

Communication

Classifying Large Wildfires in the United States by Land Cover

Brice B. Hanberry

USDA Forest Service, Rocky Mountain Research Station, Rapid City, SD 57702, USA; brice.hanberry@fs.fed.us

Received: 17 July 2020; Accepted: 10 September 2020; Published: 12 September 2020



Abstract: Fire is an ecological process that also has socio-economic effects. To learn more about fire occurrence, I examined relationships between land classes and about 12,000 spatially delineated large wildfires (defined here as uncontrolled fires ≥ 200 ha, although definitions vary) during 1999 to 2017 in the conterminous United States. Using random forests, extreme gradient boosting, and c5.0 classifiers, I modeled all fires, first years (1999 to 2002), last years (2014 to 2017), the eastern, central, and western United States and seven ecoregions. The three classifiers performed well (true positive rates 0.82 to 0.94) at modeling all fires and fires by year, region, and ecoregion. The random forests classifier did not predict to other time intervals or regions as well as other classifiers and models were not constant in time and space. For example, the eastern region overpredicted fires in the western region and models for the western region underpredicted fires in the eastern region. Overall, greater abundance of herbaceous grasslands, or herbaceous wetlands in the eastern region, and evergreen forest and low abundance of crops and pasture characterized most large fires, even with regional differences. The 14 states in the northeastern United States with no or few large fires contained limited herbaceous area and abundant crops or developed lands. Herbaceous vegetation was the most important variable for fire occurrences in the western region. Lack of crops was most important for fires in the central region and a lack of pasture, crops, and developed open space was most important for fires in the eastern region. A combination of wildlands vegetation was most influential for most ecoregions, although herbaceous vegetation alone and lack of pasture, crops, and developed open space also were influential. Despite departure from historical fire regimes, these models demonstrated that herbaceous vegetation remains necessary for fires and that evergreen forests in particular are fire-prone, while reduction of vegetation surrounding housing developments will help provide a buffer to reduce large fires.

Keywords: crops; evergreen; extreme gradient boosting; herbaceous; MTBS; NLCD; random forests; wildlands; WUI

1. Introduction

Wildfire, or generally defined as an uncontrolled or unplanned fire that typically occurs in wildlands (e.g., <https://www.nwcg.gov/glossary/a-z>), regimes have altered over time. Most reports suggest that number, area, and severity of fires are increasing in regions throughout the world [1]. Nonetheless, data may not be available yet to support these trends in some regions and agricultural expansion and intensification may be reducing overall fire activity globally [1,2]. fires appear to have increased in number, area, severity, and season in some parts of the United States during recent decades [3–5]. For example, during 1984–2010, fires ≥ 405 ha (i.e., 1000 acres) covered 24 million ha, or 10.6% of the total burnable area, of the western United States, 3 million ha of which were high severity, or 12.3% of the area burned and 1.3% of the total burnable area, although fires, like other natural disturbances, are patchy [3,6] Fire rotations therefore were about 250 years for all fires and 2100 years for high severity fires [3]. Nonetheless, current fire number and area remain a fraction of

historical fires before fire suppression, which generally became institutionalized during the early 1900s in the United States. Many historical ecosystems had fire regimes that burned every 2 to 25 years [7,8]. In historical ecosystems, abundance of herbaceous vegetation, with either no or limited tree density, meant that fire severity was low, in the form of surface fires with short flame lengths rather than high severity fires in connected tree crowns [8,9].

Land cover, vegetation type and structure, vegetation amount and continuity, fuel properties of vegetation, expansion of native and non-native species that have different fuel properties, and fuel breaks are some factors that influence fires. Recent tree and shrub increase in both eastern and western forests and in central grasslands of the United States conceptually indicate that fires may be less frequent but may be more severe when they do occur [8,10]. Many land uses, such as urbanization, grazing, and agriculture, and some land covers, such as deserts and water, may stop fire spread because they do not contain enough continuous vegetation to carry fires. Furthermore, non-native species have altered fire regime characteristics of spread, ignitability, severity, frequency, and likelihood of fire through providing different abundance and arrangement of fuels for different time intervals [11,12]. Cheatgrass (*Bromus tectorum*) in shrublands of the western United States and eastern redcedar (*Juniperus virginiana*) in central grasslands have increased fine or coarse fuel loads, respectively [9,13]. Risk of fires can be reduced by lessening stand density and fuel loads through preventative thinning or prescribed burning, letting fires burn in wildlands, reducing home ignition susceptibility, and treatment of non-native species [14].

Other factors influence fire at a variety of scales, including topography, temperature, precipitation, wind, humidity, seasonal and annual timing of wet periods that increase biomass and dry periods that cure vegetation to ignitable fuels, climate change, annual number of lightning strikes, and number of ignitions. For example, human-caused ignitions have expanded the fire season, to triple the length of lightning-caused fire season, and now account for 84% of all fires and 44% of total area burned [4]. Wildland fires that are not quickly suppressed often involve extremely high winds, prolonged dry conditions, and accumulated fuels [15]. Fire suppression has become a world-wide response to fire, but fires will occur when weather is severe and deferring fire likely will increase fire severity [1,14].

Models of burn probabilities generated through simulations are valuable tools that account for multiple factors; however, greater understanding of how land cover interacts with fire can be extracted from empirical databases of remotely sensed fires in combination with remotely sensed land cover. Recently, Donovan et al. [9] determined that fires in grasslands of the central United States were more likely to occur when woody vegetation was $\geq 20\%$ of the landscape, with fires unlikely to occur in crops, pasture, and developed areas. Fire datasets continue to accumulate fire occurrences [16] and a revised version of the National Land Cover Database (NLCD) was published during 2020 to update all previously released versions of land cover, which is not directly comparable to previous products [17]. Additionally, statistical modeling with classifiers has advanced rapidly during the past five to ten years to model big datasets that have complex spatiotemporal relationships [18], including refinements of classifiers, such as extreme gradient boosting, which reduces overfitting of models by other boosting classifiers, and has rarely been applied to ecological datasets [19].

Here, I classified the relationship between about 12,000 large wildfires (≥ 200 ha that are not known to be prescribed burns, although definitions of both wildfires and large wildfires may vary; hereafter ‘fire’ or ‘fires’) and newly revised land cover layers to determine land classes for where fires did and did not occur and importance of land class variables for fires. I analyzed different spatiotemporal scales, including the conterminous United States and by eastern, central, and western regions and seven ecoregions for all years between 1999 and 2017 and subdivided into early, middle, and late time intervals. I also assessed performance of the random forests, extreme gradient boosting, and c5.0 classifiers and whether models were consistent in time and space, by hindcasting and forecasting models and translocating models across regions and adjacent ecoregions. Generally, I expected fires to be most frequent in wildlands, particularly herbaceous vegetation, and less abundant in different land uses, such as crops. More formally, I expected to generate useful models that varied spatiotemporally

and a valuable evaluation of the classifiers to answer questions about what are the relationships between fire and land classes, how do these relationships vary over space and time, and how do the classifiers differ in modeling the relationships, particularly when predicting across time and space. Excepting Donovan et al.'s [9] study examining land cover and fires in the central United States, to my knowledge, comparable research is not available.

2. Materials and Methods

First, I divided the conterminous United States into eastern, central, and western regions for modeling by regions (Figure 1). The regions generally followed ecological divisions with some refinements, for example, moving the forested Cross Timbers area from the central region, which is primarily grasslands, to eastern forests [20]. Seven ecological divisions or “ecoregions” had >800 fires and I modeled these ecoregions and excluded ones with <300 fires. I removed 14 states, primarily in the eastern region where <15 fires were recorded during 1999 to 2017 in the fire databases ((Connecticut (0 fires), Illinois (0 fires), New Hampshire (0 fires), Rhode Island (0 fires), Vermont (0 fires), Delaware (1 fire), Iowa (1 fire), Massachusetts (1 fire), Indiana (2 fires), Maine (2 fires), Ohio (2 fires), Wisconsin (6 fires), New York (7 fires), and Pennsylvania (14 fires)).

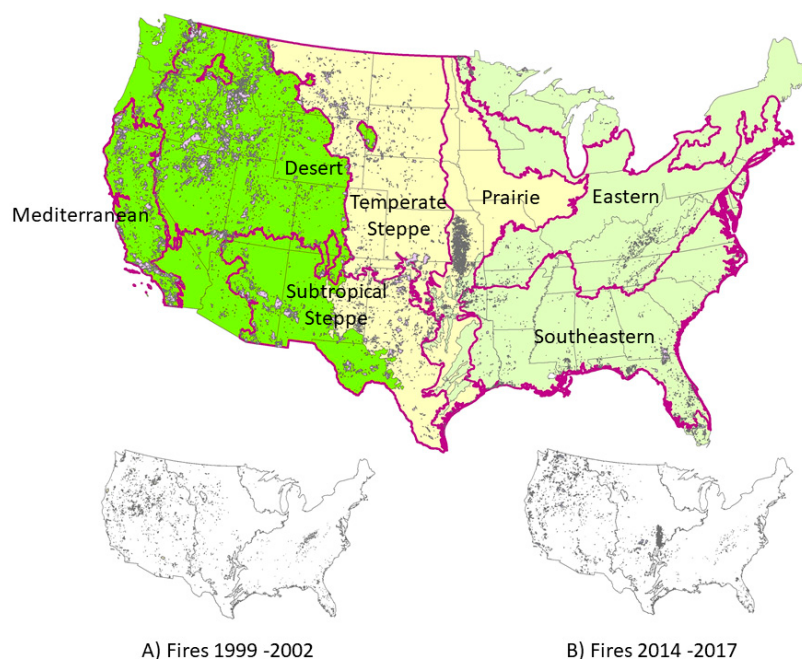


Figure 1. Three regions and seven ecoregions of the conterminous United States with all fires during 1999 to 2017. Below, fires during 1999 to 2002 (A) and fires during 2014 to 2017 (B).

Fire occurrences during 1999 to 2017 were spatially delineated fires of ≥ 200 ha. From the Monitoring Trends in Burn Severity database (MTBS; <https://www.mtbs.gov/>; [21]), I excluded prescribed burns, and I added fires of 200 to 400 ha from the Geospatial Multi-Agency Coordination database (GeoMAC; <https://www.geomac.gov/GeoMACTransition.shtml>; https://data-nifc.opendata.arcgis.com/search?tags=HistoricWildfires_OpenData; 1999 not available) that did not overlap with the MTBS fires; thereby, removing potential duplicates. A total of 11,918 fires occurred: 5895 in the western region, 2767 in the eastern region, and 3256 in the central region. Sample size by years, to match with the National Land Cover Database years, was 2301 fires during 1999–2002 for land cover during 2001, 794 fires during 2003–2004 for land cover during 2004, 1638 fires during 2005–2006 for land cover during 2006, 1235 fires during 2007–2008 for land cover during 2008, 1631 fires during 2009–2011 for land cover during 2011, 1068 fires during 2012–2013 for land cover during 2013, and 3251 fires during 2014–2017 for land cover during 2016.

I used census blocks from the Silvius WUI layer [22] that did not intersect with fires from all years of the MTBS and GeoMAC databases or fire point locations from the Fire Program Analysis fire occurrence database (about 2 million geo-referenced wildfire records during 1992–2015; [23]) for pseudoabsences, or spatial units without known fires [24]. I randomly selected 20,000 of the census blocks without known fire with housing densities $<400/\text{km}^2$, area ≥ 200 ha, and excluding states listed above with few large fires to focus the modeling efforts. To spatially match the regional ratios, the randomly selected pseudoabsences contained 10,500 blocks from the western region, 6000 from the eastern region, and 4000 from the central region. I then populated each of these blocks with land class information for 2001, 2004, 2006, 2008, 2011, 2013, and 2016 (see below; Table 1), which increased the sample size, and then I randomly selected based on matching yearly proportions.

The National Land Cover Database recently was revised to update all previously released versions of landcover, which is not directly comparable to previous products [17]. Updated years were 2001, 2004, 2006, 2008, 2011, 2013, and 2016 and for impervious surface, 2001, 2006, 2011, and 2016 and can be visualized with an interactive viewer (<https://www.mrlc.gov/viewer/>). Model variables were percent impervious surface and 20 land cover variables including land class combinations of wildlands (forests, herbaceous, wetlands, and shrubs), and herbaceous mixtures (Table 1). I calculated mean values of percent area of variables from the NLCD for samples with and without recorded fires, and also background rates by region, ecoregion (excluding states with few fires), the 14 states with few fires, and additionally nine of the fourteen states in the eastern ecoregion.

Supervised classification uses known classes (e.g., fire occurrence) and explanatory variables to develop a relationship to assign unknown observations into classes. For this modeling, I applied c5.0, random forests, and extreme gradient boosting classifiers in the caret package [25,26], which trained the model with 10-fold cross-validation and then predicted for the testing sets. Validation occurs on separate testing data (25% for this modeling), with withheld known classes, to determine how well the classifier assigned classes using explanatory variables. The true positive rate is a measure of the number of predicted observations that were accurately assigned to a class, while the true negative rate is a measure of the number of predicted observations where the response variable of interest did not occur. To some extent, absences are never known and pseudoabsences are proxies for true absences [24]. For this modeling, prevalence or the number of present samples to pseudoabsent samples was equal (e.g., 11,918 known fires and 11,918 census blocks where fire was not recorded). Prevalence, or ratio of present observations to pseudoabsent observations, may not improve the modeling of known instances when more known instances are provided [27]. However, prevalence will affect prediction probabilities, and thus interact with thresholds for prediction of occurrence [27].

I modeled all fires (1999 to 2017) for the conterminous United States, the beginning of 2000 (1999 to 2002), the last years (2014 to 2017), and bracketing 2010 (2009 to 2011). I modeled the eastern, central, and western United States and seven ecoregions. I then examined variable importance. I also assessed how well the models translated in space and time by predicting the years 2014 to 2017 with the 1999 to 2002 model and predicting the years 1999 to 2002 with the 2014 to 2017 model. I additionally used each regional model to make predictions for the other two models and each ecoregional model to predict the most similar adjacent ecoregion.

Table 1. Variables and values (% area) from the National Land Cover Database for recorded fires and samples with no known fires in the conterminous United States by years, region, and ecoregion (years 1999–2017 for each region and ecoregion).

Land class	1999–2017	None	99–02	None	09–11	None	14–17	None	East	None	Central	None	West	None
Crops	0.9	22.5	0.6	22.2	0.9	21.8	1.4	24.7	0.6	17.9	2.0	49.8	0.5	14.4
Deciduous Forest	8.7	7.5	12.3	8.0	8.3	8.2	8.7	6.9	28.7	21.1	6.3	3.3	0.6	1.3
Developed, Low Intensity	0.3	0.7	0.2	0.7	0.3	0.6	0.4	0.7	0.4	1.0	0.3	0.4	0.1	0.6
Developed, Open Space	1.3	2.5	1.1	2.6	1.2	2.5	1.8	2.5	2.1	3.5	1.7	2.7	0.7	1.8
Herbaceous Wetlands	4.0	1.5	5.0	1.5	4.2	1.6	1.9	1.5	14.7	1.8	1.3	1.5	0.3	1.3
Evergreen Forest	14.8	11.5	14.8	11.9	12.6	10.1	12.6	11.0	15.7	12.2	2.2	0.6	21.5	15.4
Firebreak = Water + Snow + Rock/Soil	0.5	1.5	0.5	1.6	0.6	1.6	0.4	1.4	0.8	0.8	0.4	0.6	0.4	2.1
Forest = Deciduous + Evergreen + Mixed	25.2	22.1	28.9	23.0	22.5	21.7	22.6	20.8	50.1	41.2	8.7	4.1	22.6	18.3
Herbaceous Vegetation	47.6	10.8	43.1	10.8	52.1	11.7	51.5	10.5	7.2	3.6	72.2	22.7	53.0	10.5
Herb1 = Grass + Pasture + Crops	50.7	41.7	44.6	42.2	54.7	41.7	57.2	43.4	10.8	38.1	79.3	81.2	53.6	28.3
Herb2 = Grass + Pasture	49.7	19.3	44.0	19.9	53.8	19.9	55.8	18.6	10.2	20.2	77.3	31.4	53.1	13.9
Herb3 = Pasture + Crops + Developed, Open	4.4	33.4	2.5	33.9	3.8	32.5	7.5	35.5	5.6	38.0	8.8	61.2	1.4	19.6
Impervious surface	0.2	0.5	0.2	0.5	0.2	0.4	0.3	0.5	0.3	0.7	0.3	0.3	0.2	0.5
Mixed Forest	1.7	3.1	1.8	3.1	1.6	3.3	1.2	2.9	5.7	8.0	0.3	0.3	0.5	1.6
Pasture	2.1	8.4	0.9	9.1	1.7	8.2	4.3	8.2	2.9	16.6	5.1	8.7	0.1	3.4
Shrub	14.0	26.2	14.6	24.7	11.3	26.4	13.6	25.8	5.0	2.7	7.8	8.9	21.7	46.5
Wetlands = Woody + Herbaceous Wetlands	8.0	5.1	10.1	5.1	9.2	5.4	4.0	5.1	30.7	12.5	1.7	2.0	0.7	2.2
Wildlands1 = Forest + Grass + Shrub + Wetlands	94.8	64.3	96.7	63.6	95.2	65.2	91.6	62.2	93.1	60.0	90.4	37.7	98.1	77.5
Wildlands2 = Forest + Grass + Wetlands	80.8	38.1	82.1	39.0	83.9	38.7	78.0	36.4	88.1	57.3	82.6	28.9	76.4	31.0
Wildlands3 = Forest + Grass	72.8	32.9	72.1	33.9	74.7	33.3	74.0	31.2	57.3	44.8	80.9	26.8	75.6	28.8
Woody Wetlands	4.0	3.6	5.1	3.6	5.0	3.8	2.1	3.6	16.0	10.7	0.4	0.6	0.4	0.9
	Eastern	None	Southeast	None	Prairie	None	Temp. Steppe	None	Subtropical Steppe	None	Desert	None	Mediterranean	None
Crops	1.3	19.3	0.3	16.9	2.1	45.1	2.7	52.9	0.4	7.0	0.5	16.8	0.5	27.0
Deciduous Forest	68.7	38.3	8.9	12.6	13.9	9.0	0.4	0.6	0.9	0.3	0.8	2.2	0.2	0.8
Developed, Low Intensity	0.4	0.9	0.5	1.1	0.5	0.6	0.1	0.3	0.1	0.3	0.1	0.5	0.4	0.7
Developed, Open Space	2.0	3.5	2.3	3.5	2.4	2.9	0.7	2.6	0.7	1.2	0.4	1.5	2.1	3.0
Herbaceous Wetlands	6.4	1.2	16.5	1.5	1.6	1.2	1.2	2.2	0.3	0.2	0.3	1.8	0.4	1.8
Evergreen Forest	2.1	1.8	25.2	20.2	0.4	0.6	5.3	1.1	20.6	5.9	17.4	12.1	22.0	24.6
Firebreak = Water + Snow + Rock/Soil	0.4	0.4	0.9	0.9	0.5	0.8	0.2	0.6	0.1	0.8	0.4	1.6	0.4	1.3
Forest = Deciduous + Evergreen + Mixed	79.4	47.3	38.7	41.4	14.8	11.0	6.1	1.8	21.6	6.3	18.5	14.6	24.0	28.7
Herbaceous Vegetation	4.2	1.1	7.2	3.9	67.6	12.9	77.0	27.8	53.5	12.4	58.6	10.4	48.5	17.0
Herb1 = Grass + Pasture + Crops	8.6	43.7	10.2	33.1	78.8	80.1	80.3	85.0	54.3	20.4	59.2	31.3	49.1	46.4
Herb2 = Grass + Pasture	7.3	24.4	9.9	16.2	76.7	35.1	77.5	32.1	53.8	13.3	58.7	14.5	48.6	19.4
Herb3 = Pasture + Crops + Developed, Open	6.3	46.1	5.2	32.7	13.7	70.2	4.0	59.8	1.4	9.2	1.0	22.4	2.7	32.4
Impervious surface	0.3	0.6	0.3	0.8	0.4	0.5	0.1	0.3	0.1	0.2	0.1	0.4	0.3	0.7
Mixed Forest	8.5	7.2	4.7	8.6	0.4	1.3	0.4	0.1	0.2	0.1	0.3	0.3	1.8	3.3
Pasture	3.0	23.3	2.7	12.3	9.1	22.2	0.6	4.3	0.3	1.0	0.1	4.1	0.1	2.4
Shrub	1.7	0.6	7.4	4.0	0.8	1.7	10.9	6.9	22.4	70.5	21.0	47.4	23.2	17.3
Wetlands = Woody + Herbaceous Wetlands	7.5	3.4	39.9	15.7	2.0	2.6	1.7	2.7	0.8	0.4	0.4	3.0	0.7	2.1
Wildlands1 = Forest + Grass + Shrub + Wetlands	92.8	52.4	93.3	65.0	85.1	28.3	95.7	39.2	98.3	89.6	98.5	75.4	96.4	65.2
Wildlands2 = Forest + Grass + Wetlands	91.2	51.8	85.9	61.0	84.3	26.5	84.8	32.3	75.9	19.0	77.5	28.0	73.3	47.8
Wildlands3 = Forest + Grass	83.6	48.4	46.0	45.2	82.3	23.9	83.1	29.6	75.1	18.7	77.1	24.9	72.5	45.7
Woody Wetlands	1.2	2.2	23.4	14.2	0.3	1.4	0.5	0.5	0.5	0.2	0.2	1.2	0.3	0.3

3. Results

According to differences in mean NLCD values for samples with and without fire presence, fires did not occur in crops, which had values greater by about a factor of 25 for where fires were not known to occur compared to where fires did occur (Table 1). Percent area for the land classes of developed open spaces (lawns, parks, and golf courses), developed low intensity (single-family housing units), impervious surface, and pasture also were greater where fire did not occur. This was true for firebreaks (open water, rock, and soil), aside from the eastern region and the eastern and southeastern ecoregions. Mixed forest had greater values where fire did not occur, except for some ecoregions.

Herbaceous vegetation had greater percent area where fire did occur for all fires, by year, region, and ecoregions. Evergreen forests had greater percent area where fire did occur for all fires, by year, region, and ecoregions, with the exception of the prairie and Mediterranean ecoregions. Wetlands, shrub, and deciduous forest varied regionally, and thus, also by ecoregions. Wetlands and shrub had greater values for where fire did occur in the eastern region. Shrubs had greater percent area where fire did not occur in the subtropical steppe and desert ecoregions particularly. Deciduous forest had greater area where fire did occur in the eastern and central regions but not in all ecoregions.

The 14 U.S. states with no or few fire occurrences, which were excluded from modeling, contained a great percentage of crop, primarily in states on the eastern edge of the central region (Table 2). The 14 U.S. states also had limited area of herbaceous wetlands and herbaceous vegetation; although, not much less area than the eastern ecoregion. Equally, percent area of developed low intensity land was a similar but greater than the eastern region. Otherwise, most values were in between the minimum and maximum values of ecoregion land classes. A subset of nine northeastern states with few fires, where crop area was lower and in the eastern ecoregion, showed reductions in crop area but also even more limited area of herbaceous wetlands and herbaceous vegetation and greater percent area of developed lands.

The three classifiers performed well (true positive rates 0.85 to 0.92) at modeling all fires and fires by year and region (Table 3). The 2014 to 2017 models predicted 1999 to 2002 fires well (true positive rates 0.81 to 0.84), but 1999 to 2002 models were less successful (true positive rates 0.71 to 0.75) at forecasting 2014 to 2017 fires (Table 4). Regionally, models for the central region generally translocated well with both the eastern and western regional models (true positive rate mean = 0.77), but models for the eastern region overpredicted fires in the western region and likewise, models for the western region underpredicted fires in the eastern region.

The three classifiers also performed well (true positive rates 0.82 to 0.94) at modeling fires by ecoregion (Table 5). The eastern and prairie ecoregions had greatest true positive rates, while the desert ecoregion, which encompasses both shrublands and forests, had the least true positive rates. Ecoregion models generally translocated well to adjacent ecoregions, although with some under- and overpredicting (the eastern and southeastern ecoregional pairs and temperate steppe to prairie).

A combination of forests, herbaceous vegetation, and wetlands land classes was the most influential variables for all fires and most years (Table 6; supplementary figures have partial dependence plots). Herbaceous vegetation was the most important variable for 2014–2017 fires and in the western region. Crops were most important in the central region and a combination of pasture, crops, and developed, open space were important in the eastern region; these variables had much greater values where areas where fire was not recorded (Table 1).

By ecoregion, the three different wildlands combinations, herbaceous vegetation, and the herbaceous land use combination of pasture, crops, and developed open area were most important, which produced overlap in ecoregions, even across regions such as the southeastern and temperate steppe ecoregions (Table 7). Except for herbaceous vegetation, single land classes overall were not influential. For example, even though deciduous forest had greater area where fire did occur in the eastern ecoregion, the classifiers relied on wildlands variables that included deciduous forest.

Table 2. Variables and values (% area) from the 2008 National Land Cover Database for 14 states with few fires, northeastern states with few fires in the eastern ecoregion, and ecoregions (years 1999–2017; ecoregions exclude 14 states with few fires).

Land class	States-few fires	Eastern-few fires	Eastern	Southeastern	Prairie	Temperate Steppe	Subtropical Steppe	Desert	Mediterranean
Crops	32.6	10.6	12.4	12.8	34.5	30.0	6.5	4.1	11.3
Deciduous Forest	23.6	33.4	41.6	10.2	9.7	0.6	0.8	2.3	0.7
Developed, Low Intensity	2.9	5.0	2.5	2.6	1.9	0.5	0.5	0.4	1.7
Developed, Open Space	4.8	7.9	5.2	5.1	3.8	1.8	1.2	0.7	3.4
Herbaceous Wetlands	1.0	0.8	1.2	2.8	1.6	1.8	0.2	0.8	1.0
Evergreen Forest	4.1	3.7	2.0	19.1	1.2	1.5	9.9	23.9	29.2
Firebreak = Water + Snow + Rock/Sand/Clay	2.7	4.2	2.2	3.8	2.4	1.3	1.0	3.5	2.4
Forest = Deciduous + Evergreen + Mixed Forest	37.1	49.0	51.8	37.0	12.4	2.2	10.9	26.5	32.7
Herbaceous Vegetation	1.0	0.8	1.4	3.5	17.4	47.8	15.2	13.6	17.4
Herb1 = Grass + Pasture + Crops	43.4	22.5	32.6	27.1	72.1	80.3	23.0	18.7	30.0
Herb2 = Grass + Pasture	10.8	11.9	20.2	14.3	37.6	50.3	16.4	14.7	18.7
Herb3 = Pasture + Crops + Developed, Open	47.2	29.6	36.4	28.6	58.5	34.3	9.0	5.9	15.9
Mixed Forest	9.3	11.9	8.3	7.7	1.5	0.1	0.1	0.3	2.9
Pasture	9.8	11.1	18.8	10.8	20.2	2.4	1.3	1.1	1.3
Shrub	0.8	0.6	0.8	4.5	2.5	11.4	62.7	48.6	26.0
Wetlands = Woody + Herbaceous Wetlands	6.5	6.5	3.5	18.8	3.6	2.3	0.5	1.4	1.3
Wildlands1 = Forest + Grass + Shrub + Wetlands	45.5	57.0	57.5	63.8	35.9	63.7	89.3	90.1	77.4
Wildlands2 = Forest + Grass + Wetlands	44.7	56.4	56.7	59.3	33.3	52.3	26.5	41.5	51.4
Wildlands3 = Forest + Grass	38.1	49.9	53.3	40.5	29.8	50.0	26.0	40.1	50.1
Woody Wetlands	5.6	5.7	2.3	16.0	2.0	0.5	0.3	0.6	0.3

Table 3. True positive rate and true negative rate of predictions for three different classifiers (C5.0; rf = random forests, xgbTree = extreme gradient boosting with trees) for test set (25% of observations) based on modeled training data of all fire occurrences during 1999 to 2017, 1999 to 2002, 2009 to 2011, 2014 to 2017, and three regions.

	All fires	Fire 1999–2002	Fire 2009–2011	Fire 2014–17	East	Central	West	Mean
C5.0								
True positive rate	0.876	0.864	0.877	0.889	0.897	0.907	0.854	0.881
True negative rate	0.864	0.831	0.855	0.858	0.852	0.919	0.869	0.864
rf								
True positive rate	0.871	0.894	0.894	0.882	0.893	0.918	0.846	0.885
True negative rate	0.868	0.852	0.848	0.860	0.837	0.909	0.881	0.865
xgbTree								
True positive rate	0.867	0.875	0.880	0.881	0.887	0.897	0.856	0.877
True negative rate	0.858	0.830	0.862	0.853	0.851	0.915	0.866	0.862

Table 4. True positive rate and true negative rate of predictions for three different classifiers (C5.0; rf = random forests, xgbTree = extreme gradient boosting with trees) of modeled training data for different years and regions.

Model	Fire 1999–2002	Fire 2014–17	East		Central		West	
Predict	2014–17	1999–2002	Central	West	East	West	Central	East
C5.0								
True positive rate	0.718	0.844	0.750	0.981	0.779	0.867	0.685	0.131
True negative rate	0.864	0.844	0.836	0.270	0.725	0.723	0.877	0.987
rf								
True positive rate	0.710	0.811	0.474	0.812	0.867	0.887	0.762	0.128
True negative rate	0.870	0.846	0.837	0.264	0.715	0.685	0.873	0.985
xgbTree								
True positive rate	0.749	0.830	0.813	0.983	0.731	0.870	0.719	0.131
True negative rate	0.858	0.836	0.804	0.240	0.767	0.716	0.885	0.980

Table 5. True positive rate and true negative rate of predictions for three different classifiers (C5.0; rf = random forests, xgbTree = extreme gradient boosting with trees) for test set (25% of observations) and for adjacent ecoregions based on modeled training data of all fire occurrences during 1999–2017 for seven ecoregions.

Test set	Eastern	Southeastern	Prairie	Temperate Steppe	Subtropical Steppe	Desert	Mediterranean	Mean
C5.0								
True positive rate	0.936	0.893	0.926	0.900	0.865	0.810	0.883	0.888
True negative rate	0.885	0.837	0.884	0.910	0.900	0.850	0.897	0.880
rf								
True positive rate	0.927	0.912	0.937	0.886	0.863	0.835	0.883	0.892
True negative rate	0.890	0.846	0.859	0.895	0.908	0.814	0.899	0.873
xgbTree								
True positive rate	0.922	0.873	0.935	0.910	0.884	0.821	0.893	0.891
True negative rate	0.881	0.861	0.879	0.905	0.876	0.839	0.888	0.875
Other ecoregions	Eastern to Southeastern	Southeastern to Eastern	Prairie to Temperate Steppe	Temperate Steppe to Prairie	Subtropical Steppe to Temperate Steppe	Desert to Mediterranean	Mediterranean to Desert	Mean
C5.0								
True positive rate	0.959	0.798	0.912	0.574	0.892	0.706	0.873	0.816
True negative rate	0.675	0.915	0.829	0.956	0.847	0.822	0.720	0.823
rf								
True positive rate	0.959	0.776	0.882	0.661	0.898	0.700	0.862	0.820
True negative rate	0.686	0.915	0.825	0.945	0.847	0.827	0.615	0.809
xgbTree								
True positive rate	0.953	0.767	0.898	0.551	0.916	0.738	0.882	0.815
True negative rate	0.701	0.913	0.845	0.960	0.841	0.787	0.660	0.815

Table 6. Most important variables and value for random forests (rf) and extreme gradient boosting (xgbTree) models for all fires and by year and region (Wildlands1= Forest + Grass + Shrub + Wetlands; Wildlands2 = Forest + Grass + Wetlands; Wildlands3 = Forest + Grass; Herb3 = Pasture + Crops + Developed, Open; and Herb2 = Grass + Pasture).

All fires		1999–2002		2009–2011		2014–17		East		Central		West	
rf													
Wildlands2	100	Wildlands2	100	Wildlands2	100	Herbaceous	100	Herb3	100	Crops	100	Herbaceous	100
Wildlands3	75	Wildlands1	71	Wildlands3	66	Wildlands2	24	Wildlands1	48	Wildlands2	82	Wildlands3	87
Herbaceous	66	Wildlands3	67	Herbaceous	62	Wildlands1	18	Wildlands2	12	Wildlands3	69	Herb2	85
Wildlands1	58	Herb3	64	Wildlands1	52	Impervious	14	Mixed Forest	11	Wildlands1	66	Wildlands2	81
Herb3	55	Herbaceous	51	Herb3	45	Crops	12	Impervious	11	Herb3	61	Wildlands1	56
Herb2	53	Herb2	36	Herb2	44	Wildlands3	7	Herbaceous Wetlands	8	Herb2	49	Herb1	51
Shrub	31	Shrub	34	Shrub	39	Shrub	7	Deciduous Forest	7	Herbaceous	44	Shrub	43
xgbTree													
Wildlands2	100	Wildlands2	100	Wildlands2	100	Herbaceous	100	Herb3	100	Crops	100	Herbaceous	100
Herbaceous	34	Wildlands1	52	Herbaceous	22	Wildlands3	30	Herbaceous Wetlands	14	Wildlands2	28	Wildlands3	55
Wildlands1	17	Shrub	27	Wildlands1	13	Wildlands2	28	Mixed Forest	12	Herbaceous	13	Wildlands1	25
Crops	15	Herb3	21	Impervious	12	Wildlands1	25	Impervious	11	Developed, Open	12	Evergreen Forest	17
Mixed Forest	12	Forest	18	Shrub	6	Crops	24	Deciduous Forest	10	Wildlands1	10	Herb3	13
Impervious	8	Wildlands3	15	Crops	6	Impervious	16	Wildlands2	9	Firebreak	8	Herb2	11
Herb3	7	Herbaceous	13	Wildlands3	6	Herb2	13	Wildlands1	9	Shrub	7	Impervious	9

Table 7. Most important variables and value for random Forests (rf) and extreme gradient boosting (XgbTree) models by ecoregion (Wildlands1= Forest + Grass + Shrub + Wetlands; Wildlands2 = Forest + Grass + Wetlands; Wildlands3 = Forest + Grass; Herb3 = Pasture + Crops + Developed, Open; and Herb2 = Grass + Pasture).

Eastern broadleaf		Southeastern mixed		Prairie		Temperate Steppe		Tropical/Subtropical Steppe		Temperate Desert		Mediterranean	
rf													
Wildlands1	100	Herb3	100	Wildlands2	100	Herb3	100	Wildlands3	100	Herbaceous	100	Herbaceous	100
Herb3	99	Wildlands1	53	Herb2	14	Wildlands1	90	Wildlands2	99	Wildlands1	97	Wildlands3	87
Wildlands2	86	Herbaceous Wetlands	17	Herbaceous	4	Crops	80	Shrub	78	Herb3	79	Wildlands2	84
Wildlands3	71	Deciduous Forest	14	Impervious	3	Wildlands2	70	Herbaceous	56	Impervious	43	Herb2	72
Pasture	70	Wildlands2	13	Crops	3	Wildlands3	51	Herb2	48	Deciduous Forest	36	Wildlands1	66
Herb1	63	Impervious	12	Wildlands3	3	Herbaceous	51	Herb1	44	Evergreen Forest	32	Herb1	48
Forest	46	Developed, Low	10	Developed, Low	3	Herb2	45	Forest	32	Shrub	32	Herb3	45
xgbTree													
Wildlands2	100	Wildlands1	100	Wildlands2	100	Wildlands1	100	Wildlands2	100	Herb3	100	Herbaceous	100
Wildlands1	41	Herb3	97	Crops	16	Herb3	64	Shrub	14	Herbaceous	90	Wildlands3	80
Herb3	36	Herbaceous Wetlands	30	Herb2	15	Wildlands2	16	Wildlands3	11	Wildlands1	87	Wildlands1	34
Herbaceous Wetlands	22	Deciduous Forest	25	Herbaceous	15	Crops	13	Wildlands1	11	Deciduous Forest	50	Herb3	16
Mixed Forest	13	Developed, Low	19	Impervious	9	Impervious	10	Impervious	10	Shrub	27	Impervious	11
Shrub	11	Shrub	17	Shrub	7	Shrub	10	Forest	6	Evergreen Forest	27	Wetlands	10
Pasture	10	Impervious	15	Wildlands1	7	Herb1	9	Herbaceous	5	Impervious	22	Shrub	5

4. Discussion

4.1. Key Findings

Greater abundance of herbaceous vegetation, forests, and wetlands and low abundance of crops, pasture, and developed open space (lawns, parks, and golf courses) generally explained large fire occurrence (wildfires ≥ 200 ha), with pronounced regional differences and less distinctive ecoregional differences. More specifically, herbaceous vegetation and herbaceous wetlands and evergreen forest classes were the primary wildland classes that carried fire. These results may help explain the lack of few large fires in 14 states, which contained limited herbaceous vegetation, either as grasslands or wetlands, and an abundance of crops or developed areas, although related factors such as fragmentation from small land parcels likely are influential in preventing fires. Herbaceous vegetation was important for fires in all regions, but in the eastern region, herbaceous vegetation was present in the form of emergent herbaceous wetlands instead of upland grasslands (Table 1). Although wetlands may be a firebreak when water is present, the herbaceous component of emergent herbaceous wetlands in the eastern region appears to be fire-prone during dry events. Wetlands were less likely to have fires in the central and western regions. Forests overall were more likely to have fires in all regions due to evergreen forests and not mixed or deciduous forests, which were less likely to have fires where these forest types were more influential variables; that is, in the eastern region and southeastern and desert ecoregions. Evergreen forests with low tree densities and an herbaceous layer, which typically result in low severity fires, occur as widespread ponderosa pine (*Pinus ponderosa*) in the western region and now relatively rare longleaf and shortleaf pine (*Pinus palustris* and *echinata*) forests and commercial pine plantations in the eastern region [8]. More densely spaced evergreen species such as spruces, firs, and eastern redcedar, which also have low branches that provide ladder fuels to carry fire vertically to crowns, typically generate higher severity crown fires [9].

Crops and pasture were the primary land classes where fire did not occur, with developed open space as a less abundant class that also appeared to prevent fire spread. These land classes can be difficult to separate from herbaceous vegetation using remote sensing methods, but differentiation was very clear in where fires did and did not occur (Table 1). Although these land classes can have fires, fires often are prescribed and small to remove vegetation residue, with limited fuels available to spread fire or escalate fire severity [2,28,29]. Crops, pasture, and developed open space are areas of reduced vegetation, in the gradient between wildlands and developed urban areas with little vegetation. Crops and pasture are harvested or grazed before plants dry and lose enough moisture to ignite easily. Lawns, parks, and golf courses similarly are kept green through watering and low in height through mowing, which limits fuels, and are intersected with impervious surfaces that fragment fuel continuity and prevent fire spread.

Land classes such as crops, pasture, and developed open space with negative influences on fire were as important for where fire did not occur as herbaceous vegetation, forests, and wetlands were greater in abundance where fire did occur, particularly in the eastern and central regions and the southeastern and temperate steppe ecoregions of these regions. In the eastern region, the most influential model variable was a combination of crops, pasture, and developed open space. These land classes covered 32% of the region, 6% of the area where fires occurred, and 38% of the area from samples where fires did not occur. Herbaceous wetlands were among the most important model variables. Herbaceous wetlands were 2% of the region and area from samples where fires did not occur, and 15% of the area where fires occurred. Likewise, in the central region, crops were the most important model variable. Crops were 33% of the region, 2% of the area where fires occurred, and 50% of the area from samples where fires did not occur. Herbaceous vegetation was among the most important model variables, albeit still minor. Herbaceous vegetation was 32% of the area, 72% of the area where fires occurred, and 23% of the area from samples where fires did not occur.

Furthermore, in the central region, forests or evergreen forests, which notably would include eastern redcedar, were not among most influential model variables, unlike grassland herbaceous

vegetation that also was more disproportionately represented where fires occurred (2.2 times greater) compared to background land class rates than other wildlands vegetation types (e.g., evergreen forests were greater by a factor of 1.5). By ecoregion, the prairie in particular showed little association of evergreen forests with fire. However, the temperate steppe did contain greater percent area of evergreen forests (5.5%) where fire occurred than background rates (1.5%), with percent area of forests with fires about 2.8-fold greater than background rates. Conversely, herbaceous vegetation was 77% of area where fire occurred compared to 48% of the region, which is a grassland with a great herbaceous vegetation background rate. Nonetheless, the overall wildlands (herbaceous and woody vegetation) and land use (crop and pasture) combinations most influenced fire. These results contrasted with Donovan et al. [9], who found that woody vegetation burned disproportionately more than other land-use types in the Great Plains. However, Donovan et al. [9], expected to find that large fires occurred in grasslands, which are more fire-prone than woodlands.

The shrub land class was minimally influential on fire, but perhaps that will change due to recent fires in the shrub chaparral of California. Fire-dependent shrub ecosystems also occur in shrubs of the Lake Wales Ridge in Florida and pine barrens of the mid-Atlantic eastern US, and albeit small remnants, these types of ecosystems are examples to explain greater abundance of shrubs where fires occurred in the eastern region. Additionally, increased fire activity occurs because of invasion by primarily non-native cheatgrass in shrublands of the Great Basin in the desert ecoregion and a complex of non-native grass invasions in shrublands of the subtropical ecoregion; non-native grasses provide horizontal fuel continuity in fuel-limited shrub systems [13]. However, fire then occurs in predominantly herbaceous vegetation rather than shrubs.

Fire relationships were not consistent in space, although the central region shared overlap with both the eastern and western regions. Models for the eastern region overpredicted fires in the western region and likewise, models for the western region underpredicted fires in the eastern region. That is, vegetation that would produce a fire in the eastern region was not sufficient for fire in the western region and vegetation that would not produce a fire in the western region was sufficient for fire in the eastern region. These differences in part may be due to more land classes associated with fires in the eastern region, including wetlands and deciduous forest, than in the western region. These results also seem to reflect that the western region typically has a drier climate, which can result in fuel limitations in vegetation but also greater probability of vegetation being cured and able to ignite for a longer time interval and extreme weather conditions suitable for fire [3,30]. Nonetheless, despite greater current rates of fires in the west, historically, the longleaf pine “grasslands” of the Coastal Plain in the southeastern U.S. had fire regimes of 2 to 5 years, making it one of the most frequently burned regions in the past, despite a humid, wet climate [8,9]. To a lesser magnitude, predictions among nearby ecoregions had the same problem, in that eastern ecoregional models overpredicted fires for the southeastern ecoregion and the southeastern models underpredicted fire for the eastern ecoregion and the temperate steppe also underpredicted fire for the prairie region.

Fire relationships were not consistent in time. The 1999 to 2002 models were less successful at forecasting 2014 to 2017 fires than 2014 to 2017 models predicted 1999 to 2002 fires. Number of fires greatly increased in the central region from 310 fires during 1999 to 2002 to 1569 fires during 2014 to 2017, while number of fires in the other regions stayed relatively stable. Compared to 1999 to 2002 models, importance of herbaceous and crop land covers increased for the 2014 to 2017 models, which still remained successful at hindcasting 1999 to 2002 fires.

The three classifiers performed near equally well at predicting based on models of the same extent or time interval, or similar ecoregions, despite a variety of potential errors arising from for example, the fire perimeters, land class misclassification, and areas that could have burned yet did not. However, in general, land classes were well-differentiated between where fire occurred and samples of where fire was not known to occur (Table 1). Predictions across time or space in different regions showed that random forests did not perform as well as the other two classifiers. Mean value of true positive rate for extreme gradient boosting was 0.790 compared 0.761 for random forests when predicting across

time intervals. Mean value of true positive rate for extreme gradient boosting was 0.708 compared 0.655 for random forests when predicting across regions, albeit if the random forests prediction for the central region based on the eastern region is excluded, the mean value increases to 0.691. These two examples demonstrate that extreme gradient boosting remained more generalized with less overfitting to specific datasets than random forests, although the sample size is too small to state that extreme gradient boosting overall is preferable to random forests when models are in development or need flexibility. Moreover, random forests performed equally well in predicting fires in adjacent ecoregions, which had fewer differences than across regions.

Analysis with multiple classifiers allows identification of better performing classifiers and selection of desirable features. Random forests is a well-established classifier with reliably accurate performance, whereas C5.0 decision trees generate explicit rulesets, which can be help explain models when there are few variables, but also produce many similar and competing models. Fernandez-Delgado et al. [18] comprehensively assessed 179 classifiers using 121 data sets and the best classifier was random forests, with C5.0 decision trees among the best classifiers. Extreme gradient boosting is newer classifier, intended to reduce overfitting that other boosting classifiers are prone to producing [19].

Further research may explore some unexpected results from these relationships. For example, future studies can examine the interaction of fire with mixed and deciduous forests, which overall were not associated with fires. These forest types were less abundant where fires occurred when these forests were among the most influential model variables; for example, mixed forests in the eastern region and deciduous forests in the southeastern and desert ecoregions. It may be that current species in mixed and deciduous forests are not fire-tolerant, and thus suppress rather than support fires, reducing fire frequency but increasing fire risk when fire does occur [31,32]. Furthermore, presence of water in mixed and deciduous forests will act as a firebreak. Deciduous species are rare in the western region and often occur along riparian networks, and likewise, deciduous and mixed forests in the southeastern U.S. often form along wetlands. Similarly, the importance of herbaceous wetlands, as a substitute for herbaceous vegetation, for fires in the eastern region also may be assessed to explore the interaction between the fire-spreading herbaceous component and the fire-breaking wetlands component.

Although this study provides summaries that clearly show fires are more prevalent in some land classes than others, varying by region, along with accurate predictions of models at least within modeling extents, limitations include errors in data and also that land cover is just one of many factors that influence fire. fires vary in severity and include unburned areas within perimeters, fires may not be detected, land classes may be misidentified, and most areas that are flammable did not burn during the time interval of this study [5]. fires also are influenced by time since last fire, vegetation properties in terms of acting as fuels, weather and climate, number of ignitions, and topography. Additionally, spatial arrangement of land classes also is an important component to whether fire will be able to spread across flammable wildlands vegetation or be limited by interspersed firebreaks that break fuel continuity. Furthermore, fire protection services in conjunction with wildlands fragmentation and lack of extreme weather may suppress fires while they are small. For example, some of the 14 states with few large fires, particularly New York, indeed have many ignitions, resulting in numerous fires <200 ha [23,33]. Land management through prescribed burns also will reduce fuels and reduce the risk of fire.

4.2. Implications

Fire is a fundamental ecological process that maintains ecosystems and associated species by filtering species that can survive fire or reproduce in response to fire. Fire favors herbaceous vegetation, which rapidly captures aboveground growing space, over woody vegetation, and thereby supports a diverse and unique assemblage of species, including insects that provide many critical ecosystem services [34–37]. Hominins have applied fire as a labor-saving land management tool for hundreds of thousands of years [38]. However, as human densities have increased, fire has become a risk to human structures and lives. Fire had been suppressed for about a century in the United States, resulting in

great departure from historical fire regimes. Nonetheless, these models demonstrated that fires still follow historical ecology that herbaceous vegetation and pine forests are fire-prone, with traits that promote fire [8].

Although fire suppression appears to be a logical response to fire, fire ecology produces a paradox that fires become more catastrophic because of, not despite, fire suppression [14,39,40]. Fuels accumulate, increasing in quantity and continuity, when fuels do not burn due to suppression of unintended fires, exclusion of prescribed fires for land management, along with passive suppression from land classes that act as fire breaks (described in this study). Changing characteristics of fires often culminates in disastrous fires that cannot be suppressed [14]. Fire suppression produces an escalating cycle of need for fire suppression and increased fire suppression costs [14]. This aspect of fire ecology has been understood for over a century. For example, prescribed burns were established practice during the late 1800s in California forests to prevent high severity crown fires [41]. Shaw and Kotok (1924, [41]) wrote,

The idea that fires could be excluded entirely from millions of acres was generally regarded as preposterous and the most gloomy pictures were drawn of any such attempt. It was claimed that the uncontrollable crown fire was to be expected as the inevitable consequence of allowing ground cover and litter to accumulate. Thus, in the early years of protection of the national forests, the forests were still open as a result of the repeated fires of the past As fire protection became an accomplished fact and the young growth began to fill up the open forest, the amount of inflammable material in the forests increased greatly.

Compounding altered fire regimes, housing developments and other infrastructure have expanded into wildlands area, increasing the area of wildland-urban interface (WUI) and intermix and concurrently, exposure of assets to fire throughout the world and resulting in structural losses to fire events [14]. Fire severity in wildlands may be relatively independent of fire in the home ignition zone; that is, homeowner management of the area immediately surrounding the determines home losses from wildland fire [14]. Thus, in addition to untreated fuel accumulation in wildlands and increased exposure and ignition rates due to development in wildlands, many residents in wildlands underinvest in mitigation actions, relying on neighboring public or private land owners for fuel treatment, government emergency services to suppress fire, and insurance or federal disaster assistance to cover any property losses [42,43]. Housing materials, such as wood shakes, can be very flammable and burning homes generate hot embers, which increase the spread of a fire. Lofted embers may cause more residential losses than direct flames or radiant heat [14].

Strategies for risk mitigation to coexist with fire include identification of communities at risk, disclosure of fire risk, reduction of vegetation (i.e., fuel for fire), outreach to homeowners to reduce both structure flammability and vegetation surrounding structures, and planning and zoning to reduce sprawl into wildlands [33]. For an example of outreach and risk disclosure, counties can provide advisory letters informing citizens about risks and homeowner responsibilities, similar to the following, “The physical characteristics of your property can be positive and negative. Trees are a wonderful environmental amenity, but can also involve your home in a forest fire. Building at the top of a forested draw should be considered as dangerous as building in a flash flood area. ‘Defensible perimeters’ are very helpful in protecting buildings from forest fire and inversely can protect the forest from igniting if your house catches on fire.” (county commissioners in Stevens County, MT, USA, as cited by [44]). Reduction of intermixing between wildlands and housing, by concentration of housing at greater densities, will reduce fire exposure. These models show that creative planning and zoning may allow placement of firebreak buffer zones of crops, grazed pastures, and open spaces that are managed by mowing and watering between wildlands and developed land classes with housing.

5. Conclusions

Information can be extracted from statistical relationship among remotely sensed imagery using many new classifier methods. While extreme gradient boosting overall did not perform better than the more established methods of random forests and c5.0 in predicting to withheld samples from models, extreme gradient boosting had improved predictions across time and space compared to random forests and produced different models that may be preferable in some cases. Based on these results, with agreement among the classifiers, the primary wildlands classes that carried fire were herbaceous grasslands and herbaceous wetlands and evergreen forest, while crops and pasture were the primary land classes that acted as firebreaks. Research needs include greater examination of the interaction between fires and mixed and deciduous forests. Despite departure from historical fire regimes, these models confirmed that herbaceous vegetation remains necessary for fires and that evergreen forests in particular are fire-prone. Reduction of vegetation surrounding housing developments, perhaps through low intensity grazing, will help provide a buffer to reduce large fires.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-4292/12/18/2966/s1>, Figure S1: Partial dependence plots for c5.0, random forests, and extreme gradient boosting models of all fires in the conterminous United States, Figure S2: Partial dependence plots for c5.0, random forests, and extreme gradient boosting models of the eastern United States, Figure S3: Partial dependence plots for c5.0, random forests, and extreme gradient boosting models of the central United States, and Figure S4: Partial dependence plots for c5.0, random forests, and extreme gradient boosting models of the western United States.

Funding: This research received no external funding.

Acknowledgments: I thank three reviewers for their time and comments.

Conflicts of Interest: The author declares no conflict of interest.

References

- Doerr, S.H.; Santin, C. Global trends in wildfire and its impacts: Perceptions versus realities in a changing world. *Philos. Trans. R. Soc. B Biol. Sci.* **2016**, *371*, 20150345. [CrossRef]
- Andela, N.; Morton, D.C.; Giglio, L.; Chen, Y.; Van Der Werf, G.R.; Kasibhatla, P.S.; DeFries, R.S.; Collatz, G.J.; Hantson, S.; Kloster, S.; et al. A human-driven decline in global burned area. *Science* **2017**, *356*, 1356–1362. [CrossRef] [PubMed]
- Liu, Z.; Wimberly, M.C. Climatic and landscape influences on fire regimes from 1984 to 2010 in the western United States. *PLoS ONE* **2015**, *10*, e0140839. [CrossRef] [PubMed]
- Balch, J.K.; Bradley, B.A.; Abatzoglou, J.T.; Nagy, R.C.; Fusco, E.J.; Mahood, A.L. Human-started wildfires expand the fire niche across the United States. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2946–2951. [CrossRef] [PubMed]
- Donovan, V.M.; Wonkka, C.L.; Twidwell, D. Surging wildfire activity in a grassland biome. *Geophys. Res. Lett.* **2017**, *44*, 5986–5993. [CrossRef]
- Kolden, C.A.; Smith, A.M.; Abatzoglou, J.T. Limitations and utilisation of Monitoring Trends in Burn Severity products for assessing wildfire severity in the USA. *Int. J. Wildland Fire* **2015**, *24*, 1023–1028. [CrossRef]
- Landfire, 2020. Biophysical Settings. Available online: https://www.landfire.gov/national_veg_models_op2.php (accessed on 4 May 2020).
- Hanberry, B.B.; Bragg, D.C.; Alexander, H.D. Open forest ecosystems: An excluded state. *For. Ecol. Manag.* **2020**, *472*, 118256. [CrossRef]
- Donovan, V.M.; Wonkka, C.L.; Wedin, D.A.; Twidwell, D. Land-use type as a driver of large wildfire occurrence in the U.S. Great Plains. *Remote Sens.* **2020**, *12*, 1869. [CrossRef]
- Hanberry, B.B.; Hansen, M.H. Advancement of tree species across ecotonal borders into non-forested ecosystems. *Acta Oecologica* **2015**, *68*, 24–36. [CrossRef]
- D’Antonio, C.M.; Vitousek, P.M. Biological invasions by exotic grasses, the grass/fire cycle, and global change. *Annu. Rev. Ecol. Syst.* **1992**, *23*, 63–87. [CrossRef]
- Lippincott, C.L. Effects of *Imperata cylindrica* (L.) Beauv. (Cogongrass) invasion on fire regime in Florida Sandhill (USA). *Nat. Areas J.* **2000**, *20*, 140–149.

13. Balch, J.K.; Bradley, B.A.; D'Antonio, C.M.; Gómez-Dans, J. Introduced annual grass increases regional fire activity across the arid western USA (1980–2009). *Glob. Chang. Biol.* **2013**, *19*, 173–183. [[CrossRef](#)] [[PubMed](#)]
14. Calkin, D.E.; Cohen, J.D.; Finney, M.A.; Thompson, M.P. How risk management can prevent future wildfire disasters in the wildland-urban interface. *Proc. Natl. Acad. Sci. USA* **2013**, *111*, 746–751. [[CrossRef](#)] [[PubMed](#)]
15. Finney, M.A. The wildland fire system and challenges for engineering. *Fire Saf. J.* **2020**, 103085, in press.
16. Long, T.; Zhang, Z.; He, G.; Jiao, W.; Tang, C.; Wu, B.; Zhang, X.; Wang, G.; Yin, R. 30 m Resolution Global Annual Burned Area Mapping Based on Landsat Images and Google Earth Engine. *Remote Sens.* **2019**, *11*, 489. [[CrossRef](#)]
17. Homer, C.; Dewitz, J.; Jin, S.; Xian, G.; Costello, C.; Danielson, P.; Gass, L.; Funk, M.; Wickham, J.; Stehman, S.; et al. Conterminous United States land cover change patterns 2001–2016 from the 2016 National Land Cover Database. *ISPRS J. Photogramm. Remote Sens.* **2020**, *162*, 184–199. [[CrossRef](#)]
18. Fernández-Delgado, M.; Cernadas, E.; Barro, S.; Amorim, D. Do we need hundreds of classifiers to solve real world classification problems? *J. Mach. Learn. Res.* **2014**, *15*, 3133–3181.
19. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining—KDD '16, San Francisco, CA, USA, 13–17 August 2016; ACM Press: New York, NY, USA, 2016; pp. 785–794. [[CrossRef](#)]
20. Cleland, D.T.; Freeouf, J.A.; Keys, J.E., Jr.; Nowacki, G.J.; Carpenter, C.; McNab, W.H. *Ecological Subregions: Sections and Subsections of the Conterminous United States [1:3,500,000]*. U.S.; Department of Agriculture, Forest Service: Washington, DC, USA, 2007.
21. Picotte, J.J.; Bhattarai, K.; Howard, D.; Lecker, J.; Epting, J.; Quayle, B.; Benson, N.; Nelson, K. Changes to the Monitoring Trends in Burn Severity program mapping production procedures and data products. *Fire Ecol.* **2020**, *16*, 1–12. [[CrossRef](#)]
22. Radeloff, V.C.; Helmers, D.P.; Kramer, H.A.; Mockrin, M.H.; Alexandre, P.M.; Bar-Massada, A.; Butsic, V.; Hawbaker, T.J.; Martinuzzi, S.; Syphard, A.; et al. Rapid growth of the US wildland-urban interface raises wildfire risk. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 3314–3319. [[CrossRef](#)]
23. Short, K.C. *Spatial Wildfire Occurrence Data for the United States, 1992–2015 [FPA_FOD_20170508]*, 4th ed.; Forest Service Research Data Archive: Fort Collins, CO, USA, 2017.
24. Hanberry, B.B.; He, H.S.; Palik, B.J. Pseudoabsence generation strategies for species distribution models. *PLoS ONE* **2012**, *7*, e44486. [[CrossRef](#)]
25. Kuhn, M. Building Predictive Models in R Using the caret Package. *J. Stat. Softw.* **2008**, *28*, 1–26. [[CrossRef](#)]
26. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.
27. Hanberry, B.B.; He, H.S. Prevalence, statistical thresholds, and accuracy assessment for species distribution models. *Web Ecol.* **2013**, *13*, 13–19. [[CrossRef](#)]
28. Cruz, M.G.; Hurley, R.J.; Bessell, R.; Sullivan, A.L. Fire behaviour in wheat crops—effect of fuel structure on rate of fire spread. *Int. J. Wildland Fire* **2020**, *29*, 258–271. [[CrossRef](#)]
29. Tulbure, M.G.; Wimberly, M.C.; Roy, D.P.; Henebry, G.M. Spatial and temporal heterogeneity of agricultural fires in the central United States in relation to land cover and land use. *Landsc. Ecol.* **2011**, *26*, 211–224. [[CrossRef](#)]
30. Riley, K.L.; Abatzoglou, J.T.; Grenfell, I.C.; Klene, A.E.; Heinsch, F.A. The relationship of large fire occurrence with drought and fire danger indices in the western USA, 1984–2008: The role of temporal scale. *Int. J. Wildland Fire* **2013**, *22*, 894–909. [[CrossRef](#)]
31. Hanberry, B.B. Recent shifts in shade tolerance and disturbance traits in forests of the eastern United States. *Ecol. Process.* **2019**, *8*, 32. [[CrossRef](#)]
32. Hanberry, B.B.; Bragg, D.C.; Hutchinson, T.F. A reconceptualization of open oak and pine ecosystems of eastern North America using a forest structure spectrum. *Ecosphere* **2018**, *9*, e02431. [[CrossRef](#)]
33. Hanberry, B.B. Reclassifying the wildland—urban interface using fire occurrences for the United States. *Land* **2020**, *9*, 225. [[CrossRef](#)]
34. Vander Yacht, A.L.; Keyser, P.D.; Buehler, D.A.; Harper, C.A.; Buckley, D.S.; Applegate, R.D. Avian occupancy response to oak woodland and savanna restoration. *J. Wildl. Manag.* **2016**, *80*, 1091–1105. [[CrossRef](#)]
35. Pausas, J.G.; Parr, C.L. Towards an understanding of the evolutionary role of fire in animals. *Evol. Ecol.* **2018**, *32*, 113–125. [[CrossRef](#)]

36. Semanova-Nelsen, T.A.; Platt, W.J.; Patterson, T.R.; Huffman, J.; Sikes, B.A. Frequent fire reorganizes fungal communities and slows decomposition across a heterogeneous pine savanna landscape. *New Phytol.* **2019**, *224*, 916–927. [[CrossRef](#)]
37. Grundel, R.; Dulin, G.S.; Pavlovic, N.B. Changes in conservation value from grasslands to savannas to forests: How a temperate canopy cover gradient affects butterfly community composition. *PLoS ONE* **2020**, *15*, e0234139.
38. Roos, C.I.; Bowman, D.M.J.S.; Balch, J.K.; Artaxo, P.; Bond, W.J.; Cochrane, M.; D’Antonio, C.M.; DeFries, R.; Mack, M.; Johnston, F.H.; et al. Pyrogeography, historical ecology, and the human dimensions of fire regimes. *J. Biogeogr.* **2014**, *41*, 833–836. [[CrossRef](#)]
39. Bento-Gonçalves, A.; Vieira, A.; Úbeda, X.; Martin, D. Fire and soils: Key concepts and recent advances. *Geoderma* **2012**, *191*, 3–13. [[CrossRef](#)]
40. Moreira, F.; Ascoli, D.; Safford, H.; Adams, M.A.; Moreno, J.M.; Pereira, J.M.; Catry, F.X.; Armesto, J.; Bond, W.; González, M.E.; et al. Wildfire management in Mediterranean-type regions: Paradigm change needed. *Environ. Res. Lett.* **2020**, *15*, 011001. [[CrossRef](#)]
41. Show, S.B.; Kotok, E.I. *The Role of Fire in the California Pine Forests*; Bull. No. 1294; U.S. Department of Agriculture: Washington, DC, USA, 1924.
42. Ayres, A.; Degolia, A.; Fienup, M.; Kim, Y.; Sainz, J.; Urbisci, L.; Viana, D.; Wesolowski, G.; Plantinga, A.J.; Tague, C. Social science/natural science perspectives on wildfire and climate change. *Geogr. Compass* **2016**, *10*, 67–86. [[CrossRef](#)]
43. Office of Inspector General. *Audit Report: Forest Service Large Fire Suppression Costs*; Rep. 08601-44-SF, 2006. Available online: <http://www.usda.gov/oig/webdocs/08601-44-SF.pdf> (accessed on 18 September 2019).
44. Stein, S.M.; Menakis, J.; Carr, M.A.; Comas, S.J.; Stewart, S.I.; Cleveland, H.; Bramwell, L.; Radeloff, V.C. *Wildfire, Wildlands, and People: Understanding and Preparing for Wildfire in the Wildland-Urban Interface—A Forests on the Edge Report*; U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station: Fort Collins, CO, USA, 2013.



© 2020 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).