. 2015. Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. Remote Sensing of Environment 170: 121-132.

Hijmans, R.J. et al. raster: geographic data analysis and modeling. 2019. R package version 2.9-23. Available at; https://cran.r-project.org/web/packages/raster/index.html (accessed 10 Aug 2019).

Hu, T., Y. Su, B. Xue, J. Liu, X. Zhao, J. Fang, and Q. Guo. 2016. Mapping global forest aboveground biomass with spaceborne LiDAR, optical imagery, and forest inventory data. Remote Sensing 8: 565.

Huang, C., B.K. Wylie, L. Yang, C. Homer, and G. Zylstra. 2002. Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. International Journal of Remote Sensing 23: 1741-1748.

Huang, Y., B. Yu, J. Zhou, C. Hu, W. Tan, and J. Wu. 2013. Toward automatic estimation of urban green volume using airborne LiDAR data and high resolution remote sensing images. Frontiers in Earth Sciences 7: 43-54.

Hudak, A.T., M.A. Lefsky, W.B. Cohen, and M. Berterretche. 2002. Integrationg of lidar and Landsat ETM+ data for estimating and mapping forest canopy height. Remote Sensing of Environment 82: 397-416.

Hyde, P., R. Dubayah, W. Walker, J.B. Blair, M. Hofton, and C. Hunsaker. 2006. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. Remote Sensing of Environment 102: 63-73.

Ichii, K., A. Kawabata, and Y. Yamaguchi. 2002. Global correlation analysis for NDVI and climatic variables and NDVI trends: 1982-1990. International Journal of Remote Sensing 23: 3873-3878.

Jakubowski, M.K., Q. Guo, and M. Kelly. 2013. Tradeoffs between lidar pulse density and forest measurement accuracy. Remote Sensing of Environment 130: 245-253.

Kauth, R.J. and G.S. Thomas. 1976. The tasseled cap-a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. InLARS Symposia p.159.

Keane, R.E., E.D. Reinhardt, J. Scott, K. Gray, and J. Reardon. 2005. Estimating forest canopy bulk density using six indirect methods. Canadian Journal of Forest Research 35(3): 724-739.

Kennedy, R.E., Z. Yang, and W.B. Cohen. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr – Temporal segmentation algorithms. Remote Sensing of Environment 114: 2897-2910.

Key, C.H. and N.C. Benson. 2003. The normalized burn ratio (NBR): a Landsat TM radiometric measure of burn severity. US Geological Survey Northern Rocky Mountain Science Center, Bozeman, MT.

LaRue, E.A., J.W. Atkins, K. Dahlin, R. Fahey, S. Fei, C. Gough, B.S. Hardiman. Linking Landsat to terrestrial LiDAR: vegetation metrics of forest greenness are correlated with canopy structural complexity. International Journal of Applied Earth Observation and Geoinformation 73: 420-427.

Lawrence, R.L. and C.J. Moran. 2015. The AmericaView classification methods accuracy comparison project: a rigorous approach for model selection. Remote Sensing of Environment 170: 115-120.

LeDell et al. h2o: R interface for 'H2O'. 2019. R package version 3.26.02. Available at https://cran.r-project.org/web/packages/h2o/index.html (accessed 19 Aug 2019).

Lefsky, M.A., W.B. Cohen, D.J. Harding, G.G. Parker, S.A. Acker, S.T. Gower. 2002. Lidar remote sensing of above-ground biomass in three biomes. Global Ecology and Biogeography 11: 393-399.

Lefsky, M.A., D.P. Turner, M. Guzy, and W.B. Cohen. 2005. Combining lidar estimates of aboveground biomass and Landsat estimates of stand age for spatially extensive validation of modeled forest productivity. Remote Sensing of Environment 95: 549-558.

Liang, J., D.E. Calkin, K. M. Gebert, T.J. Venn, and R.P. Silverstein. 2008. Factors influencing large wildland fire suppression expenditures. International Journal of Wildland Fire 17: 650-659.

Lim, K., P. Treitz, M. Wulder, B. St-Onge, and M. Flood. 2003. LiDAR remote sensing of forest structure. Progress in Physical Geography 27: 88-106.

Lundquist, J.D., N. Pepin, C. Rochford. 2008. Automated algorithm for mapping regions of cold-air pooling in complex terrain. Journal of Geophysical Research-Atmospheres 113: 1984-2012.

Luraschi et al. sparklyr: R interface to Apache Spark. 2019. R package version 1.0.2 Available at https://cran.r-project.org/web/packages/sparklyr/index.html (accessed 10 Aug 2019).

Luther, J.E., R.A. Fournier, O.R. van Lier, and M. Bujold. 2019. Extending ALS-based mapping of forest attributes with medium resolution satellite and environmental data. Remote Sensing 11: 1092.

Lutz, J.A., J.W. van Wagtendonk, and J.F. Franklin. 2010. Climatic water deficit, tree species ranges, and climate change in Yosemite National Park. Journal of Biogeography 37: 936-950.

Makela, H. and A. Pekkarinen. 2004. Estimation of forest stand volumes by Landsat TM imagery and stand-level field-inventory data. Forest Ecology and Management 196: 245-255.

Maltamo, M., O.M. Bollandsas, T. Gobakken, and E. Naesset. 2016. Large-scale prediction of aboveground biomass in heterogeneous mountain forests by means of airborne laser scanning. Canadian Journal of Forest Research 46: 1138-1144.

Margolis, H.A., R.F. Nelson, P.M. Montesano, A. Beaudoin, G. Sun, H. Andersen, and M.A. Wulder. 2015. Combining satellite lidar, airborne lidar, and ground plots to estimate the amount and distribution of aboveground biomass in the boreal forest of North America. Canadian Journal of Forest Research 45: 838-855.

Masek, J.G., E.F. Vermote, N.E. Saleous, R. Wolfe, F.G. Hall, K.F. Huemmrich, F. Gao, J. Kutler, and T-K. Lim. 2006. A Landsat surface reflectance dataset for North America, 1990-2000. IEEE Geoscience and Remote Sensing Letters 3: 68—72.

Matasci, G., T. Hermosilla, M.A. Wulder, J.C. White, N.C. Coops, G. W. Hobart, and H.S.J. Zald. 2018. Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using Landsat composites and lidar plots. Remote Sensing of Environment 209: 90-106.

McGaughey, R.J. 2015. FUSION v3.5. USDA Forest Service, Pacific Northwest Research Station, Olympia, WA. http://forsys.cfr.washington.edu/

McRoberts, R.E., E. Tomppo, E. Naesset. 2010. Advances and emerging issues in national forest inventories. Scandinavian Journal of Forest Research 25: 368-381.

Moran, C.J., E. Rowell, and C.A. Seielstad. 2018. A data-driven framework to identify and compare forest structure classes using LiDAR. *Remote Sensing of Environment* 211: 154-166.

Mousivand, A., M. Menenti, B. Gorte, and W. Verhoef. 2014. Global sensitivity analysis of spectral radiance of a soil-vegetation system. Remote Sensing of Environment 145: 131-144.

Natekin, A., and A. Knoll. 2013. Gradient boosting machines, a tutorial. Frontiers in Neurorobotics 7: 1-21.

Noonan-Wright, E.K., T.S. Opperman, M.A. Finney, G.T. Zimmerman, R.C. Seli, L.M. Elenz, D.E. Calkin, and J.R. Fielder. 2011. Developing the US Wildland Fire Decision Support System. Journal of Combustion 168473.

Ohmann, J.L., M.J. Gregory, and H.M. Roberts. 2014. Scale considerations for integrating forest inventory plot data and satellite image data for regional forest mapping. Remote Sensing of Environment 151: 3-15.

Pascual, C., A. Garcia-Abril, W.B. Cohen, S. Martin-Fernandez. 2010. Relationship between LiDAR-derived forest canopy height and Landsat images. International Journal of Remote Sensing 31: 1261-1280.

Peterson, B., K.J. Nelson, C. Seielstad, J. Stoker, W.M. Jolly, and R. Parsons. 2015. Automated integration of lidar into the LANDFIRE product suite. Remote Sensing Letters 6: 247-256.

Reeves, M.C., K.C. Ryan, M.G. Rollins, and T.G. Thompson. 2009. Spatial fuel data products of the LANDFIRE project. International Journal of Wildland Fire 18: 250-267.

Rollins, M.G., B.C. Ward, G. Dillon, S. Pratt, and A. Wolf. 2007. Developing the LANDFIRE fire regime data products. Available at: https://landfire.cr.usgs.gov/documents/Developing_the_LANDFIRE_Fire_Regime_Data_P roducts.pdf Accessed July 20, 2019.

Rollins, M.G. 2009. LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment. International Journal of Wildland Fire 18: 235-249.

Rouse, JW, RH Haas, JA Scheel, and DW Deering. 1974. Monitoring vegetation systems in the great plains with ERTS. *Proceedings of the 3rd Earth Resource Technology Satellite Symposium* p.48-62.

Roy, D.P., V. Kovalskyy, H.K. Zhang, E.F. Vermote, L. Yan, S.S. Kumar, and A. Egorov. 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. Remote Sensing of Environment 185: 57-70.

Ryan, K.C. and T.S. Opperman. 2013. LANDFIRE – a national vegetation/fuels data base for use in fuels treatment, restoration, and suppression planning. Forest Ecology and Management 294: 208-216.

Tomppo, E., H. Olsson, G. Stahl, M. Nilsson, O. Hagner, and M. Katila. 2008. Combining national forest inventory field plot and remote sensing data for forest databases. Remote Sensing of Environment 112: 1982-1999.

Torgo, L., R. P. Ribeiro, B. Pfahringer, and P. Branco. 2013. Smote for regression. Portuguese Conference On Artificial Intelligence. Springer, Berlin, Heidelberg. Pp. 378-389.

Schoennagel, T., J.K. Balch, H. Brenkert-Smith, P.E. Dennison, B.J. Harvey, M.A. Krawchuk, N. Mietkiewicz, P. Morgan, M.A. Moritz, R. Rasker, M.G. Turner, and C. Whitlock. 2017. Adapt to more wildfire in western North American forests as climate changes. Proceedings of the National Academy of Sciences 114: 4582-4590.

Schmidt, G.L., C.B. Jenkerson, J. Masek, E. Vermote, and F. Gao. 2013. Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm description: US Geological Survey Open-File Report 2013-1057, 17 p.

Sorenson, R., U. Zinko, and J. Seibert. 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. Hydrology and Earth System Sciences Discussion, European Geosciences Union 10: 101-112.

Stojanova, D., P. Panov, V. Gjorgjioski, A. Kobler, and S. Dzeroski. 2010. Estimating vegetation height and canopy cover from remotely sensed data with machine learning. Ecological Informatics 5: 256-266.

Stratton, R.D. 2009. Guidebook on LANDFIRE fuels data acquisition, critique, modification, maintenance, and model calibration. Gen. Tech. Rep. RMRS-GTR-220. Fort Collins, CO: US Dept of Ag, Forest Service, Rocky Mountain Research Station, 54 p.

Wickham, H. et al. dplyr: a grammar of data manipulation. 2019. R package version 0.8.3. Available at: https://cran.r-project.org/web/packages/dplyr/index.html (accessed 10 Aug 2019).

Wickham, H. et al. ggplot2: create elegant data visualisations using the grammar of graphics. 2019. R package version 3.2.1. Available at: https://cran.r-project.org/web/packages/ggplot2/index.html (accessed 10 Aug 2019).

Wieinmyer, C. and M.D. Hurteau. 2010. Prescribed fire as a means of reducing forest carbon emissions in the western United States. Environmental Science Technology 44: 1926-1932.

Wilkes, P., S.D. Jones, L. Suarez, A. Mellor, W. Woodgate, M. Soto-Berelov, A. Haywood, and A.K. Skidmore. 2015. Mapping forest canopy height across large areas by upscaling ALS estimates with freely available satellite data. Remote Sensing 7: 1-x.

Williams, J. 2013. Exploring the onset of high-impact mega-fires through a forest land management prism. Forest Ecology and Management 294: 4-10.

Wulder, M.A., R.J. Hall, N.C. Coops, and S.E. Franklin. 2004. High spatial resolution remotely sensed data for ecosystem characterization. BioScience 54: 511-521.

Zald, HSJ, MA Wulder, JC White, T Hilker, T Hermosilla, GW Hobart, and NC Coops. 2016. Integrating Landsat pixel composites and change metrics with lidar plots to predictively map forest structure and aboveground biomass in Saskatchewan, Canada. *Remote Sensing of Environment* 176: 188-201.

Zaharia, M., R.S. Xin, P. Wendell, T. Das, M. Armbrust, A. Dave, X. Meng, J. Rosen, S. Venkataraman, M.J. Franklin, and A. Ghodsi. 2016. Apache spark: a unified engine for big data processing. Communications of the ACM 59: 56-65.

Chapter 4. Deriving Fire Behavior Metrics from UAS Imagery

1) Introduction

Although laboratory and modeling approaches offer greater control of environmental conditions and better replicability than is possible in field settings, empirical studies of fire behavior remain important to a range of scientific inquires [1–4], including supporting theory of fire spread dynamics [5,6], evaluation/validation of mathematical fire models [7–10], and assessments of fire and other disturbance interactions [11,12], fuel treatment design and effectiveness [13–15], and fire behavior process-vegetation pattern relationships [16–18]. In each of these areas of study, field observations and experiments provide opportunities to quantify actual fire behavior providing validation of numerical experiments in an intellectual environment that still prefers real-world corroboration of modeling research [6,19]. Remote sensing from handheld devices, tripods, manned aircraft, satellite, and, increasingly, unmanned aerial systems (UAS), are fundamental to the systematic measurement of fire behavior in the field [20].

The most common fire behavior metrics that are derived from empirical studies can be grouped into two distinct types: (1) energy release as characterized by intensity and/or power [21] and (2) fire movement such as rate of spread (RoS), spread direction, and residence time. Measurements of fire energy have significant scientific legacy from a multitude of small-scale laboratory experiments [22–24], in-situ field observations [25–27], and regional to global assessments of satellite imagery [28–31]. Significant limitations and assumptions are inherent in the estimates of fire radiative energy (FRE), fire radiative power (FRP), and fire intensity, primarily related to difficulties in collecting data over the full temporal range of combustion [32–34]. However, the performance of these metrics is well-documented in the literature, and many of the caveats are well-summarized in Table 6 of Hudak et al. [35].

Derivations of fire movement metrics are less common in the literature, but they are experiencing a renaissance of sorts [36–39]. Fire RoS, an essential metric for characterizing fire intensity [40], and more generally fire behavior, has become obtainable at high temporal and spatial resolutions (~0.1–1 Hz and 1–50 cm, respectively) from UAS. The ability of UAS to collect data from new perspectives and variable scales has created a need for more robust algorithms to characterize fire progression and RoS. Such fire movement data are increasingly valuable for characterizing relationships between fire behavior and environmental patterns in fuels, moisture, wind, and topography, as well as in providing better estimates of fire intensity and real-time data for managing wildfires.

UAS confer a fresh, dynamic, and relatively safe and inexpensive perspective for studying wildland fire in field settings [4,41,42]. A rotor-wing UAS hovering above a fire provides a static, scalable scientific measurement platform, with some advantages over other systems. For example, fixed-wing aircraft must be moving, and thus face challenges in providing spatially and temporally consistent measurements [39]. Helicopters have the same hovering capabilities with longer flight durations [36,43,44], but they come with high operating costs, complex logistics, increased risk, and the potential to produce rotor wash that affects fire behavior. Boom lifts [27], towers [45], large tripods [46], or trees can raise

sensors to overhead perspectives but constrain the spatial extents of measurement and must be resistant to high temperatures.

There has been growing interest in the capabilities of UAS as a fire research tool, but their limitations are still not widely known [41]. Hardin et al. [47] outline the six primary challenges to using UAS for environmental remote sensing. Relatively short flight times (~10–30 min) for small rotor-wing UAS is the primary challenge and it largely determines mission parameters, such as launch and landing locations and the amount of time for uninterrupted data collection. Payload limits restrict sensor weight and, often as a result, sensor performance capabilities (e.g., uncooled microbolometers vs. cooled shortwave thermal cameras). Visual line-of-sight rules limit operating extents in many locations and further constrain launch and landing locations [41]. Consequently, pilots, visual observers, and researchers are constantly at risk of interfering with fire operations. UAS operations have not yet been fully integrated into fire operations either, and most fire personnel have little experience working with UAS when compared to traditional manned aircraft [48]. All of these factors increase the operational complexity in an already complex fire environment. Deploying UAS in prescribed fire settings simplifies operations somewhat compared to operating in uncontrolled wildfire situations, enhances research efficacy, and exposes scientists and fire managers to strengths and weaknesses associated with UAS use more broadly.

UAS are also proving useful for collecting vegetation data, resulting in opportunities for acquiring coincident fuels-fire behavior datasets with increasingly finer grains and larger spatial extents. Relatively cheap UAS hardware combined with structure-from-motion photogrammetric techniques can be used to build detailed three-dimensional (3-D) models of the environment [49–51]. Lidar systems, which are being mounted to UAS platforms, have also produced 3-D datasets of similar quality and detail [4,52]. Thus, the analysis of pattern-process ecological relationships is now possible in unprecedented detail and extent. There are at least two major and interrelated concerns for these types of pattern-process analyses. First, the size of analytical units becomes an important consideration [3]. Point cloud data are often summarized over two-dimensional (2-D) pixels or 3-D voxels with a number of metrics that characterize the arrangement or presence/absence of points over defined areas or volumes [52,53]. The choice of area or volume of these units can significantly affect results [54]. Secondly, and specific to study presented here, the estimation of fire movement suffers from a lack of data between consecutive images. This can lead to a violation of statistical assumptions, particularly sample independence, and thus the determination of appropriate analytical units is necessary.

Fire has inherent spatial and temporal dependence as a self-perpetuating chemical reaction [55], which should be accounted for or leveraged in analyses. In particular, for RoS, if the values are attributed to individual pixels at the original resolution of the imagery, interpolation is necessary between flame fronts locations. These are essentially redundant data introducing artificial spatial autocorrelation (SA), which confounds the true SA signal. This effect is spatially variable, as missing data increases when fire covers a larger area between images (i.e., higher RoS). Statistical inference and model performance will undoubtedly be affected by inflating independent sample sizes (i.e., pseudo-replication) and the aforementioned confounding of the true SA signal [56]. This general issue is coined the 'change of support problem' or 'modifiable areal unit problem' (MAUP), where

'support' refers to the geometrical size, shape, and spatial orientation of the region associated with the value or measurement of interest [57,58].

Here, in conjunction with estimating fire spread metrics from UAS imagery, we introduce a method to derive more robust analytical units, while also maintaining inherent SA, in order to improve pattern-process analyses, such as fuel-fire behavior interactions. Our approach is methods-oriented with the intent of sharing techniques in sufficient reproducible detail, and with the expectation that fire research with UAS will continue to develop and expand in use. Our paper addresses the following specific objectives:

1. Collect spatially and temporally consistent images of free-burning fire using UAS-borne thermal IR, NIR, and visible cameras on spatial domains of 0.01–1 Ha.
2. Generate image time-series with known radiometric and geometric fidelity.
3. Extract fire progression, rate of spread (RoS), and spread direction in complex fire behavior (e.g., interacting fire lines, multiple heads) using automated methods.
4. Determine appropriate analytical units from fire behavior data to improve statistical analysis of fire-environment interactions (e.g., fire behavior-fuels).
5. Share lessons-learned with the intent of flattening the learning curve for others adopting UAS technology for wildland fire operational and research use.

2) Materials and Methods

The following methods were derived from UAS deployments on a dozen prescribed fires in the states of GA, FL, MT, and OR, USA. Here, we focus on data collected from seven field plots in three prescribed fires conducted in dry ponderosa pine forests of western MT and southern OR. Five plots were located at the University of Montana's Lubrecht Experimental Forest (referred to as Lubrecht henceforth) in May 2017. Two plots were located at The Nature Conservancy's Sycan Marsh Preserve (Sycan henceforth) in October 2018. Plot dimensions for the Lubrecht experiments were 10×10 m. The Sycan experiments expanded plot sizes to 100×100 m. Data were collected by positioning a UAS with nadir-viewing cameras above plot center and imaging fire as it burned through a plot. We include relevant background and literature pertinent to the specific methods below.

2.1. UAS Platforms and Sensors

After testing many rotor-wing UAS and sensors (e.g., platforms: DJI (Shenzhen, GD, China) Phantom, Phantom Pro, M600, M100; Skyfish (Missoula, MT, USA) M4; ICI (Beaumont, TX, USA) Halo; 3DR (Berkeley, CA, USA) Solo, X8 (DIY); GoPro (San Mateo, CA, USA) Karma; sensors: FLIR (Wilsonville, OR, USA) XT; ICI 8640; ICI SWIR 640; DJI X3, X5; Sony (Tokyo, Japan) A7R, QX1, QX30, A6000; GoPro Hero; MicaSense (Seattle, WA, USA) RedEdge), we selected the DJI Matrice M100 with a dual battery configuration and the FLIR XT and MicaSense RedEdge sensors. The Matrice provided reliability at relatively low cost along with a good balance of size, speed, hovering stability, flight time, payload, software capability and stability, camera integration, availability of spare parts, and ease of repair. A bench-calibrated (prior to both field missions) FLIR XT thermal camera (7.5–13.5 µm spectral band) mounted on a DJI gimbal was used for the Lubrecht experiment. A multispectral Micasense RedEdge camera on a fixed mount (no gimbal) was added for the Sycan burns. These two sensors provide multispectral,

radiometric assessment of fire and vegetation at wavelengths of 0.465–0.86 µm and 7.5–13.5 µm. Table 1 shows the sensor specifications.

2.2. Plot Selection and Layout

Eight ground control points (GCPs) were established at the corners of two nested squares for each plot. At Lubrecht, the outer square had 10x10m dimensions with the inner square at one-third the area (3.33 × 3.33 m). For Sycan, the outer squares were 100×100 m and the inner squares were 10×10 m. The internal geometry of the GCPs was established with measurements from a TruePulse laser rangefinder. GCP locations were geotagged using an Emlid Reach RS rover-base station dGPS setup. Plot layout and UAS camera orientation were consistent to ease interpretation and georectification of imagery. Plots were selected to minimize canopy occlusion but also to represent a variety of western US surface fuels that typically experience low and mixed severity fire. Plots were ignited using a drip torch with the intention of achieving a coherent, steady-state fire front spreading perpendicular to a plot edge. In practice, heterogeneous fuels and shifting wind speed and direction led to ignitions at variable distances and directions from plot edges. For example, fire in Sycan plot 1 failed to carry into the plot, and it had to be re-ignited in receptive fuels within the plot boundaries.

2.3. Data Collection

The UAS hovered at a fixed altitude above plot center with cameras viewing nadir. Table 1 reports the flying altitudes, corresponding pixel sizes, and other relevant data collection parameters.

Flight altitudes were established to capture the extent of the plot with a buffer of 2–5 m at Lubrecht and 10–15 m at Sycan. Data were collected for at least one UAS battery cycle, until active fire spread within the plot ceased, or progression of fire operations necessitated travel to the next plot. Temporal resolution was fixed at 0.2 Hz at Sycan and 1 Hz for Lubrecht, although image capture rates were delayed to 0.13 Hz and slightly variable (SD of 0.13 s) at Lubrecht due to memory card issues. The FLIR XT was the primary sensor for the Lubrecht burn. Both the Micasense RedEdge and FLIR XT were used at Sycan although the RedEdge was the primary camera for RoS derivation.

2.4. UAS Imagery Stabilization and Georectification

An important consideration for deriving useable fire data from UAS is the stabilization of imagery to mitigate platform drift and jitter [59]. Gimbaled sensors reduce these effects, but correction is still necessary. The first step is to establish GCPs that can be identified in different spectral bands with fire in the field of view. Either 'cold' (i.e., low emissivity) or 'hot' (usually charcoal beds) targets relative to background temperatures are options. After extensive experimentation, we settled on low-emissivity, 40×40 cm polished aluminum targets (for Sycan) and 12 cm diameter aluminum foil-wrapped circular targets (for Lubrecht), which proved to be reliably identifiable in visible and infrared imagery. Target visibility was tested at multiple altitudes in sunlit and diffuse lighting as well as in active fire conditions.

Table 1. Sensor and data collection information for the two prescribed burns (Lubrecht and Sycan) in this study.

Sensor	Spectral Band for Fire Behavior	Radiometric Temperature Range [1]	Array Size	Plot	Altitude Above Ground	Ground Sample Distance	Temporal Resolution	Time Series Length
(focal length)	(μm)	(°C)	(pixels)	Name	(m)	(cm)	(Hz)	(s)
FLIR XT (9 mm) [2]	7.5–13.5	100–1100	640 × 512	Lubrecht 1	19	3.59	0.13	725
FLIR XT (9 mm) [2]	7.5–13.5	100–1100	640 × 512	Lubrecht 2	20	3.78	0.13	1216
FLIR XT (9 mm) [2]	7.5–13.5	100–1100	640 × 512	Lubrecht 3	20	3.78	0.13	1449
FLIR XT (9 mm) [2]	7.5–13.5	100–1100	640 × 512	Lubrecht 4	18.5	3.49	0.13	1014
FLIR XT (9 mm) [2]	7.5–13.5	100–1100	640 × 512	Lubrecht 5	18	3.41	0.12	928
FLIR XT (9 mm)	7.5–13.5	100–190 [3]	640 × 512	Sycan 1	120	22.7	0.2	250
MicaSense RedEdge (5.4 mm)	0.82–0.86	650–1150 [4]	1280 × 960	Sycan 1	120	8.2	0.2	250
MicaSense RedEdge (5.4 mm)	0.82–0.86	650–1150 [4]	1280 × 960	Sycan 2	180	12.5	0.2	240

[1] Tested on a Mikron M300 blackbody calibration source with a 100–1150 °C temperature range; [2] Higher temperature range achieved using a custom neutral-density filter; [3] High gain setting without neutral-density filter saturated pixels at 190 °C; [4] Lower limit determined as blackbody radiance became clearly distinguishable from background radiance in a laboratory setting.

Two basic methods were evaluated to create co-incident, georectified imagery: (1) georectification alone, which effectively corrects image jitter while rectifying, but requires each image to be separately processed and (2) image stabilization of the time series, followed by separate georectification. Georectification alone may seem faster, but automated GCP and tie-point locators struggle with the rapid change between images due to fire spread. This created the need for manual identification of GCPs and other visible tie-points in every individual image necessitating significant time and resources. GCPs were also obscured by fire and smoke in some images, which led to inconsistent results across the time series. The use of image stabilization algorithms, specifically those designed for video stabilization, resolved these issues [59]. We used the warp stabilizer algorithm within Adobe After Effects (version 15.1.2) for this purpose, although open-source variants are also available. By locating objects that are relatively invariant throughout a time-series or only using the previous image to identify tie-points, image stabilization algorithms efficiently produce image stacks that need GCPs to be visible in only one image (often pre- or post-fire) for effective georectification. The downsides of video stabilization techniques are their tendency to transform data to into video compliant data types (8 or 16 bits), which often involve data scaling.

2.5. Data Pre-Processing

Before calculating fire behavior metrics, aligned imagery was converted to radiance (Micasense RedEdge) and radiant temperature (FLIR XT) and organized in raster stack form. The FLIR XT sensor requires proprietary software to extract radiometric data and calculates radiant temperature based on calibrated equations specific to the sensor. Although improvements in radiant temperature estimates can be made using a set of user inputs, the inputs can be temporally and spatially variable (e.g., emissivity) [60] and subject to radiometric saturation. Thus, accurate spatial calibration is unlikely. We used default values besides ambient temperature and relative humidity, which were measured at take-off for each plot. An emissivity constant of 0.98 was utilized. Calibration of imagery from the Micasense RedEdge is integrated into several software packages (e.g., Agisoft Metashape), although its use for observing fire requires customized scripts. A GitHub webpage provides Python scripts and tutorials [61]. The RedEdge camera has five distinct sensors for each spectral band, and the image rasters must be aligned post collection. While Micasense provides an automated alignment script, the results proved inconsistent. We used the GCPs to create an alignment model using an affine transformation within ERDAS Imagine software (version 16.5.0) to achieve satisfactory band-to-band alignment. We then followed Micasense's algorithms for radiometric calibration and calculation of radiance for each spectral band.

2.6. Flaming Combustion Determination

Characterizing the movement of flaming combustion begins with a definition of flaming, as measured by the sensors employed. Most commonly, raw digital numbers from the sensor are converted to radiance and then transformed to radiant temperature. A static threshold is then applied to the radiant temperatures for a binary classification. The value of this temperature threshold varies considerably in the literature from 150 °C [44], 326.85 °C

[36], 426.85 °C [43], 499.85 °C [38,40], and the Draper point at 525 °C [62]. Another approach is to define the flaming front using edge detection algorithms, which rely on gradient change in the imagery [37,45,63]. Major advantages of edge detection approaches are the ability to apply them to image data that do not have robust radiant temperature transformations available and in situations where absolute values of radiance may be affected, such as smoky atmospheric conditions. With the rapid development of UAS-specific sensors, we prefer the gradient approaches, as datasets are likely to be variable in scale, resolution, and sensor-specific parameters, such as spectral bands and sensitivity. Theoretically, one gradient-based approach could be applied to a variety of imagery and yield similar results, whereas methods reliant on radiant temperature thresholds would likely have to be customized for each sensor and perhaps for different spatial resolutions as well.

However, gradient-based approaches typically define only the leading flame front edge. As the active flame front is often discontinuous and complex in heterogeneous fuels, the edge-smoothing needed to account for gaps, such as in Ononye et al. [63], simplifies subsequent fire behavior calculations and risks reducing the variability seen at the fine scales capable from UAS perspectives. Additionally, fire spread into areas not ignited by the initial flaming front will often be ignored in these gradient-based approaches.

To overcome the limitations of both methods, we combined temperature threshold and edge detection techniques to maintain observed variability and to maximize indifference to resolution, sensor, and spectral band differences. First, edge detection following the Canny method was applied to each thermal image [64]. Valero et al. [37] showed that the Canny method discriminates between the flaming front, transient flames, and pre-frontal heat. Instead of directly using these edges, we extracted pixel values along the defined flame front edge from the Canny algorithm and applied a two class k-means clustering algorithm to automatically determine the flaming combustion threshold [65].

From one perspective, determination of flaming combustion (and any binary classification) requires optimizing the balance between errors of commission (false positives) and omission (false negatives) [66]. Errors of commission are abundant in the pixels that immediately precede the flaming front due to preheated fuel and soot particles. Errors of omission are high in situations with high RoS, low residence time, or occlusion by vegetation or smoke. In our study, for example, areas of sparse grass and litter overlaying rock burned quickly but also cooled quickly once the flame front passed. This required logic to attribute pixels that clearly burned in the flaming front but were never detected above the flaming combustion threshold (omission error). Similarly, if temperature thresholds are set too high, omission error rates increase leading to erroneously discontinuous flaming fronts. In such cases, fire spread direction loses coherency, and estimates of RoS will erroneously decrease. Alternatively, low thresholds cause high rates of commission error, which increase RoS, reduce variability in spread direction, and incorporate unburned pixels.

We tested a dual-threshold approach to account for these situations: one to define the leading edge weighted towards a low error of commission and another weighted to minimize error of omission. This approach allows for the use of data from one or multiple sensors and a single spectral band or a split window using different spectral bands. In our experiments at Sycan, we erroneously acquired data from the FLIR-XT in the high-gain

95

setting for plot 1, which saturated the pixels at 190 °C. This plot also had the grass-overlying-rock fuel arrangement causing the residence times to be shorter than the temporal resolution of the imagery. We thus used the Micasense NIR (0.82–0.86 µm) radiance values to define the leading flame edge through Canny gradient detection and k-means clustering. Within each particular timestep polygon, any pixels that did not reach the threshold were then compared to the FLIR XT data. If at the lower threshold (190 °C), the pixels were assumed to have experienced flaming combustion at that particular timestep. A single threshold was applied as defined by the k-means clustering for the other plots (Lubrecht plots 1–5 and Sycan plot 2), since the residence times were consistently longer than the temporal resolution of the imagery.

2.7. Fire Progression

Each pixel in the time-series stack was assigned the timestep at which flaming combustion initially occurred after a temperature threshold was applied individually to each image. This is termed a fire progression or time-of-arrival data layer, which is the basis for all subsequent fire behavior calculations presented in this study.

2.8. Spread Rate and Direction

Of all the metrics based on fire movement, RoS and spread direction may have the most utility for characterizing fire behavior. RoS can be estimated in the field, is commonly used, and exerts strong influence on the estimation of fire intensity [40]. Johnston et al. [38] showed that sensor and algorithm choice have significant influence on these estimates. Regardless of how flaming combustion is determined, the most popular approaches to estimate RoS utilize vector contours with movement speed and direction estimated from straight lines perpendicular to adjacent contours. In cases where threshold temperatures are used, contours are derived from the progression map, while edge detection approaches inherently create contour vectors. Often, the flaming front is broken and discontinuous and steps are needed to reduce the resulting complex geometry of the resulting vectors. Paugam et al. [36] found that areas of complex vector shapes were difficult to characterize and had to average groups of pixels at the cost of losing data by increasing pixel resolution to 1.44 m from the initial 18 cm. Valero et al. [59] and Onyaye et al. [63] smoothed the contours and applied logic to traverse flame front gaps in order to create continuous vector lines.

Our perspective was to characterize the variability and complexity observed in wildland fire settings and to compliment the diversity of imagery that can be collected with UAS. We formulated a new algorithm building on the previous work discussed above. While we utilize UAS-derived imagery, the approach could be applied to nearly any overhead, coincident time-series imagery. We explicitly consider complex fire behaviors, such as multiple distinct firelines interacting, which are common in heterogeneous fuels typical of western US forest ecosystems. The approach is primarily vector-based relying on pairing points defined as lead (i.e., where the fire is heading) or back edge (i.e., where the fire came from). Figure 1 illustrates the algorithm. Scripting was completed using R statistical software [67], with the additional packages 'raster' [68], 'sp' [69], 'rgeos' [70], and 'circular' [71]. First, all spatially connected pixels of the same arrival time were grouped into individual polygons. Each of these progression polygons can be assumed to be

the actual native sampling resolution, which is determined by the spatial and temporal resolution of the sensor and the RoS. For example, we observed several distinct, elliptical-shaped fire heads in our plots at Lubrecht and Sycan plot 1, and the polygons describing them are much larger than those for the flanking and backing portions of the fire (e.g., Figure 2 Box E). The statistical implications are important and are discussed below.

We regularly sampled along each progression polygon edge to create a series of focal points, similar to Onyaye et al. [63], but at regular spacing (with random start locations) rather than select specific locations. Paugam et al. [36] used a similar approach, but our pairing of the back and lead edge points overcomes problems with perpendicular lines not representing the fire spread direction when fire edges create complex polygon shapes. We also smoothed vector lines just enough to remove the jagged edges created by the square shape of the pixel, which are artifacts of the sensor array (Raster to Polygon tool with the 'simplify polygons' option selected, ArcGIS 10.6.1). Using regularly spaced points introduces a sampling bias, but we iterated the algorithm twenty times with random start locations for the points and averaged the results.

Each focal point was defined as either a lead or back edge point. This was determined by examining the adjacent polygon to the point and assessing whether it burned before or after in time. Leading edge points were connected with lines to back edge points with the logic demonstrated in Figure 1. Back edge points can be connected to multiple lead edge points. This situation represents fire growth in an expanding elliptical pattern, which most spatial fire spread models assume (e.g., FARSITE) [72]. The inverse situation is also possible where the fire closes in from multiple directions resulting in more back edge points than lead edge points. This special situation is highlighted in Figure 1 Box D. Once each point is paired, a straight line connecting the pair was created, which allowed for distance and direction to be determined. The line was buffered with a width of two pixels, and all of the pixels within this area were attributed with the direction and a RoS. The RoS was determined by dividing the length of the line by the length of time between the previous and current images. The buffering of the line leads to some pixels being attributed multiple RoS and directions. In these cases, we attributed the minimum RoS and its associated direction with the assumption that the fire traveled the shortest route.

Another special situation is fire spread of one pixel per timestep, which is not amenable to the vector method above. We borrowed logic from calculation of topographical aspect from digital elevation rasters for these cases. First, we calculated the potential RoS to all the adjacent pixels and decomposed RoS into separate x, y components. The most likely spread direction was taken as the two-argument arc-tangent of the x, y components ('atan2' function) [67]. Therefore, this method is built on the assumption that the direction of fastest spread is the most likely. We could not use the same assumption that the shortest route is most likely here, as only one pixel is traversed per timestep. We then placed a line pointed in this resulting direction, found the intersecting pixel, calculated the pixel centroid to centroid distance, determined the length of time to traverse this line, and divided the distance by time to calculate the RoS. The intersection of the spread direction line can clip the corners of pixels that may not be the pixel with the most likely travel path. Thus, we added 20° to the calculated direction line in both directions, iterated the process, and selected the pixel with the highest resulting RoS as the most likely travel path. As pixel resolutions become larger and fire spread slows, individual pixel movement gains in

frequency. This method becomes increasingly important for the accurate characterization of fire spread in those situations.
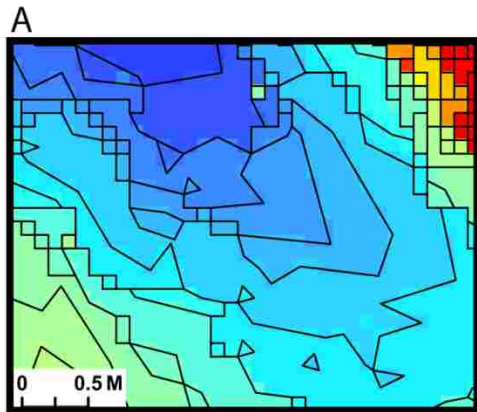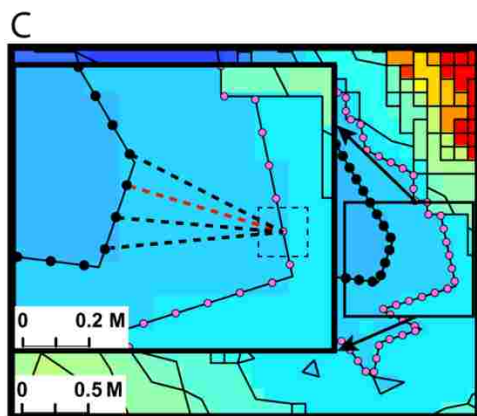

Progression Polygons

Take progression raster layer and dissolve to polygons. Smooth polygons enough to excise at least a portion of square pixel edges. Time progression represented from cooler (blue, earlier) to warmer (red, later) color gradients.
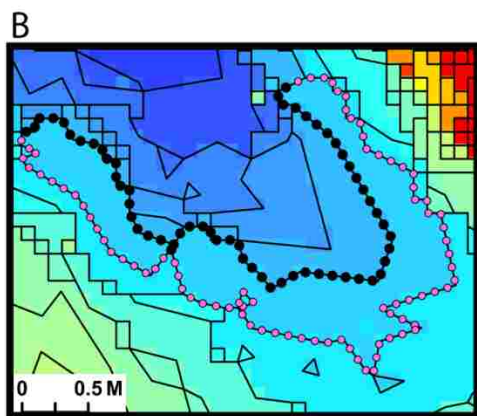
Focal Points

Create regularly spaced (equal to pixel size) focal points along each polygon. Buffer each point to create a square polygon (also the size of a pixel) and intersect to determine timestep of the adjacent polygon. Attribute either lead (pink points) or back edge (black points) to the point if the adjacent polygon is preceding or succeeding in time, respectively.

Lead Edge Points

Calculate a normal angle using the polygon created in the previous step and the two points of intersection with the progression polygon. Calculate the distance and direction to the four closest back edge points and pair the closest point within 30^0 of the normal angle and a length within 1.5 times the distance of the closest point. Reject all points if spread direction defined as $>89^0$ to the normal angle.

Back Edge Points

For points not paired to a lead point, calculate the normal line as above and extend the line until intersected with (1 pixel width buffer, black dotted lines) the progression polygon (solid black line). If a lead point intersects this line, then pair. If not, allow back points to pair. If multiple points intersect, select point with angle closest to the normal angle. Half the attributed distance assuming fire traveled from both back edge points equally (red dotted line and yellow point). Reject line if spread direction defined as $>89.9^0$ to the normal angle.

↓

Per Pixel Rate of Spread

Take each line from previous two steps, buffer (2 pixel buffer), and extract pixels intersecting this buffered line (black squares). Take the distance of the line and divide by length of time between the previous and current timesteps to calculate rate of spread. Due to complex geometry and overlapping buffered lines, pixels may be extracted multiple times. Attribute the minimum rate of spread and associated direction for each pixel.

↓

Single Pixel Rate of Spread

For single pixels, calculate the potential rate of spread to each adjacent pixel (centroid to centroid, black dotted squares) and decompose into x,y gradient components. Take the two-argument arc-tangent of the x,y components to find the direction of spread (red arrowed line). Draw a line along the resulting angle and select the intersecting pixel (red square). Vary this angle by 20^0 in both directions, iterate, and select the intersecting pixel that confers the highest rate of spread.
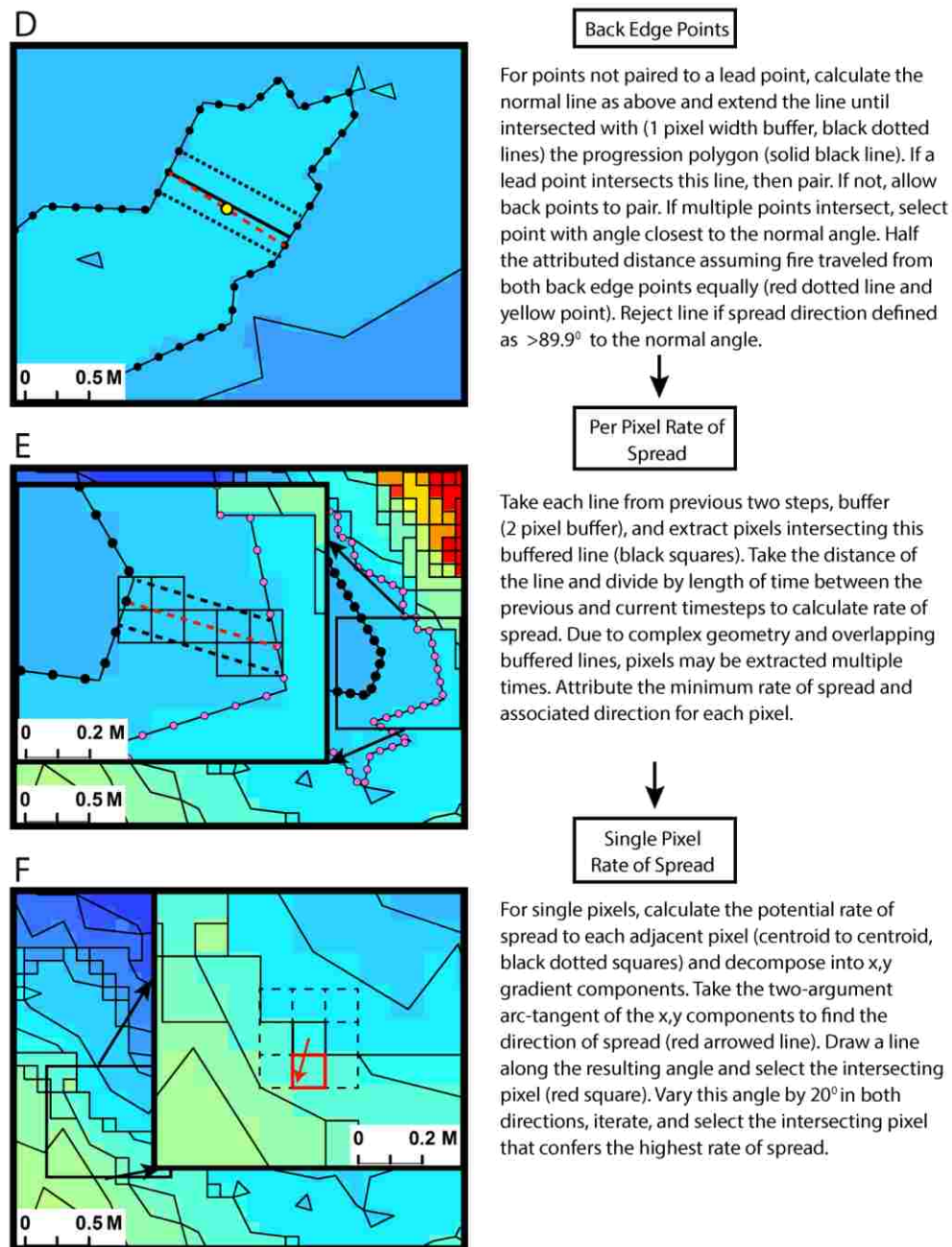
Figure 1. Description of the rate of spread algorithm including situations requiring additional consideration when experiencing complex fire behavior (Box D and F). Illustrated examples taken from Lubrecht plot 5.

2.9. Analytical Units

The geometrical size and shape of analytical units are important considerations [58]. For RoS, the minimum areal unit of measurement could be considered the point-to-point vector line and intersecting pixels (i.e., Figure 1, Box E). However, considerable overlap occurs between these buffered lines and the point-to-point spacing is arbitrary. Conversely, the progression polygons created from the connected timestep pixels could also be considered a measurement unit. However, these polygons can be quite large and incorporate backing, flanking, and heading fire behavior. We sought to aggregate (or

disaggregate depending on the perspective) pixels (or progression polygons) to achieve a medium between these two extremes.

We chose an automated approach using the idea of connected components from graph theory [73,74], hypothesizing that sharp changes in RoS between adjacent pixels delineate the boundaries of relatively independent analytical units. Pixels are considered to be connected when adjacent, within the same timestep, and below a defined change (or tolerance) parameter. The change parameter determines the degree of pixel aggregation. We evaluated a series of values settling on 0.015 m/s for the parameter as reasonable to aggregate to polygons that were relatively homogeneous in terms of the backing, flanking, and heading fire behavior. Tuning this parameter is likely necessary for higher RoS and different degrees of variability though this value was acceptable for the range of RoS observed in this study. We also required a minimum polygon area of 6x6 pixels for progression timesteps with high RoS (polygons sizes > 1 m^2) to prevent individual and small groups of pixels from becoming their own units. In these cases, pixels that displayed large deviations in RoS estimates from neighboring pixels were predominantly artifacts of the algorithm (especially areas of complex geometry) and were assumed to not be different units.

3) Results

We evaluated two cold target sizes as GCPs, circular with 12 cm diameter, and square with 40 cm sides, at six altitudes above ground level (AGL) up to 150 m. For polished aluminum with an approximate emissivity coefficient of 0.04, targets must be roughly 1.1 times the size of the pixel to be reliably visible in optical, thermal IR, and NIR imagery assuming line of sight is maintained. At altitudes of 150 m, a ratio of 1.2 is a safer minimum threshold as atmospheric effects become larger. At nadir and lower AGL (<100 m), targets can be half the size of pixels and still have visibility. However with flames surrounding, we observed targets five times larger than a pixel obscured by heat. Eight ground targets distributed throughout the image in our nested square plot layout provided accurate georectification results following image stabilization techniques (RMSE 0.14–0.28 m for Lubrecht and 0.27–0.94 m for Sycan).

Fire behavior metrics were produced after the images were aligned, calibrated, and georectified. First, fire progression maps from gradient-based threshold techniques were derived (Figure 2, Box A and B). Similar to the results of Valero et al. [37], gradients successfully defined fire movement in thermal imagery. The use of NIR also matched well with optical flame imagery (Figure 2, Box C and D). A 1–3 pixel (8.2–24.6 cm) error of commission was observed at certain portions of the fireline though. This was most conspicuous at the heading portions and along the wind vector (e.g., right side of fireline in Figure 2, Box D). The use of the dual threshold technique filled the burned portions with low residence times and followed the natural flow of the progression (cf. Figure 2, Box A and B), but left unburned areas when visually compared to post-fire imagery (data not shown).

We observed complex fire behavior with multiple heading fires and interacting flanking fires in nearly every plot, despite single strip ignition near each plot. Sycan plot 2 had the most coherent fireline with relatively consistent topography and grass fuels. RoS ranged from 0–2.7 m/s at the two Sycan plots and 0–0.1 m/s for the Lubrecht plots (Figure

3). RoS distributions all had heavy-tailed, positively-skewed distributions with area-weighted mean spread rates of 0.013–0.02 m/s for the Lubrecht plots and 0.305–0.404 m/s for the Sycan plots. Predictably, RoS was highest along the primary vectors of fire travel (heading fire) and lower along the flanks. At Lubrecht, field estimates of RoS were made between the visible targets (GCPs) within each plot. These estimates were significantly correlated with our image-derived measurements (Pearson correlation, $n = 12$, $r = 0.71$, $p < 0.05$). Spread direction followed expectations; though in areas of highly irregular and complex firelines, inconsistencies still emerged despite designs to reduce them.

Spatial aggregation of RoS produced polygon numbers averaging 35.4% of the original pixel count for Lubrecht and 3.1% of the pixels for Sycan (Figure 3), highlighting the large amount of redundant data and the strong influence of spread rate. The areas of inconsistent spread direction due to complex fireline geometry tended to have more polygons as spread rates showed increased variability along sharper gradients within each progression timestep. Aggregated polygons had a mean size of 0.38 m^2 for Sycan and 0.033 m^2 for Lubrecht. The maximum polygon size was 23.6 m^2 for Sycan and 2.8 m^2 for Lubrecht.
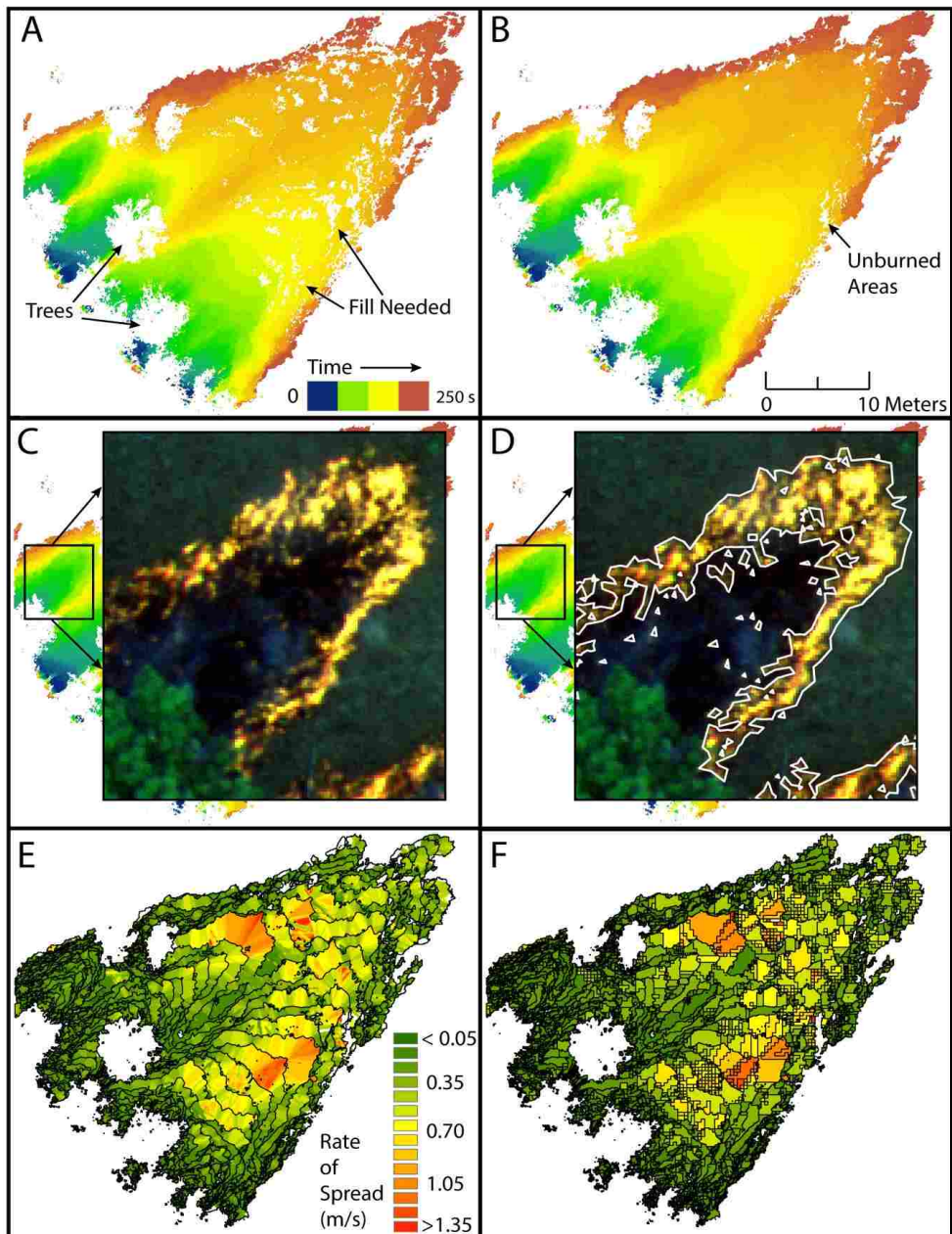
Figure 2. UAS data and derived metrics from Plot 1, Sycan Marsh, OR. (**A**) Fire progression with tree occlusion and additional missing data due to low flame residence time. (**B**) Data gaps filled using dual-threshold flaming combustion definition. (**C**) Optical (RGB) imagery from Micasense RedEdge camera highlighting one of the multiple heading fires within the plot. (**D**) Outline (white lines) of flaming combustion as defined by the method described in Section 2.6. (**E**) Calculated rate of spread delineated by image timesteps (black lines). (**F**) Automated aggregation of polygons to create analytical units.
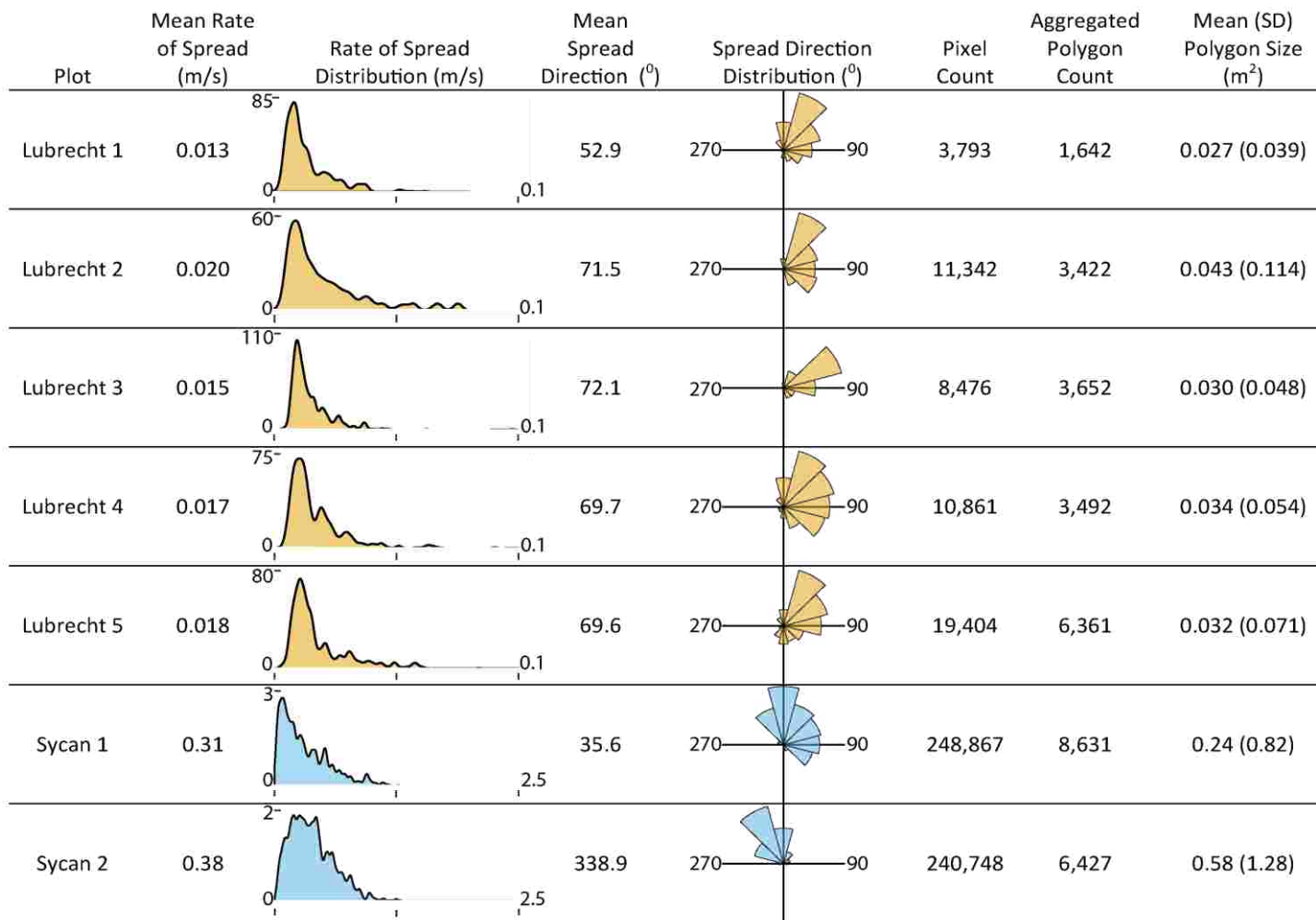
Plot	Mean Rate of Spread (m/s)	Rate of Spread Distribution (m/s)	Mean Spread Direction (^{0})	Spread Direction Distribution (^{0})	Pixel Count	Aggregated Polygon Count	Mean (SD) Polygon Size (m^2)
Lubrecht 1	0.013		52.9		3,793	1,642	0.027 (0.039)
Lubrecht 2	0.020		71.5		11,342	3,422	0.043 (0.114)
Lubrecht 3	0.015		72.1		8,476	3,652	0.030 (0.048)
Lubrecht 4	0.017		69.7		10,861	3,492	0.034 (0.054)
Lubrecht 5	0.018		69.6		19,404	6,361	0.032 (0.071)
Sycan 1	0.31		35.6		248,867	8,631	0.24 (0.82)
Sycan 2	0.38		338.9		240,748	6,427	0.58 (1.28)

Figure 3. Fire spread rate and direction plots from Lubrecht, MT and Sycan Marsh, OR. Pixel count, aggregated polygon count, and mean polygon size shown to highlight the large amount of redundant data and the loss of perceived sample size (pixels vs. aggregated polygons) after removal. All y-axes are scaled independently for ease of viewing and represent probability density functions.

103

4) Discussion

Fire metrics based on fire progression estimates are common in the scientific literature though rarely validated [38]. Our estimates of RoS compared well with visual, one-dimension estimates ($r = 0.71$) and also with mathematical fire model predictions. For example, using observed weather and fuel moisture parameters at Lubrecht and the timber litter and understory fuel model, Rothermel's [75] fire spread equation predicts heading fire spread rates of 0.0045–0.060 m/s, which compares favorably to the 0–0.1 m/s range observed.

Spread direction followed general expectations with heading fire moving in the observed primary spread direction. However, our results show an interesting pattern related to fireline coherency and the relative amounts of flanking and heading fire. As mentioned, Sycan plot 2 had the most coherent fireline and the direction histogram shows the majority of spread in the same direction. In contrast, the majority of Lubrecht plot 5 experienced spread directions that were nearly orthogonal (mean 69.6° from North) to the direction of the heading fires (c. 135°). All of the Lubrecht plots show this spread direction distribution to varying degrees with plot 5 as the most prominent. The discontinuous surface fuels are likely the primary cause of this behavior. The burn occurred in a previously thinned area having numerous stumps, abundant 100-to 1000-hr fuel class (2.54–20.32 cm diameter) woody debris, and skid trails left from mechanized equipment. The heading fires followed paths with sufficient fine fuels and then flanked and backed into areas where the barriers (e.g., larger fuels) had stopped the initial flame front. The associated changes in fire behavior have implications for fire effects [18].

Multiple heading fires also led to multiple firelines interacting, and the design of the algorithm properly identified and estimated RoS in these situations with one significant caveat. Some of these interaction zones, for example, the center of Sycan plot 1 (Figure 2, Box A), had lower RoS than anticipated. The preceding polygons showed a rise in RoS as the two flanking fires interacted, but RoS then decreased at the converging polygon. This artifact arose as a result of the temporal resolution of data capture being considerably slower than the RoS, and from the assumption that the fire spread at an equal rate from each fire line. Although between-image fire behavior can be inferred, assumptions are inherently necessary. For example, total radiant energy or intensity metrics could be used to infer higher than observed RoS or to determine the predominant direction of spread rather than assuming equal rates from both directions.

The area traversed by fire between consecutive images causes a lack of data as fire behavior in the intervening time period is unknown. As RoS is a derivative of fire progression and fire progression is a derivative of the time-series imagery, it is possible and tempting to attribute each pixel from the progression or time-series layers with a RoS value. However, our analysis shows that the majority of these pixels are not measurement units nor statistically independent. The statistical consequences are inflated or introduced spatial autocorrelation and pseudo-replication [56]. The approach proposed here does risk a loss of relevant data if the change parameter is too high in value or RoS is spatially homogenous. This could lead to large analytical units. For example, long, linear fire fronts with consistent RoS could be aggregated to one or a few analytical units within a timestep (we emphasize that the units here were aggregated within

individual timesteps and not beyond). This issue could also be exacerbated by coarse temporal resolution imagery. We focused on creating robust analytical units for subsequent pattern-process relationships. From this perspective, consistent RoS implies either consistent underlying patterns or little influence of pattern on fire processes, either case would be discoverable with analysis using larger analytical units. The analytical unit derivation strategy may need to be reconsidered for other analytical objectives. A potential research direction would be to follow the general approach of Openshaw and Taylor [57] and systematically vary analytical units and assess the effect on resulting spatial relationships. See Gelfand [76] for an in-depth review of existing literature and strategies for the change of the support problem.

We hypothesized that this issue could be leveraged to help inform other data sampling methods. The analytical unit size is dependent on the interaction of RoS and the temporal and spatial resolution of the time-series imagery. Thus, with RoS predictions and known camera parameters, estimates of the analytical unit sizes are possible pre-fire. This information could provide guidance for field sampling protocols, especially plot sizes, and unit sizes for other remotely-sensed data. For example, UAS platforms have enabled increasingly fine-grained pre- and post-fire 3-D vegetation data [50]. The data density from these photogrammetrically-derived point clouds can be hundreds to thousands of points per square meter, and thus the level of spatial aggregation that minimizes the loss of pertinent information while removing redundancy is becoming a research priority. The analytical units of fire behavior are likely to be larger and of variable size when compared to this vegetation data as we show in this study. Ultimately, these datasets will need to be transformed and aligned if characterization of fire behavior and fire environment interaction is the objective. With additional research, derivation of a generalized relationship between RoS and camera parameters and the resulting analytical unit sizes can be useful for the production and analysis of comprehensive datasets.

4.1. Lessons Learned

UAS promise relatively cheap and low risk aviation platforms that provide new remote sensing capabilities to a variety of users. In the course of our experiences over the last few years, we evaluated many off-the-shelf and custom-built UAS. These experiences enable us to provide some suggestions for fire researchers interested in employing UAS in their work. First, we advise to not focus on flight capabilities alone. Flight time, speed, maneuverability, and payload are important, but sensor hardware integration, quality of software, spare parts availability, and ease of repair are equally, if not more, consequential to successful data collection. Custom or boutique UAS often promise improved or specific performance capabilities but frequently lack available spare parts and are often beset with hardware and software issues. Better flight performance and larger payloads also require larger UAS, which are cumbersome to transport, an important consideration in dynamic and time sensitive field settings. We place a premium on small, cheap, reliable systems, which often means working with consumer-grade technology provided by established companies.

We found that software issues were common in nearly all of the systems that we tested to varying degrees. At one point, three different software applications, two being third-party, were required to execute pre- and during-fire data collection within a single mission. How software interacts is unknown after any updates to the myriad of systems, and mission workflows must then be retested in entirety. Automatic software updates are often applied without notification and would frequently change system behavior. Considering the resources needed to plan data collection, coordinate with fire operations, meet acceptable burn conditions (weather, fuel moisture, time of season, etc.), and have needed resources available, the chance of hardware or software failure must be minimized. Often, this level of coordination and timing only comes together for a few days out of one or more years. Thus, the importance of a dependable UAS cannot be understated with emphasis on the system (e.g., platform, sensors, software, planning).

During our prescribed burn campaigns, the UAS experienced extremes ranging from heat from fire, freezing temperatures, heavy smoke and dust, high winds, attacks from territorial birds, and nearly all forms of precipitation. Given the likelihood of such situations in field conditions, we recommend building redundancy, such as having duplicate UAS available with exactly replicated hardware and software configurations. We also recommend platforms with weather resistance. The sensors described in this study cost more than the UAS and, if cost is an issue, we recommend having duplicate platforms over sensors as the probability of platform malfunction is usually greater.

For the use of UAS in prescribed fire experiments, plot locations must strike a balance of achieving desired experimental results while also operating within the context of prescribed burn operations. Field-based fire research generally needs to be opportunistic and in harmony with burn objectives. The majority of prescribed burn units in the US are fired in one day, and coordination and timing with the operations team throughout the day is of utmost importance. Unless multiple UAS teams are available, each research plot must be positioned in order to allow enough time to collect data at a plot, disassemble, and then reassemble at a new launch location before burn operations are ready to fire the next plot. Alternatively, it may be logical to fly multiple plots from one location. This strategy can significantly reduce the flight time for data collection though. We chose to travel the fire perimeter to locations near each plot where the UAS was viewable by the pilot. The pre-planning for these logistics can be complex as operational firing plans change based on weather conditions, particularly wind directions. Plot layouts designed for a single spread direction due to fuel or topography arrangement often do not follow expectations if winds shift. For example, a plot could be burned at either the start or end of burn operations, depending on wind directions, with large subsequent changes in fire behavior. Successful implementation of UAS in prescribed burn experiments requires flexibility to weather conditions while also ensuring that research efforts do not impede firing or holding operations.

In our study, plots were selected and fired to reach steady state fire spread and produce a coherent fireline before reaching the plot edge. In practice, heterogeneous fuels and shifting winds meant ignition at variable distances and directions in relation to the plots, including

through the plot in one instance (Sycan plot 1). Timing UAS takeoff in conjunction with ignition in these conditions proved to be difficult, requiring consideration of different tradeoffs. If takeoff is too early, then battery life will be expended without gathering useful data; conversely, if takeoff is too late, fire may enter the plot before data collection begins. The time to prepare a UAS for launch is also variable depending on multiple devices booting, cameras initializing, and achieving a GPS fix. A vantage point where the UAS pilot can view burn operations is optimal for timing, but often not likely due to fire, smoke, and tree interference. Prompt, succinct, and informative radio communication is vital, which usually necessitates a radio liaison communicating between pilots, visual observers, operational personnel, and the actual plot ignitor. Ideally, personnel with the necessary fire qualifications and sufficient familiarity with the research and plot design are inserted into the operational command structure to facilitate the necessary coordination.

5) Conclusions

UAS can give an unprecedented perspective for data collection in active fire environments at favorable spatial and temporal scales if the software, hardware, and fire operations conflicts are resolved or minimized. Robust data collection workflows must constantly evolve while still maintaining coherent scientific rigor due to the rapid and ongoing development of UAS and sensor technology.

With the new perspective provided by UAS, we are able to image complex fire behavior that necessitated updated algorithms capable of characterizing such behavior. Fire behavior is missed between sample intervals requiring dynamic analytical unit sizes. The nature of this relationship is largely dependent on RoS and the temporal and spatial resolution of the sensor. Characterizing this relationship in a generalized fashion is likely possible with additional research and could potentially inform other pre- and post-fire sampling methods for the ultimate goal of comprehensive datasets of the fire environment.

The collection of such comprehensive datasets that characterize fuels, fire behavior, weather, and, if possible, emissions and fire effects, are essential for evaluation of models and for cross-disciplinary knowledge gains in the broader field of wildland fire science. Research projects that attempt to create such datasets are substantial endeavors. UAS could be a unifying remote sensing platform that collects such datasets in a safe and relatively inexpensive manner. Ultimately, UAS provide complementary capabilities that enhance our ability to understand how fires burn.

Literature Cited

1. Stocks, B.J.; Alexander, M.E.; Lanoville, R.A. Overview of the international crown fire modelling experiment (ICFME). *Can. J. For. Res.* **2004**, *34*, 1543–1547.

2. Clements, C.B.; Zhong, S.; Goodrick, S.; Li, J.; Potter, B.E.; Bian, X.; Heilman, W.E.; Charney, J.J.; Perna, R.; Jang, M.; et al. the dynamics of wildland grass fires: FireFlux—A field validation experiment. *Bull. Am. Meteor. Soc.* **2007**, *88*, 1369–1382.

3. Kremens, R.L.; Smith, A.M.S.; Dickinson, M.B. Fire metrology: Current and future direction in physics–based measuremens. *Fire Ecology* **2010**, *6*, 13–35.

4. Ottmar, R.D.; Hiers, J.K.; Butler, B.W.; Clements, C.B.; Dickinson, M.B.; Hudak, A.T.; O'Brien, J.J.; Potter, B.E.; Rowell, E.M.; Strand, T.M.; Zajkowski, T.J. Measurements, datasets and preliminary results from the RxCADRE project–2008, 2011 and 2012. *Int. J. Wildl. Fire* **2016**, *25*, 1–9.

5. Finney, M.A.; Cohen, J.D.; Forthofer, J.M.; McAllister, S.S.; Gollner, M.J.; Gorham, D.J.; Saito, K.; Akafuah, N.K.; Adam, B.A.; English, J.D.; Role of buoyant flame dynamics in wildfire spread. *Proc. Nat. Acad. Sci.* USA **2015**, *112*, 9833–9838.

6. Finney, M.A.; Cohen, J.D.; McAllister, S.S.; Jolly, W.M. On the need for a theory of wildland fire spread. *Int. J. Wildl. Fire* **2013**, *22*, 25–36.

7. Sauvagnargues–Lesage, S.; Dusserre, G.; Robert, F.; Dray, G.; Pearson, D.W. Experimental validation in Mediterranean shrub fuels of seven wildland fire rate of spread models. *Int. J. Wildl. Fire* **2001,** *10*, 15–22.

8. Mell, W.; Jenkins, M.A.; Gould, J.; Cheney, P. A physics–based approach to modelling grassland fires. *Int. J. Wildl. Fire* **2007,** *16*, 1–22.

9. Achtemeier, G.L. Field validation of a free–agent cellular automata model of fire spread with fire–atmosphere coupling. *Int. J. Wildl. Fire* **2012,** *22*, 148–156.

10. Alexander, M.E.; Cruz, M.G. Are the applications of wildland fire behavior models getting ahead of their evaluation again? *Environ. Model. Softw.* **2013**, *41*, 65–71.

11. Buma, B.; Wessman, C.A. Disturbance interactions can impact resilience mechanisms of forests. *Ecosphere* **2011**, *2*, 1–13.

12. Jenkins, M.J.; Runyon, J.B.; Fettig, C.J.; Page, W.G.; Bentz, B.J. Interactions among the mountain pine beetle, fires, and fuels. *For. Sci.* **2014**, *60*, 489–501.

13. Finney, M.A.; McHugh, C.W.; Grenfell, I.C. Stand– and landscape–level effects of prescribed burning on two Arizona wildfires. *Can. J. For. Res.* **2005**, *35*, 1714–1722.

14. Cochrane, M.A.; Moran, C.J.; Wimberly, M.C.; Baer, A.D.; Finney, M.A.; Beckendorf, K.L.; Eidenshink, J.; Zhu, Z. Estimation of wildfire size and risk changes due to fuels treatments. *Int. J. Wildl. Fire* **2011**, *21*, 357–367.

15. Ziegler, J.P.; Hoffman, C.; Battaglia, M.; Mell, W. Spatially explicit measurements of forest structure and fire behavior following restoration treatments in dry forests. *For. Ecol. Manag.* **2016**, *286*, 1–12.

16. Thaxton, J.M.; Platt, W.J. Small–scale fuel variation alters fire intensity and shrub abundance in a pine savanna. *Ecology* **2006**, *87*, 1331–1337.

17. Loudermilk, E.L.; O'Brien, J.J.; Mitchell, R.J.; Cropper, W.P.; Hiers, J.K.; Grunwald, S.; Grego, J.; Fernandez–Diaz, J.C. Linking complex forest fuel structure and fire behavior at fine scales. *Int. J. Wildl. Fire* **2012**, *21*, 882–893.

18. O'Brien, J.J.; Loudermilk, E.L.; Hiers, J.K.; Pokswinski, S.M.; Hornsby, B.; Hudak, A.T.; Strother, D.; Rowell, E.; Bright, B.C. Canopy–derived fuels drive patterns of in–fire energy release and understory plant mortality in a longleaf pine (*Pinus palustris*) sandhill in northwest Florida, USA. *Can. J. Remote Sens.* **2016**, *42*, 489–500.

19. Johannsson, N. *Numerical Experiments a Research Method in Fire Science*; Lund University: Lund, Sweden, 2013.

20. Lentile, L.B.; Holden, Z.A.; Smith, A.M.S.; Falkowski, M.J.; Hudak, A.T.; Morgan, P.; Lewis, S.A.; Gessler, P.E.; Benson, N.C. Remote sensing techniques to assess active fire characteristics and post–fire effects. *Int. J. Wildl. Fire* **2006**, *15*, 319–345.

21. Wooster, M.J.; Roberts, G.; Perry, G.L.W.; Kaufman, Y.J. Retrieval of biomass combustion rates and totals from fire radiative power observations: FRP derivation and calibration relationships between biomass consumption and fire radiative energy release. *J. Geophys. Res. Atmos.* **2005**, *110*, doi:10.1029/2005JD006018

22. Freeborn, P.H.; Wooster, M.J.; Hao, W.M.; Ryan, C.A.; Nordgren, B.L.; Baker, S.P.; Ichoku, C. Relationships between energy release, fuel mass loss, and trace gas and aerosol emission during laboratory biomass fires. *J. Geophys. Res. Atmos.* 2008, *113,* doi:10.1029/2007JD008679.

23. Dupuy, J.-L.; Maréchal, J.; Portier, D.; Valette, J.-C. The effects of slope and fuel bed width on laboratory fire behavior. *Int. J. Wildl. Fire* **2011**, *20*, 272–288.

24. May, N.; Ellicott, E.; Gollner, M. An examination of fuel moisture, energy release, and emissions during laboratory burning of live wildland fuels. *Int. J. Wildl. Fire* **2019**, *28*, 187–197.

25. Wooster, M.J. Small–scale experimental testing of fire radiative energy for quantifying mass combusted in natural vegetation fires. *Geophys. Res. Lett.* **2002**, *29*, 1–23.

26. Frankman, D.; Webb, B.W.; Butler, B.W.; Jimenez, D.; Forthofer, J.M.; Sopko, P.; Shannon, K.S.; Hiers, J.K.; Ottmar, R.D. Measurements of convective and radiative heating in wildland fires. *Int. J. Wildl. Fire* **2012**, *22*, 157–167.

27. Dickinson, M.B.; Hudak, A.T.; Zajkowski, T.; Loudermilk, E.L.; Schroeder, W.; Ellison, L.; Kremens, R.L.; Holley, W.; Martinez, O.; Paxton, A.; et.al. Measuring radiant emissions from entire prescribed fires with ground, airborne and satellite sensors-RxCADRE 2012. *Int. J. Wildl. Fire* **2016**, *25*, 48–61.

28. Ichoku, C.; Giglio, L.; Wooster, M.J.; Remer, L.A. Global characterization of biomass–burning patterns using satellite measurements of fire radiative energy. *Remote Sens. Environ.* **2008**, *112*, 2950–2962.

29. Roberts, G.; Wooster, M.J.; Lagoudakis, E. Annual and diurnal African biomass burning temporal dynamics. *Biogeosciences* **2009**, *6*, 849–866.

30. Vermote, E.; Ellicot, E.; Dubovik, O.; Lapyonok, T.; Chin, M.; Giglio, L.; Roberts, G.J. An approach to estimate global biomass burning emissions of organic and black carbon from MODIS fire radiative power. *J. Geophys. Res. Atmos.* **2009**, *114*, D18.

31. Sparks, A.M.; Kolden, C.A.; Smith, A.M.S.; Boschetti, L.; Johnson, D.M.; Cochrane, M.A. Fire intensity impacts on post–fire temperate coniferous forest net primary productivity. *Biogeosciences* **2018**, *15*, 1173–1183.

32. Boschetti, L.; Roy, D.P. Strategies for the fusion of satellite fire radiative power with burned area data for fire radiative energy derivation. *J. Geophys. Res. Atmos.* **2009**, *114*, doi:10.1029/2008JD011645

33. Kumar, S.S.; Roy, D.P.; Boschetti, L.; Kremens, R.L. Exploiting the power law distribution properties of satellite fire radiative power retrievals: A method to estimate fire radiative energy and biomass burned from sparse satellite observations. *J. Geophys. Res. Atmos.* **2011**, *116*, doi:10.1029/2011JD015676

34. Freeborn, P.H.; Wooster, M.J.; Roy, D.P.; Cochrane, M.A. Quantification of MODIS fire radiative power (FRP) measurement uncertainty for use in satellite–based active fire characterization and biomass burning estimation. *Geophys. Res. Lett.* **2014**, *41*, 1988–1994.

35. Hudak, A.T.; Freeborn, P.H.; Lewis, S.A.; Hood, S.M.; Smith, H.Y.; Hardy, C.C.; Kremens, R.L.; Butler, B.W.; Teske, C.; Tissell, R.G.; et.al. The Cooney Ridge Fire Experiment: An Early Operation to Relate Pre–, Active, and Post–Fire Field and Remotely Sensed Measurements. *Fire* **2018**, *1*, 10.

36. Paugam, R.; Wooster, M.J.; Roberts, G. Use of handheld thermal imager data for airborne mapping of fire radiative power and energy and flame front rate of spread. *IEEE Tran. Geosci. Remote Sens.* **2013**, *51*, 3385–3399.

37. Valero, M.M.; Rios, O.; Pastor, E.; Planas, E. Automated location of active fire perimeters in aerial infrared imaging using unsupervised edge detectors. *Int. J. Wildl. Fire* **2018**, *27*, 241–256.

38. Johnston, J.M.; Wheatley, M.J.; Wooster, M.J.; Paugam, R.; Davies, G.M.; DeBoer, K.A. Flame–front rate of spread estimates for moderate scale experimental fires are strongly influenced by measurement approach. *Fire* **2018**, *1*, 16.

39. Stow, D.; Riggan, P.; Schag, G.; Brewer, W.; Tissell, R.; Coen, J.; Storey, E. Assessing uncertainty and demonstrating potential for estimating fire rate of spread at landscape scales based on time sequential airborne thermal infrared imaging *Int. J. Remote Sens.* **2019**, *40*, 4876–4897.

40. Johnston, J.M.; Wooster, M.J.; Paugam, R.; Wang, X.; Lynham, T.J.; Johnston, L.M. Direct estimation of Byram's fire intensity from infrared remote sensing imagery. *Int. J. Wildl. Fire* **2017**, *26*, 668–684.

41. Watts, A.C.; Ambrosia, V.G.; Hinkley, E.A. Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use. *Remote Sens.* **2012**, *4*, 1671–1692.

42. Yuan, C.; Zhang, Y.; Liu, Z. A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques. *Can. J. Remote Sens.* **2015**, *45*, 783–792.

43. Pérez, Y.; Pastor, E.; Planas, E.; Plucinksi, M.; Gould, J. Computing forest fires aerial suppression effectiveness by IR monitoring. *Fire Saf. J.* **2011**, *46*, 2–8.

44. McRae, D.J.; Jin, J.–Z.;Conard, S.G.; Sukhinin, A.I.; Ivanova, G.A.; Blake, T.W. Infrared characterization of fine–scale variability in behavior of boreal forest fires. *Can. J. For. Res.* **2005**, *35*, 2194–2206.

45. Mueller, E.V.; Skowronski, N.; Clark, K.; Gallagher, M.; Kremens, R.L.; Thomas, J.C.; Houssami, M.E.; Filkov, A.; Hadden, R.M.; Mell, W.; et.al. Utilization of remote sensing techniques for the quantification of fire behavior in two pine stands. *Fire Safety, J.* **2017**, *91*, 845–854.

46. Loudermilk, E.L.; Achtemeier, G.L.; O'Brien, J.J.; Hiers, J.K.; Hornsby, B.S. High–resolution observations of combustion in heterogeneous surface fuels. *Int. J. Wildl. Fire* **2014**, *23*, 1016–1026.

47. Hardin, P.J.; Lulla, V.; Jensen, R.R.; Jensen, J.R. Small unmanned aerial systems (sUAS) for environmental remote sensing: Challenges and opportunities revisited. *GISCI Remote Sens.* **2019**, *56*, 309–322.

48. Twidwell, D.; Allen, C.R.; Detweiler, C.; Higgins, J.; Laney, C.; Elbaum, S. Smokey comes of age: Unmanned aerial systems for fire management. *Front. Ecol. Environ.* **2016**, *14*, 333–339.

49. Bright, B.C.; Loudermilk, E.L.; Pokswinksi, S.M.; Hudak, A.T.; O'Brien, J.J. Introducing close–range photogrammetry for characterizing forest understory plant diversity and surface fuel structure at fine scales. *Can. J. Remote Sens.* **2016**, *42*, 460–472.

50. Shin, P.; Sankey, T.; Moore, M.M.; Thode, A.E. Evaluating unmanned aerial vehicle images for estimating forest canopy fuels in a ponderosa pine stand. *Remote Sens.* **2018**, *10*, 1266.

51. Samiappan, S.; Hathcock, L.; Turnage, G.; McCraine, C.; Pitchford, J.; Moorhead, R. Remote sensing of wildfire using small unmanned aerial system: Post–fire mapping, vegetation recovery and damage analysis in Grand Bay, Mississippi/Alabama, USA. *Drones* **2019**, *3*, 43.

52. Loudermilk, E.L.; Hiers, J.K.; O'Brien, J.J.; Mitchell, R.J.; Singhania, A.; Fernandez, J.C.; Cropper, W.P. Jr.; Slatton, K.C. Ground–based LIDAR: A novel approach to quantify fine–scale fuelbed characteristics. *Int. J. Wildl. Fire* **2009**, *18*, 676–685.

53. Lefsky, M.A.; Cohen, W.B.; Parker, G.G.; Harding, D.J. Lidar remote sensing for ecosystem studies. *Bioscience* **2002**, *52*, 19–30.

54. Hiers, J.K.; O'Brien, J.J.; Mitchell, R.J.; Grego, J.M.; Loudermilk, E.L. The wildland fuel cell concept: An approach to characterize fine–scale variation in fuels and fire in frequently burned longleaf pine forests. *Int. J. Wildl. Fire* **2009**, *18*, 315–325.

55. Anderson, W.R.; Catchpole, E.A.; Butler, B.W. Convective heat transfer in fire spread through fine fuel beds. *Int. J. Wildl. Fire* **2010**, *19*, 284–298.

56. Bataineh, A.L.; Oswald, B.P.; Bataineh, M.; Unger, D.; Hung, I.-K.; Scognamillo, D. Spatial autocorrelation and pseudo replication in fire ecology. *Fire Ecol.* **2006**, *2*, 107–118.

57. Openshaw, S.; Taylor, P. A Million or so Correlation Coefficients: Three Experiments on the Modifiable Areal Unit Problem, In *Statistical Applications in the Spatial Sciences* Wrigley, N., Ed.; Pion: London, UK, 1979, Volume 1; pp. 127–144.

58. Gotway, C.A.; Young, L.J. Combining incompatible spatial data. *J. Am. Stat. Assoc.* **2002**, *97*, 632–648.

59. Valero, M.M.; Jimenez, D.; Butler, B.W.; Mata, C.; Rios, O.; Pastor, E.; Planas, E. On the use of compact thermal cameras for quantitative wildfire monitoring. In Proceedings of the 8th International Conference on Forest Fire Research Proceedings, Coimbra, Portugal, 9–16 November 2018.

60. Johnston, J.M.; Wooster, M.J.; Lynham, T.J. Experimental confirmation of the MWIR and LWIR grey body assumption for vegetation fire flame emissivity. *Int. J. Wildl. Fire* **2014**, *23*, 463–479.

61. Micasense Image Processing. Available online: https://github.com/micasense/imageprocessing (accessed on 1 December 2018).

62. O'Brien, J.J.; Loudermilk, E.L.; Hornsby, B.; Hudak, A.T.; Bright, B.C.; Dickinson, M.B.; Hiers, J.K.; Teske, C.; Ottmar, R.D. High–resolution infrared thermography for capturing wildland fire behavior—RxCADRE 2012. *Int. J. Wildl. Fire* **2016**, *25*, 62–75.

63. Ononye, A.; Vodacek, A.; Saber, E. Automated extraction of fire line parameters from multispectral infrared images. *Remote Sens. Environ.* **2007**, *108*, 179–188.

64. Canny, J. A Computational Approach to Edge Detection. In *Readings in Computer Vision*; Fischler, M.A.; Firschein, O., Eds.; Morgan Kaufmann: Burlington, USA, 1987; pp. 184–203.

65. Lloyd, S.P. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **1982**, *28*, 129–137.

66. Boschetti, L.; Flasse, S.P.; Brivio, P.A. Analysis of the conflict between omission and comission in low spatial resolution dichotomic thematic products: The Pareto Boundary. *Remote Sens. Environ.* **2004**, *91*, 280–292.

67. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria, 2010. Available online: http://www.R–project.org/ (accessed on 1 January 2018) .

68. Hijmans, R.J. Raster: Geographic Data Analysis and Modeling. 2018, R Package Version 2.8-4. Available online: https://cran.r–project.org/package=raster (accessed on 15 December 2018).

69. Pebesma, E.J.; Bivand, R.S. Classes and Methods for Spatial Data in R. 2005, R News 5(2). Available online: https://cran.r–project.org/doc/Rnews (accessed on 15 December 2018) .

70. Bivand, R.S.; Rundel, C. Rgeos: Interface to Geometry Engine—Open Source ('GEOS'). 2018, R Package Version 0.4-2. Available online: https://cran.r–project.org/packages=rgeos (accessed on 15 December 2018).

71. Agostinelli, C.; Lund, U. Circular: Circular Statistics. 2017, R Package Version 0.4-93. Available online: https://r–forge.r–project.org/projects/circular. (accessed on 15 December 2018)

72. Finney, M.A. FARSITE: *Fire Area Simulator–Model Development and Evaluation*. *Res. Pap.* RMRS-RP-4; Department of Agriculture, Forest Service, Rocky Mountain Research Station: Ogden, UT: US 1998, 47p.

73. Tarjan, R. Depth–first search and linear graph algorithms. *SIAM J. Computing* **1972**, *1*, 146–160.

74. Barthelme, S. Imager: Image Processing Library Based on 'CImg'. 2018, R Package Version 0.41.1. Available online: https:// cran.r–project.org/package=imager (accessed on 15 December 2018) .

75. Rothermel, R.C. *A Mathematical Model for Predicting Fire Spread in Wildland Fuels.* Res. Pap. INT–115, Department of Agriculture, Forest Service, Rocky Mountain Research Station: Ogden, UT: US 1972**;** 40p.

76. Gelfand, A.E. Misaligned spatial data: The change of support problem. *In Handbook Spatial Statistics;* 1st ed.; Gelfand, A.E., Diggle, P.J., Fuentes, M., Guttorp, P., Eds.; CRC Press: Boca Raton, FL, USA; 2010; pp. 517–539.

<u>Chapter 5.</u> Deriving 3D Fuel Variability from UAS Imagery and Characterizing Trends in Fuel Pattern-Fire Process Spatial Associations

1) Introduction

Fire and fuel science has progressed from relying on broad assumptions of homogeneity in fuel and fire metrics to explicit consideration of their variability (Hiers et al. 2009, Parsons et al. 2011, Keane et al. 2012, Loudermilk et al. 2012, Alexander and Cruz 2013, O'Brien et al. 2016a, Vakili et al. 2016, Fernandes et al. 2018). Fuel and fire have been shown to covary at fine spatial and temporal scales with significant direct and indirect relationships (Thaxton and Platt 2006, Archibald et al. 2009, Hiers et al. 2009, O'Brien et al. 2016a, O'Brien et al. 2016b). However, obtaining concurrent data on fine-scale fuel heterogeneity and fire behavior remains a research challenge (Kremens et al. 2010, O'Brien et al. 2016b, Ottmar et al. 2016a). Integrated datasets characterizing the entire fire environment are the gold standard for parameterizing and validating the new generation of fire models and for testing a multitude of management and ecological research questions (Ottmar et al. 2016a, Hudak et al. 2018) , but they are extremely difficult to obtain. Unmanned aerial systems (UAS) offer a new technology to potentially overcome many of these challenges in measure fire behavior, terrain, and fuels concurrently. In Chapter 4, fire rate of spread and direction was quantified at multiple scales from a UAS. Here, the pre-fire fuel environment will be characterized using novel, gradient-based techniques on 3D point clouds derived from structure-from-motion photogrammetry (Westoby et al. 2012). Then, a measure of bivariate spatial association, the L-index, is used to quantify trends in the influence of fuels variability on fire rate of spread (RoS) (Lee 2001).

Until now, explicit characterization of fire and fuel spatial variance has relied primarily on semivariograms relating variation at multiple spatial lags (Hiers et al. 2009, Keane et al. 2012, O'Brien et al. 2016b, Valkili et al. 2016). This has revealed the scales at which certain fuel or fire components are spatially correlated from a univariate perspective, but the underlying question that tends to be indirectly addressed is the degree at which fuel patterns of interest covary with fire behavior from a bivariate or multivariate perspective. Loudermilk et al. (2012) showed that the importance of LiDAR-derived fuel metrics for predicting fire behavior varied as a function of low, moderate, or high intensity fire. Parsons et al. (2017) showed similar results in that the effect of canopy fuel aggregation on RoS was reduced at higher wind speeds (and by extension, increased overall fire behavior). These results imply non-stationarity in the relative influence of fuel variation on fire behavior processes. The working hypothesis for this research is therefore:

Fuels variability exerts less influence on fire spread as rate of spread increases.

If true, the ramifications extend to climate change, disturbance, and vegetation interactions and feedbacks (Schoennagel et al. 2004), fuel treatment effectiveness and design (Agee and Skinner 2005), and fire behavior modeling assessment and improvement (Finney et al. 2013).

Measures of 3D fuels often require intensive field campaigns to model fuel loading, structure, and composition (Rowell et al. 2016a, Ottmar et al. 2016b, Hawley et al. 2018). I hypothesized that for objectives focused on fuel and fire variability characterization, metrics derived directly from remotely sensed data can provide the necessary information to derive fire-fuel relationship trends. Fuel heights, vertical distributions metrics, and surface area to volume ratios can be directly measured and R,G,B color information can indirectly characterize composition and moisture variation. Following the theme of developing new technology for fire science and management, UAS again provides a means for characterizing fuel complexes in 3D at reduced costs compared to traditionally-used LiDAR. Structure-from-motion photogrammetry has successfully quantified surface (Bright et al. 2016) and canopy fuels (Shin et al. 2018). UAS flown subcanopy has the potential to extend fine scale characterization of surface fuels as seen in Bright et al. (2016) but at broader extents.

Lee (2001) developed an index to characterize spatial co-patterning by combining an aspatial bivariate correlation measure, Pearson's correlation coefficient, with Moran's I, a univariate spatial association measure, termed the L-index. This can be assessed at a localized scale analogous to the suite of local indicators of spatial association (LISA) (Anselin 1995). Trends in the L-index using measures of fuel variability and fire RoS as the two variables of interest provides a statistic for testing the working hypothesis. The formulation of the L-index is simply the replacement of the cells' values in the numerator of the Pearson equation with spatial lag values. Equation 1 shows the formulation of the Pearson correlation coefficient for variables X and Y:

$$r_{X,Y=} \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \tag{1}$$

The spatial lag is defined as the weighted average of neighbors created from a spatial weights matrix (Anselin 1998):

$$\tilde{x}_i = \sum_j w_{ij} x_j \tag{2}$$

Finally, one formulation of L-index is:

$$L_{X,Y} = \frac{\sum_i (\tilde{x}_i - \bar{x})(\tilde{y}_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \tag{3}$$

A local version of L-index can be calculated as well:

$$L_i = \frac{n(\tilde{x}_i - \bar{x})(\tilde{y}_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \tag{4}$$

The study presented here can be viewed as preliminary and exploratory in nature as the technologies and methods are relatively new, and the timescales for collecting significant data on prescribed burns in the western US often cover many years. Indeed, the software, sensors, and spatial scales of measurement have all been modified from the onset of this research to the finalization of the dissertation. The work at the Sycan Marsh, OR presented in the previous chapter versus that presented here highlights a portion of the evolution in this scientific endeavor. Regardless, the ideas, methods, and working hypothesis are still relevant to the fire science community, which is rapidly adopting new technology for scientific advancement. This study's objectives are as follows:

1. Develop techniques to acquire imagery and derive 3D point clouds of surface fuels from UAS at moderate scales (100 m^2)
2. Quantify the fuel complex directly from point cloud-derived metrics and through gradient change metrics representing variability
3. Characterize the trends in fire-fuel spatial co-patterning as a function of fire spread rates

2) Methods

The general outline of the methods is as follows: characterize pre-fire fuel patterns using 3-D photogrammetry from UAS, burn plots in systematic fashion and calculate RoS, derive analytical units based on RoS patches, aggregate fuel variability and RoS to analytical units, and then estimate spatial co-patterning of fuel variability and RoS.

2.1 Site Description and Plot Layout

The 5.1 ha prescribed burn unit was located at Lubrecht Experimental Forest 20 miles NE of Missoula, MT. Surface fuels were dominated by logging slash, ponderosa pine and Douglas fir needle-litter, and relatively short grass and forbs. The prescribed burn is the same as described in Chapter 4.

Six, 10m x 10m square plots were placed within the burn unit with several conditions for placement. Plots were placed at least 20 meters from the unit boundaries to minimize interference from the burn's blacklining operations. Plots were spatially paired (three sets of two) in order to reduce the number of UAS takeoff and landing locations and the amount of equipment transfer necessary during the burn. Plots were also placed in canopy openings to reduce canopy occlusion and to reduce the complexity of UAS flight. One plot was removed as fuel conditions did not produce any appreciable fire spread.

As described in Chapter 4, plots were laid out with an outer 10 m x 10 m and a nested 3.33 m x 3.33 m square. At each corner of the squares, a length of rebar was driven into the ground leaving 20.3 cm exposed, and a circular compact disc (CD) covered in high-visibility

duct tape was attached on top. These were the ground control points (GCPs) that assisted the 3-D point generation process. The geometry of the plot as defined and measured from the CD locations was also used to create eighteen scale constraints for calibration and estimating horizontal and vertical error in the point cloud. Outer and inner square edges (n=8/plot) and outer square diagonals (n=2/plot) were used for horizontal calibration and error estimation, and the height of the rebar above ground was used for the vertical calibration and error estimation (n=8/plot). When the plots were burned, the CDs were replaced with CDs covered in aluminum foil (low emissivity) which gave them visibility in thermal imagery (cold targets) and allowed the pre-fire fuel data and thermal data to be spatially rectified together.

2.2 Fuel Patterns

Fuels were characterized using a DJI (Shenzhen, GD, China) Phantom P4 Pro UAS and processed in Pix4D (v. 3.2, Prilly, Switzerland) photogrammetry software using structure-from-motion algorithms (Westoby et al. 2012). Photos were taken near mid-day in cloudy, overcast conditions to minimize shadowing and ensure consistent illumination. For each plot, the UAS pilot manually flew a flight pattern with at least four gridlines in two perpendicular directions (cross-grid). Two cross-grids were flown at approximately 4 and 9 meters above ground. Oblique and nadir imagery was taken at each image collection point. Flights could not be automated due to the nearby trees and small branches that may not be sensed by the UAS avoidance system causing a crash and the multiple angles of image exposure. Thus, acquisition parameters related to altitude and image overlap were not consistent in part from the manual UAS control, but also the trees and branches that had to be avoided. A general rule was to attempt to maintain 75% overlap between successive images. Figure 1 shows the general flight pattern taken for each plot. Additional cross-grids with other patterns and flight altitudes were experimented with but tended to cause confusion when the algorithm grouped tiepoints among sets of images that did not tie into the larger set of tiepoints. This had the effect of creating false surfaces of tiepoints at irregular angles or distances relative to the broader set ultimately reducing the quality of the point cloud. Images from the lower altitude cross-grid tended to not have enough common tiepoints to the higher cross-grid. Many of those images were not used as a result. The lower cross-grid was likely not necessary for fuel metrics ultimately derived at a 10 cm resolution.
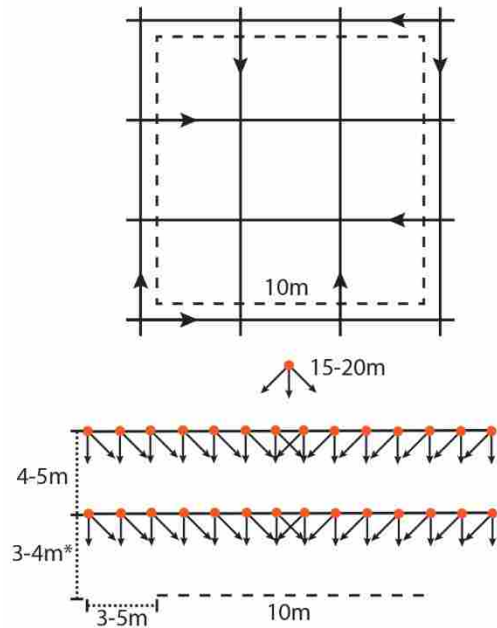
Figure 1. Flight patterns for image collection from UAS. Top pattern is from nadir perspective and bottom is a horizontal view. Orange dots are the exposure stations and arrows depict the camera angle. *Optional additional flight altitude that did not add significant value for deriving 10 cm resolution fuel metrics.

Pix4D software creates a point cloud with each point having x,y,z location and R,G,B color data derived from the set of images. Settings were set to maximum resolution (original image resolution) and ultra-high density of points. Points had to be identified in at least three images to be created. In areas of sparse automatic tiepoints, additional manual tiepoints consisting of vegetation features recognizable from multiple perspectives were identified to boost the number and quality of rectified images. Thirty-three to sixty-three total (mean 46.4) manual tiepoints (including GCPs and scale constraints) were added for each plot. Each tiepoint was manually identified in 4 to 144 images adding additional images until the software confidently placed each tiepoint in the correct location.

The aforementioned issue of erroneous sets of tiepoints not matched to the larger set was a significant obstacle to overcome. To correct these errors, the process became one of trial-and-error by adding additional manual tiepoints, changing parameters, removing images with error prone or low numbers of automatic tiepoints, and then restarting the tiepoint matching process and checking results. One nadir image at an altitude showing the entire plot was useful for identifying all GCPs, additional tiepoints, and scale constraints within one image.

Point cloud error estimates were calculated using these GCPs and scale constraints as the source of truth. However, Pix4D also uses these data as a source of calibration and thus produces a biased, though still useful, measure of error that likely underestimates true error in the point cloud. Figure 2 shows an oblique perspective of plot 5. The set of GCPs, manual tiepoints, and scaling features are visible. Image acquisition took approximately two hours per plot, GCPs one to two hours each (~40-60 hours/plot), eight to twenty-four hours per plot for image matching, and finally point cloud creation required about 72 hours of processing time per plot.
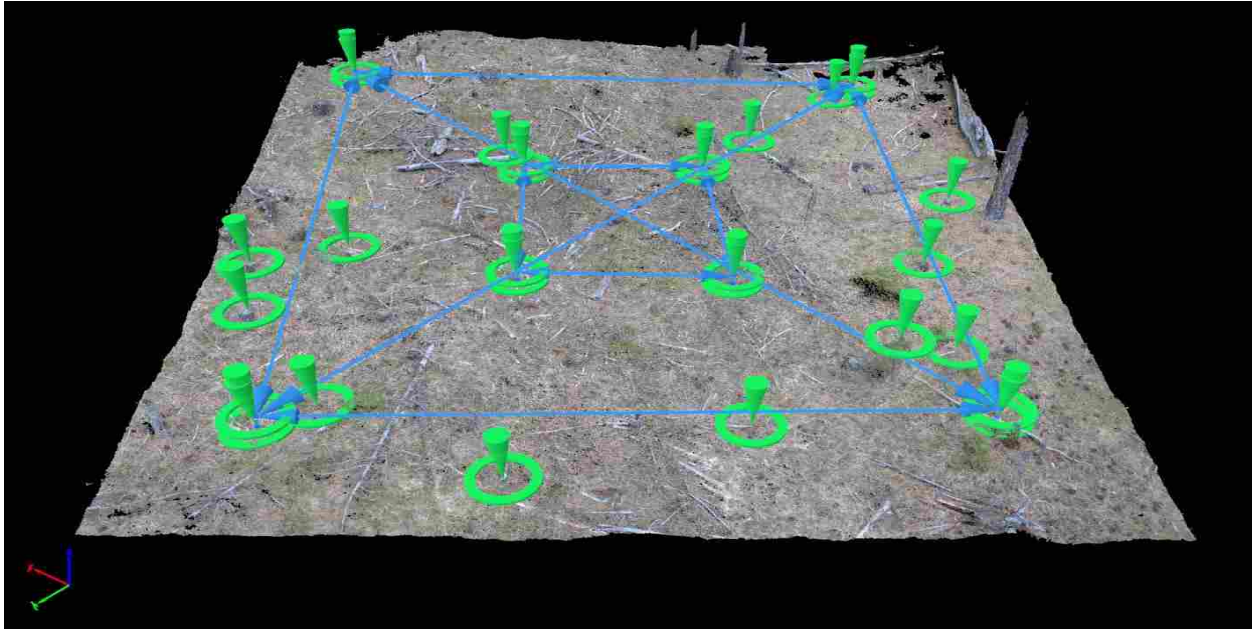
Figure 2. Example of plot layout and underlying point cloud (plot 5 shown). Blue lines represent scale constraints added for calibration and error estimation. Green pointers show manually identified GCPs and tiepoints used for image matching. Outer square dimensions are 10 m x 10 m and the inner square dimensions are 3.33 m x 3.33 m. Each corner of the two squares have two manual tiepoints: one at the ground surface and one at 20.3 cm above ground at the CD location to create vertical scale constraints.

After point cloud generation, data was rasterized to a 10 cm grid matching the thermal image-derived RoS data, and a parsimonious set of metrics was chosen to characterize the surface fuel mosaic. All points within each grid were exported to individual files using LASTools (Isenburg 2018). The rest of the analysis was completed within R software (version 3.5). Each point cloud file was uploaded and metrics calculated using the rLiDAR package (Silva et al. 2017). The package calculates the majority of the metrics calculated by the commonly used FUSION software (McGaughey 2015). Many of these metrics are conditional on first estimating a ground surface and then attributing a height above ground to each point. However, ground classification algorithms were designed primarily for aerial lidar scanning (ALS) data, which do not have the density of points created here (mean 235,516,443 points per plot) and also have the penetrating ability of an active laser sensor. A conventional TIN-based approach implemented in LasTools resulted in the majority of points classified as ground, thus removing surface variability

from resulting fuel height metrics. Therefore, ground was defined as the minimum height for each 10 cm cell, and all height metrics were defined from that point. This produced a subtle difference from the TIN approach as a majority of points were still at or near the surface but more variability was ultimately retained in resulting metrics. Eight metrics were chosen as the initial set of fuel variables with Pearson correlation used as the feature selection metric to reduce data redundancy and analytical complexity. Table 1 shows the resulting four variables chosen to characterize the fuel complex. Figure 3 shows the Pearson correlation comparison for the eight evaluated metrics.

Table 1. Chosen fuel variables to characterize the fuel complex

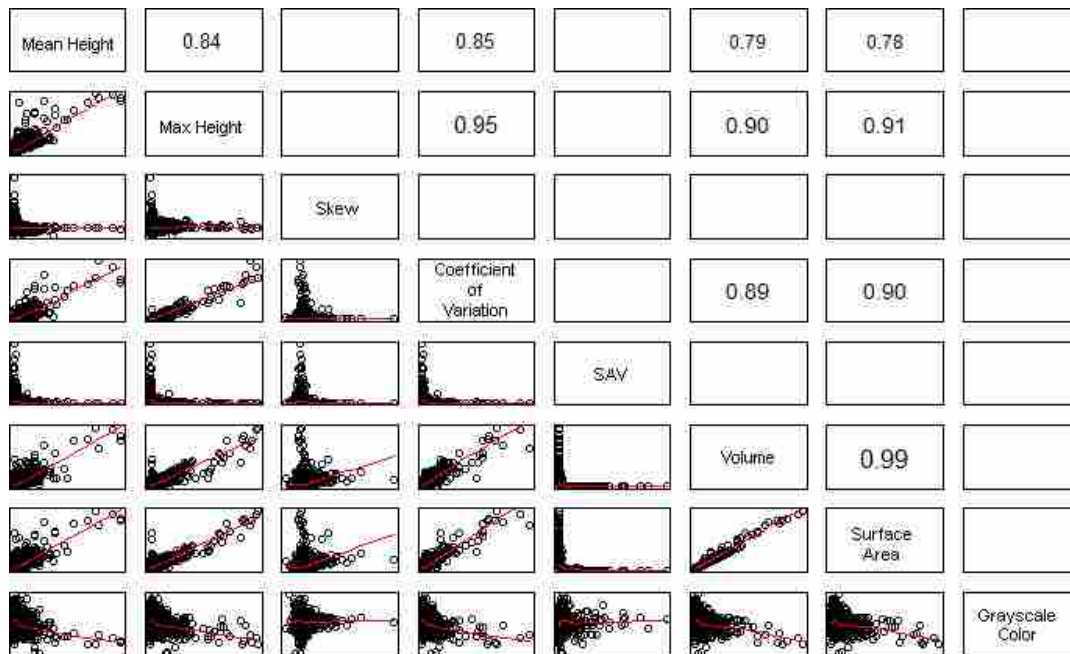| Metric	Definition
Mean Height	Mean height of points above 1st percentile height
SAV	Surface area to volume ratio
Skew	Skewness of the vertical distribution above the 1st percentile height
Grayscale Color	(0.299 * Red) + (0.587 * Green) + (0.114 * Blue)

Figure 3. Fuel variable pairwise comparison. Significant Pearson correlation coefficients displayed in upper right boxes with corresponding scatterplots in the bottom left. Mean height, skew, surface area to volume (SAV) and grayscale color were selected as fuel variables for the analysis.

The influence of fuel variability on fire rate of spread was the objective for this study. Thus, the fuel variables were converted to gradients characterizing their change rather than absolute values. To quantify this fuels variability, the norm of the gradient for each of the fuel metrics was calculated using the following equation:

$$\sqrt{\left(\frac{\partial}{\partial_y}I\right)^2 + \left(\frac{\partial}{\partial_x}I\right)^2} \qquad (5)$$

Where $I(x,y)$ is the fuel value at cell x,y. $\frac{\partial}{\partial_y}I$ and $\frac{\partial}{\partial_x}I$ partial derivatives were estimated using the imager package in R (Barthelme 2018). The gradient fuel variables were then aggregated to the analytical units derived in Chapter 4 in order to coincide with the fire behavior data. Each variable was scaled from 0 to 1 to facilitate comparisons. The scaled variables were also averaged together to create one combined variable assessing fuel variability.

2.3 Fire Rate of Spread

See Chapter 4 for in-depth description of methods for the fire RoS and direction data layers (Moran et al. 2019). These layers were spatially rectified to the fuel layers using the eight GCPs identifiable in both sets of data.

2.4 Statistical Analysis

Analytical units were derived from the RoS data following methods described in Chapter 4 (Moran et al. 2019). Fuels data were aggregated from 10 cm to the variable-sized analytical units using summary statistics. The L-index, the measure to assess the bivariate spatial co-patterning (Lee 2001), was then calculated using the sp package in R (Pebesma et al. 2018). L-indices were calculated for fuel variability layers and also individually for each of the fuel variables listed in Table 1 (after conversion to gradients). Random permutation tests (n=10,000) were conducted to assess the significance of spatial co-patterns.

3) Results

The point clouds produced representations of the fuel complex in exquisite detail (Fig. 4). Table 2 shows the calibrated images, total images taken, number of 3D points, mean ground sampling distance, and root mean square error (RMSE) for the vertical and horizontal planes. The mean RMSE for all plots was 6.97 cm in the horizontal and 1.53 cm in the vertical plane. Ground sampling distances ranged from 0.11 to 0.15 cm, and the number of 3D points created ranged from 206 million to 352 million. The number of calibrated images correlated to the number of 3D points generated. Plot 2 had the lowest number of calibrated images (182) and the

lowest number of 3D points (205,898,726). Plot 1 had the highest number of calibrated images (340) and highest number of 3D points (351,691,433). Table 3 describes the fuel metrics results for each plot. Figure 5 provides an example of the calculated fuel metrics and their fuel variability counterparts. Figure 6 shows the fuel variables in their final form after being aggregated to the analytical units.

Figure 4. Left: 3-D point cloud of plot burned in May, 2017, created from structure-from-motion photogrammetry containing x, y, z spatial location and R, G, B color intensity for each point. Right: Image of same location. Spatial differences are largely due to geometric distortion, which is rectified in the point cloud.

Table 2. Point Cloud Generation Results

Plot	Calibrated Images	Total Images	3D Points	Mean Ground Sampling Distance (cm)	RMSE Horizontal Error (cm)	RMSE Vertical Error (cm)
1	340	395	351,691,433	0.11	8.05	2.14
2	182	344	205,898,726	0.14	9.13	1.18
3	196	338	223,094,765	0.15	4.51	1.44
4	286	474	239,915,318	0.14	7.45	0.85
5	270	457	243,369,340	0.11	5.72	2.05

The statistical analysis showed statistically significant areas of both negative and positive local L-index spatial co-patterning (Fig. 7). Plots 5 and 4 did not show significance in the global L-index ($p > 0.05$), plot 1 showed marginal significance ($p = 0.03$), and plots 2 and 3 showed high significance ($p < 0.01$). Two major hotspots of spatial co-patterning were identified where moderately high rates of spread aligned with high and low fuel variability in plots 2 and 3, respectively. Other plots showed local regions with the same relationships but at reduced L-index

values. In plot 5, the lower-left third showed sustained negative L-index following head fire movement (upper left to lower right general spread direction, black arrows in Fig. 6).

Figure 8 shows the local L-index values compared to their RoS. Similar patterns are seen for each fuel variables' L-index and the combined variable. The vast majority of samples (n=14,445) are at lower rates of spread (<0.025 m/s) and low levels of spatial co-patterning (<0.5) owing to the highly skewed distribution of rates of spread observed (Moran et al. 2019, Chapter 4). The highest levels of spatial co-patterning tend to be at the lower rates of spread but the majority of sample points are at lower rates as well.

Table 3. Fuel metrics summary by plot.

Plot	Mean Height (m)			Grayscale Color			Surface Area to Volume			Skewness		
	Range	Median	SD	Range	Median	SD	Range	Median	SD	Range	Median	SD
1	0-2.71	0.04	0.10	14903-57566	31239	3501	40.6-803.9	71.2	25.9	-16.3-104.8	0.27	2.5
2	0-2.69	0.02	0.17	17428-52972	36158	3424	40.9-327.9	100.8	29.4	-7.9-78.2	0.21	1.8
3	0-2.4	0.02	0.07	13318-52369	33859	3682	41.2-220.7	96.4	26.0	-19.7-64.2	0.20	1.9
4	0-2.36	0.02	0.07	17405-51501	34591	3576	41.4-1028.4	79.7	22.6	-33.8-67.3	0.24	1.8
5	0-1.15	0.02	0.03	13332-56263	36827	3533	42.6-1185.5	84.2	30.2	-8.2-52.1	0.21	1.3
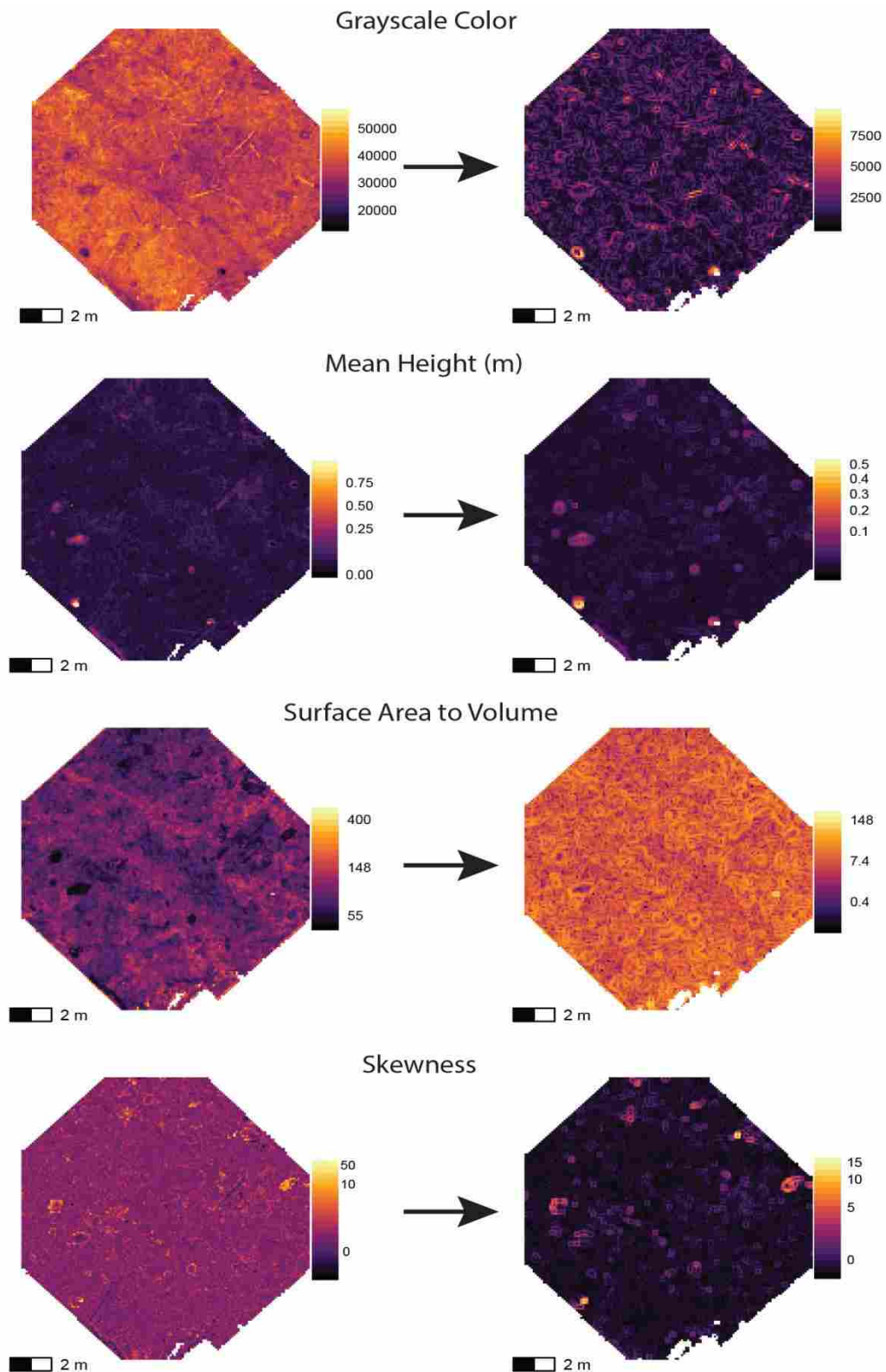
Figure 5. Derived fuel metrics for plot 5 (left column). Fuel layers converted to gradient layers shown on right. Color ramps are transformed (except grayscale) to highlight observed variability.
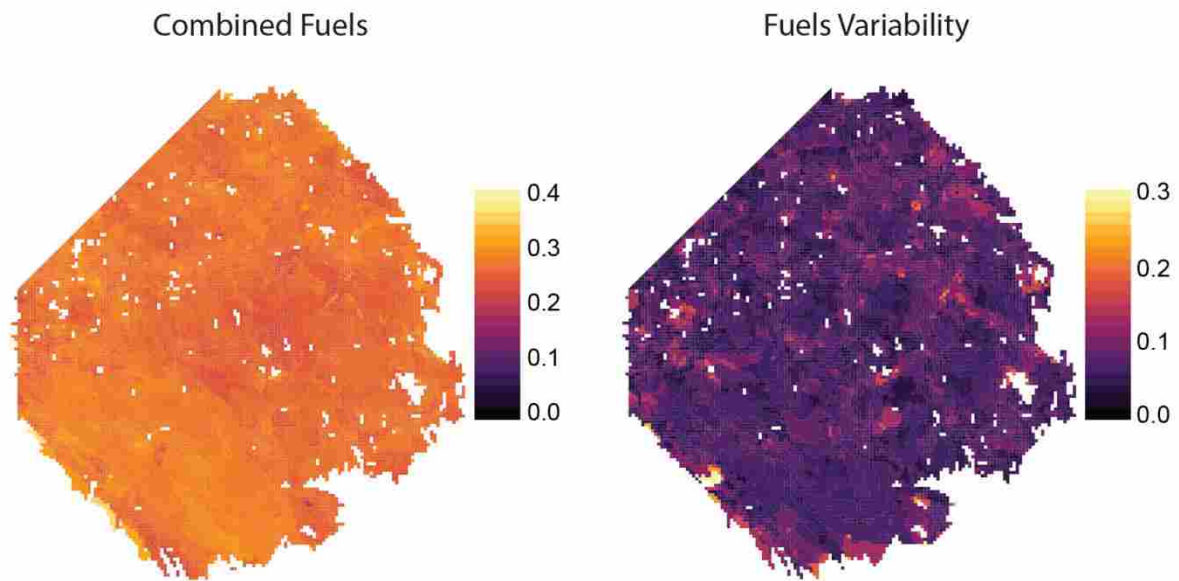
Figure 6. Derived fuels (left) and gradient (right) metrics. Each fuel layers is aggregated to analytical units, scaled from zero to one, and then averaged to create the combined variable.
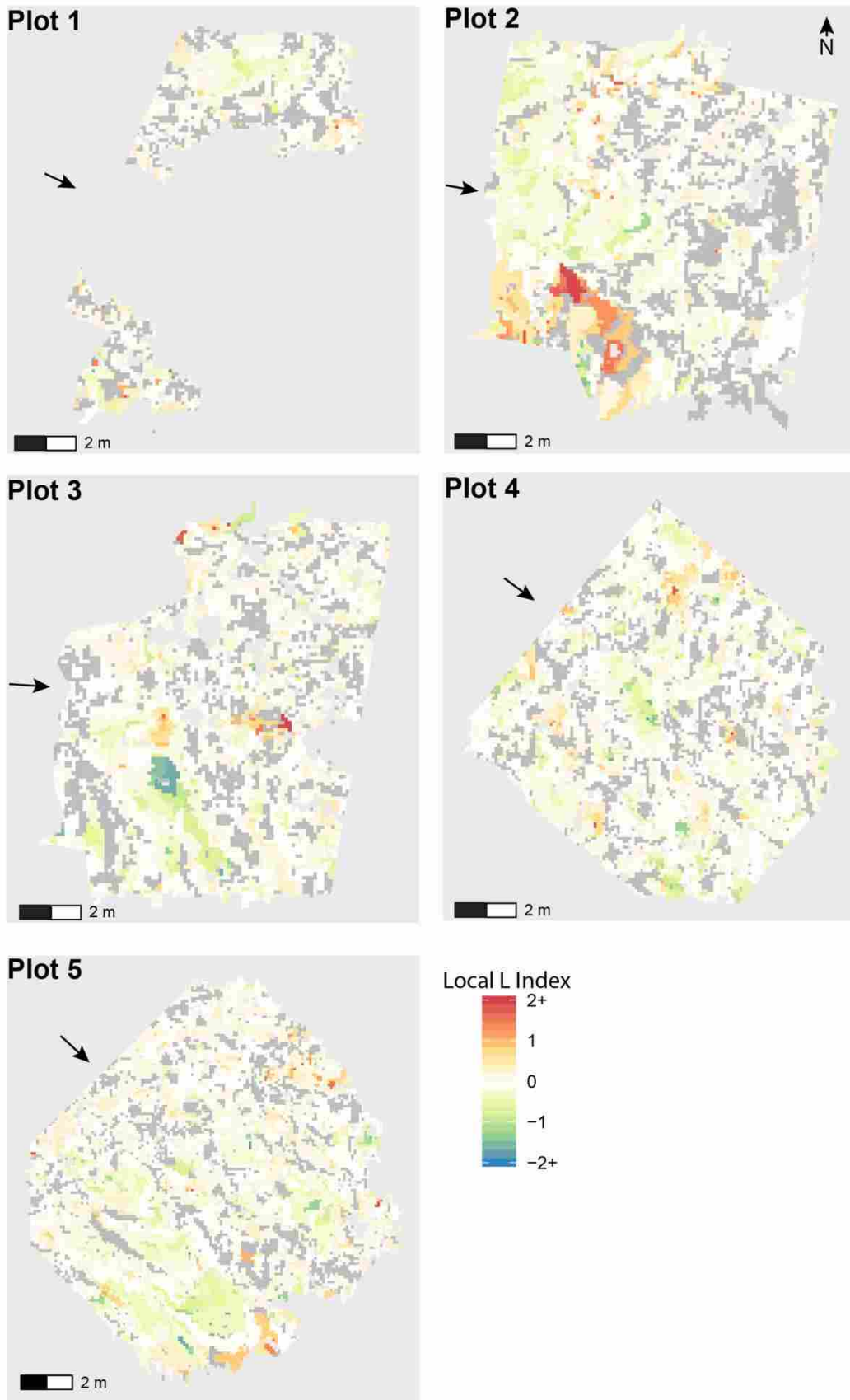
Figure 7. Local L-index values correlating fuel variability and fire rate of spread (RoS). Dark grey indicates areas of no statistical significance while colored areas are significant at p < 0.05. Black arrows show predominant fire spread direction.
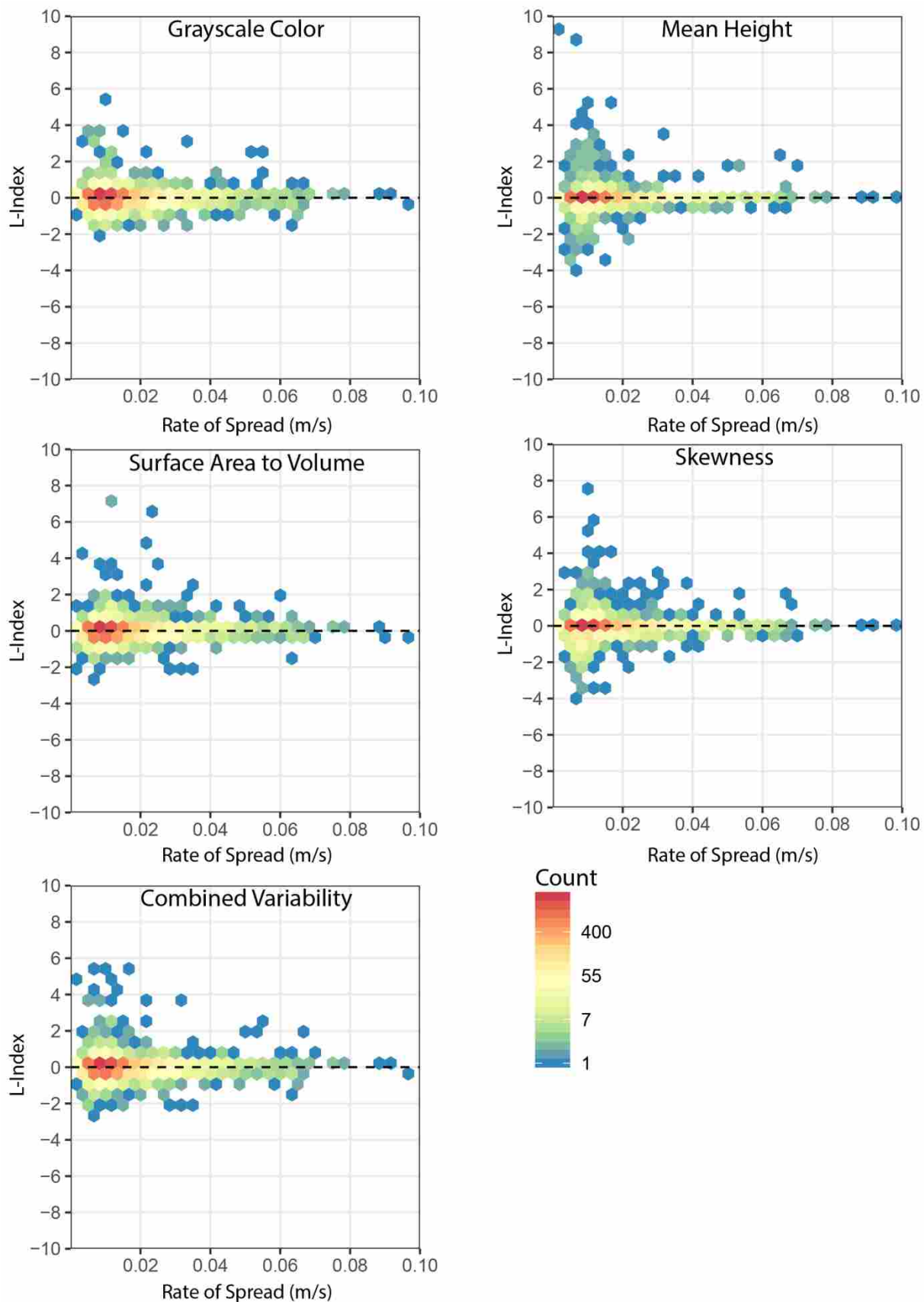
127

Figure 8. Local L-index values compared to rate of spread for all plots. The top four are formulated using each particular metric's variability layer while the combined variability is a function of all four metrics (see methods). Note that the color ramp is on a logarithmic scale.

4) Discussion

The results suggest a trend in the decoupling of fuels and fire correlation as RoS increases. The high L-index values tend to be at low RoS, and high RoS tends to have low L-index. However, the large majority of samples at low RoS (<0.02 m/s) ultimately do not provide the necessary data distribution for definitive conclusions. The methodology can characterize fuel and fire dynamics though. With a better distribution of RoS values and a diversity of fuel types beyond those measured in this study, consistent and replicated trends may appear. Better characterization of other variables influencing fire spread (i.e. weather and topography) could also help isolate the influence of fuel variability on fire spread. While differences in topography were minimal between plots, weather varied and was not explicitly considered in the analysis. The spatial correlations can reduce the effect of weather on analyses, but this 'lurking' variable is ever-present in field studies attempting to characterize fuel-fire behavior relationships (Wimberly et al. 2009).

The L-index is an integrated measure of the aspatial, Pearson's bivariate correlation and Moran's I univariate spatial association (Lee et al. 2001). While it allows assessment of the spatial co-patterning of two variables, its integrated nature can reduce interpretability of L-index values. For example, splitting the Pearson's coefficient and spatial association is necessary to determine if point-to-point or spatial lag is driving L-index variation. However, the strong spatial association seen in fire behavior necessitates this type of approach to understand the true influence of the underlying variability in fuels, weather, and topography (Bataineh et al. 2006).

Several potential improvements to the analysis were identified but were beyond the scope of the dissertation. The following will be evaluated after additional fuel and fire data are acquired, such as those from the ongoing Sycan Marsh prescribed burns described in Chapter 4. First, the neighborhood of analytical units used in the formulation of the spatial weights matrix was defined as the eight nearest-neighbors. Fire spread has inherent directionality and the spread direction estimates could inform the neighborhood definition. Two separate associations could be derived, neighborhoods defined by previous timesteps and neighborhoods defined by subsequent timesteps. Second, RoS processes change based on heading, flanking, and backing fire spread types. Fire direction could again classify these fire types, and L-index values stratified by fire type classification. Third, the analytical units vary in size and while designed to create relatively independent samples, each unit has a variable effect on the landscape. Faster RoS also tend to have larger analytical units and thus are inherently under-represented in the sample count. A weighting scheme using the area of the analytical units would likely correct for this sample bias.

The most comparable results assessing fuel heterogeneity and fire behavior are from physics-based fire models with either simulated or measured fuel complexes (e.g. Lynn et al. 2002, Hoffman et al. 2012, Parsons et al. 2017), though discrete spatial analyses have not been conducted and fuel patterns restricted to broad classifications such as spatially aggregated or homogenous classes. Parsons et al. (2017) provides explicit characterizations of fuel variability

effects on fire behavior at varying levels of fire behavior (modulated by varying wind speeds), and their results corroborate the trends seen here. In the low wind speed scenario, the mean forward RoS decreased by 22% due to high versus low levels of spatial aggregation in fuels. In the high wind speed scenario, the low to high spatial aggregation effect was only a 3.7% mean decrease in RoS implying less influence of fuel variability on fire RoS. Loudermilk et al. (2012) field-based analysis shows variable importance metrics as a function of low, moderate, and high fire intensity classes implying non-stationarity in the fuel-fire relationships. The ability of models to control confounding variables suggests the trends hypothesized here are likely to be better quantified in model space, and the L-index and associated methods presented here can be applied to modeled, laboratory, or field derived data.

The 3D point cloud and fuel variability layer produced a detailed and precise representation of the pre-fire fuel environment. This represents the first attempt at fine-scale surface fuel characterization from a UAV flown below the canopy at moderate (100 m^2) spatial extents. Small-scale photogrammetry (3 m^2 plots) using ground-based cameras has been attempted (Bright et al. 2016) but is not feasible at broader scales. Given the mean analytical unit size of 332 cm^2 derived from the rate of spread data described in Moran et al. (2019), producing plots with a mean of 235,816,443 points is not worth the time and computational resources if fire processes and fuel pattern relationships are the objective. Indeed, experimental trials show the image matching and tiepoint process is easier if attempted with images of reduced resolutions, which would then require less manual tiepoint identification and use a more automated processing workflow in general at the cost of losing some fine-scale detail.

While the horizontal and vertical error assessment and comparison to images show an accurate representation of the surface fuels, one major issue was identified. Tall, vertical objects, in this case tree boles, can be misrepresented leading to errors in the subsequent fuel variability metrics. The area of high variability and high RoS spatial co-patterning in plot 2 (Fig. 6) was caused by the tree bole leaning off the vertical axis. This caused the 3D points representing the tree bole to appear in the wrong grid cells producing high measures of variability. In reality, the affected area was relatively flat with high abundance of needle litter, which should have been characterized as a low variability area. This would have generated a negative L-index similar to the patch in plot 3 (Fig. 6). The error was likely caused by several factors including high levels of image distortion and parallax given the close distance of the UAS imagery in combination with the change in image overlap and perspective due to the UAS pilot avoiding the trees' branches. Even without the error, resolving 3D to 2D metrics (i.e. area-based approaches) will likely cause misrepresentations when overhead branches, shrubs, tree boles, etc. are within the same x,y plane but do not materially affect the fire behavior. Identifying which components of 3D point clouds are most important for characterizing fuel-fire behavior relationships is an ongoing research theme (Rowell et al. 2016b).

The methods presented here could be used to quantify which of the remotely-sensed 3D components co-pattern with observed fire behavior. The L-index was derived for all four of the

variables individually and compared (Figure 8). The general patterns are largely the same for each individual fuel variable in this case. A larger suite could be compared in detail to evaluate of which structural components of the fuel complex have the largest influence on RoS. Development of the ideas on fuel characterization and the variability defined as first derivative gradients presented here could eventually obviate the time-consuming process of field-based 3D fuel measurement (e.g. Hawley et al. 2018) and provide a direct link from remotely-sensed data to fuels and fire behavior characterization.

The fuel metrics mean height, skew, SAV, and grayscale color, can be linked intuitively to physical components of the fuel complex, specifically structure and composition. Both Loudermilk et al. (2012) and Hiers et al. (2009) found structure derived from 3D data to be significant variables in explaining fine-scale fire behavior. The relative influence of each fuel metrics is not necessarily known and an average of all four represents an adequate first attempt at a combined fuel characterization but can be further refined. A sensitivity analysis of existing operational or more complex physics-based fire models could provide support for variable weighting. However, Liu et al. (2015) found nonstationarity in the sensitivities across different fuel types, which complicates the issue. For example, fuel depth (correlative to mean height used here) explained 40% of the Rothermel fire model output variance in light logging slash but only 12% in short grass fuel models. Given that each fuel metric showed similar trends in L-index and RoS relationships, an equal weighting remains reasonable until better information can inform a different scheme, which would likely have to change in different fuel types. The point clouds also do not account for subsurface fuel variability. Litter depth is an essential component of calculating surface fuel load, a standard input to fire behavior models, although the actual effect of litter depth on rates of spread may be minimal in certain situations (e.g. Kreye et al. 2012).

Scale is another important factor to consider when interpreting the results of this work. I expect the strength of spatial co-patterning-and perhaps the general nature-to change at difference scales of fuel variability characterization and at different levels of aggregation of the RoS analytical units. Further research into this subject is planned and measurements at coarser grains are ongoing (e.g. Sycan plots described in Chapter 4). A motivation for this continuing research is the potential for trends in spatial co-patterning to inform the scales of field sampling and fire modeling. Fuel variability at fine scales can show the effect of individual objects but at management-relevant scales, characterizing the effect of fuel type changes or clump-gap canopy arrangements (Larson and Churchill 2012) requires broader scales of variability assessment. Variable-sized, moving-window averages of variability and/or larger focal areas for gradient calculation could delineate stark, courser changes in the fuel complex.

Hiers et al. (2009) also found significant spatial correlation in both fuel and fire behavior with semivariograms range values at approximately 0.5 m^2 (but analyzed separately). The L-index approach could be used to produce a corroborative analysis to better elucidate the variability in the fuel-fire relationships due to scale, although the effect of analytical units and their size variation would need to be considered in tandem. Complicating the issue is the

apparent heterogeneity in the spatial scaling of at least the fuel component. Keane et al. (2012) and Vakili et al. (2016) found significant differences in spatial scaling of fuels among field sites and a small number of fuel types have been assessed in general. Clearly, more research is necessary into the spatial relationships of fuel and fire behavior. This research provides a method for producing an integrated characterization that considers the spatial components of fire and fuel variability together.

5) Conclusions

Detailed, 3D point clouds characterizing fuels variability can be derived from UAS imagery. The close-range imagery can create errors in fuel locations especially for tall objects and when deviations in flight patterns are necessary to avoid obstacles. A mean of 235,516,443 3D points generated per 100 m^2 plot with a mean ground sampling distance of 0.13 cm had unnecessary detail considering the fire spread rate analytical unit size averaged 332 cm^2 and processing at coarser grains would decrease the necessary number of manual tiepoints and time for processing. Bivariate measures of spatial association (L-index) suggest a decoupling of fuel variability and RoS co-patterning as RoS increases, but further research is necessary in a diversity of fuel types, levels of fuel variability, scales of observations, and ranges of RoS before robust conclusions on the trends in fuel and fire behavior co-patterning can be drawn. The analytical methods developed here could be applied to either model- or laboratory-based experiments, which could better control confounding variables such as weather variation.

Literature Cited

Agee, J.K. and C.N. Skinner. 2005. Basic principles of forest fuel reduction treatments. Forest Ecology and Management 211: 83-96.

Alexander, M.E. and M.G. Cruz. 2013. Limitations on the accuracy of model predictions of wildland fire behavior: a state-of-the-knowledge overview. The Forestry Chronicle 89: 370-381.

Anselin, L. 1995. Local indicators of spatial association: LISA. Geographical Analysis 27: 93-115.

Anselin, L. 1988. Spatial econometrics: methods and models. Kluwer Academic Publishers, Boston, MA, USA.

Archibald, S., D.P. Roy, B.W. Van Wilgen, and R.J. Scholes. 2009. What limits fire? An examination of drivers of burnt area in Southern Africa. Global Change Biology 15: 613-630.

Barthelme, S. Imager: image processing library based on 'CImg'. 2018, R Package, Version 0.41.1. Available online: https://cran.r-project.org/packager=imager (accessed on 15 December 2018).

Bright, B.C., E.L. Loudermilk, S.M. Pokswinksi, A.T. Hudak, and J.J. O'Brien. 2016. Introducing close-range photogrammetry for characterizing forest understory plant diversity and surface fuel structure at fine scales. Canadian Journal of Remote Sensing 42: 460-472.

Fernandes, P.M., A. Sil, D. Ascoli, M.G. Cruz, M.E. Alexander, C.G. Rossa, J. Baeza, N. Burrows, G.M. Davies, A. Fiedelis, J.S. Gould, N. Govender, M. Kilinc, and L. McCaw. 2018. Drivers of wildland fire behavior variation across the Earth. VIII International Conference on Forest Fire Research, Columbra, Portugal.

Finney, M.A., J.D. Cohen, S.S. McAllister, and W.M. Jolly. 2013. On the need for a theory of wildland fire spread. International Journal of Wildland Fire 22: 25-36.

Hawley, C.M., L. Loudermilk, and E. Rowell. 2018. A novel approach to fuel biomass sampling for 3D fuel characterization. MethodsX 5: 1597-1604.

Hiers, J.K., J.J. O'Brien, R.J. Mitchell, J.M. Grego, and E.L. Loudermilk. The wildland fuel cell concept: an approach to characterize fine-scale variation in fuels and fire in frequently burned longleaf pine forests. International Journal of Wildland Fire 18: 315-315.

Hijmans, R.J. Raster: Geographic Data Analysis and Modeling. 2018, R Package Version 2.8-4. Available
online: https://cran.r-project.org/package=raster (accessed on 15 December 2018).

Hudak, A.T., P.H. Freeborn, S.A. Lewis, S.M. Hood, H.Y. Smith, C.C. Hardy, R.L. Kremens, B.W. Butler, C. Teske, R.G. Tissel, et al. 2018. The Cooney Ridge Fire Experiment: An early operation to relate pre-, active, and post-fire fiel and remotely sensed measurements. Fire 1: 10.

Isenburg, Martin. 2018. LAStools - efficient tools for LiDAR processing.
version 111216, http://lastools.org.

Keane, R.E., K. Gray, V. Bacciu, and S. Lierfallom. 2012. Spatial scaling of wildland fuels for six forest and rangeland ecosystems of the northern Rocky Mountains, USA. Landscape Ecology 27: 1213-1234.

Kremens, R.L., A.M.S. Smith, and M.B. Dickinson. 2010. Fire metrology: current and future directions in physics-based measurements. Fire Ecology 6: 13-35.

Kreye, J.K., L.N. Kobziar, W.C. Zipperer. 2012. Effects of fuel load and moisture content on fire behavior and heating in masticated litter-dominated fuels. International Journal of Wildland Fire 22: 440-445.

Lee, S-II. 2001. Developing a bivariate spatial association measure: an integration of Pearson's r and Moran's I. Journal of Geographical Systems 3: 369-385.

Larson, A.J. and D. Churchhill. 2012. Tree spatial patterns in fire-frequent forests of western North America, including mechanisms of pattern formation and implications for designing fuel reduction and restoration treatments. Forest Ecology and Management 267: 74-92.

Loudermilk, E.L., J.K. Hiers, J.J. O'Brien, R.J. Mitchell, A. Singhania, J.C. Fernandez, W.P. Cropper, and K.C. Slatton. 2009. Ground-based LIDAR: a novel approach to quantify fine-scale fuelbed characteristics. International Journal of Wildland Fire 18: 676-685.

Liu, Y., Y. Hussaini, and G. Okten. 2015. Global sensitivity analysis for the Rothermel model based on high-dimensional model representation. Canadian Journal of Forest Research 45:1474-1479.

McGaughey, R.J. 2015. FUSION v3.5. USDA Forest Service, Pacific Northwest Research Station, Olympia, WA. http://forsys.cfr.washington.edu/

Moran, C.J., E. Rowell, and C.A. Seielstad. 2018. A data-driven framework to identify and compare forest structure classes using LiDAR. *Remote Sensing of Environment* 211: 154-166.

Moran, C.J., C.A. Seielstad, M.R. Cunningham, V. Hoff, R.A. Parsons, L. Queen, K. Sauerbrey, and T. Wallace. 2019. Deriving fire behavior metrics from UAS imagery. *Fire* 2: 36.

O'Brien, J.J. E.L. Loudermilk, J.K. Hiers, S. Pokswinski, B. Hornsby, A. Hudak, D. Strother, E. Rowell, B.C. Bright. 2016a. Canopy-derived fuels drives patterns of in-fire energy release and understory plant mortality in a longleaf pine (Pinus palustris) sandhill in northwest Florida, USA. Canadian Journal of Remote Sensing 42: 489-500.

O'Brien, J.J., E.L. Loudermilk, B. Hornsby, A.T. Hudak, B.C. Bright, M.B. Dickinson, J.K. Hiers, C. Teske, and R.D. Ottmar. 2016b. High-resolution infrared thermography for capturing wildland fire behavior – RxCADRE 2012. International Journal of Wildland Fire 25: 62-75.

Ottmar, R.D., J.K. Hiers, B.W. Butler, C.B. Clements, M.B. Dickinson, A.T. Hudak, J.J. O'Brien, B.E. Potter, E.M. Rowell, T.M. Strand, et al. 2016a. Measurements, datasets and preliminary results from the RxCADRE project-2008, 2011 and 2012. International Journal of Wildland Fire 25: 1-9.

Ottmar, R.D., A.T. Hudak, S. Prichard, J. Wright, C.S. Restaino, J.C. Kennedy, and R.E. Vihnanek. 2016b. Pre-fire and post-fire surface fuel and cover measurements collected in the south-eastern United States for model evaluation and development-RxCadre 2008, 2011 and 2012. International Journal of Wildland Fire 25: 10-24.

Parsons, R.A., W.E. Mell, and P. McCauley. 2011. Linking 3D spatial models of fuels and fire: Effects of spatial heterogeneity on fire behavior. Ecological Modeling 222: 679-691.

Parsons, R.A., R.R. Linn, F. Pimont, C. Hoffman, J. Sauer, J. Winterkamp, C.H. Sieg, and W.M. Jolly. 2017. Numerical investigation of aggregated fuel spatial pattern impacts on fire behavior. Land 6: 43.

Pebesma, E. et al. Sp: classes and methods for spatial data. 2018. R package version 1.3-1. Available online: https://cran.r-project.org/web/packages/sp/index.html (accessed on 8 Aug 2019).

Rowell, E.M., C.A. Seielstad, and R.D. Ottmar. 2016a. Development and validation of fuel height models for terrestrial lidar—RxCadre 2012. International Journal of Wildland Fire 25: 38-47.

Rowell, E., E.L. Loudermilk, C. Seielstad, and J.J. O'Brien. 2016b. Using simulated 3D surface fuelbeds and terrestrial laser scan data to develop inputs to fire behavior models. Canadian Journal of Remote Sensing 00: 1-17

Schoennagel, T., T.T. Veblen, and W.H. Romme. 2004. The interaction of fire, fuels, and climate across Rocky Mountain forests. BioScience 54: 661-676.

Silva, S.A. et al. rLiDAR: LiDAR data processing and visualization. 2018, R Package version 0.1.1. Available online: https://cran.r-project.org/package=rLiDAR (accessed on 8 Aug 2019).

Shin, P., T. Sankey, M.M. Moore, and A.E. Thode. 2018. Evaluating unmanned aerial vehicle images for estimating forest canopy fuels in a ponderosa pine stand. Remote Sensing 10: 1266.

Vakili, E., C.M. Hoffman, R.E. Keane, W.T. Tinkham, and Y. Dickinson. 2016. Spatial variability of surface fuels in treated and untreated ponderosa pine forests of the southern Rocky Mountains. International Journal

Westoby, M.J., J. Brasington, N.F. Glasser, M.J. Hambrey, and J.M. Reynolds. 2012. Structure-from-motion photogrammetry: a low-cost, effective tool for geoscience applications. Geomorphology 179: 300-314.

Wimberly, M.C., M.A. Cochrane, A.D. Baer, and K. Pabst. 2009. Assessing fuel treatment effectiveness using satellite imagery and spatial statistics. Ecological Applications 19: 1377-1384.