



A weekly, continually updated dataset of the probability of large wildfires across western US forests and woodlands

Miranda E. Gray¹, Luke J. Zachmann^{1,2}, and Brett G. Dickson^{1,2}

¹Conservation Science Partners, Inc., Truckee, CA 96161, USA

²Lab of Landscape Ecology and Conservation Biology, Northern Arizona University, Flagstaff, AZ 86011, USA

Correspondence: Miranda E. Gray (miranda@csp-inc.org)

Received: 13 December 2017 – Discussion started: 9 January 2018

Revised: 15 August 2018 – Accepted: 20 August 2018 – Published: 20 September 2018

Abstract. There is broad consensus that wildfire activity is likely to increase in western US forests and woodlands over the next century. Therefore, spatial predictions of the potential for large wildfires have immediate and growing relevance to near- and long-term research, planning, and management objectives. Fuels, climate, weather, and the landscape all exert controls on wildfire occurrence and spread, but the dynamics of these controls vary from daily to decadal timescales. Accurate spatial predictions of large wildfires should therefore strive to integrate across these variables and timescales. Here, we describe a high spatial resolution dataset (250 m pixel) of the probability of large wildfires (> 405 ha) across forests and woodlands in the contiguous western US, from 2005 to the present. The dataset is automatically updated on a weekly basis using Google Earth Engine and a “continuous integration” pipeline. Each image in the dataset is the output of a random forest machine-learning algorithm, trained on random samples of historic small and large wildfires and represents the predicted conditional probability of an individual pixel burning in a large fire, given an ignition or fire spread to that pixel. This novel workflow is able to integrate the near-term dynamics of fuels and weather into weekly predictions while also integrating longer-term dynamics of fuels, the climate, and the landscape. As a continually updated product, the dataset can provide operational fire managers with contemporary, on-the-ground information to closely monitor the changing potential for large wildfire occurrence and spread. It can also serve as a foundational dataset for longer-term planning and research, such as the strategic targeting of fuels management, fire-smart development at the wildland–urban interface, and the analysis of trends in wildfire potential over time. Weekly large fire probability GeoTiff products from 2005 to 2017 are archived on the Figshare online digital repository with the DOI <https://doi.org/10.6084/m9.figshare.5765967> (available at <https://doi.org/10.6084/m9.figshare.5765967.v1>). Weekly GeoTiff products and the entire dataset from 2005 onwards are also continually uploaded to a Google Cloud Storage bucket at <https://console.cloud.google.com/storage/wffr-preds/V1> (last access: 14 September 2018) and are available free of charge with a Google account. Continually updated products and the long-term archive are also available to registered Google Earth Engine (GEE) users as public GEE assets and can be accessed with the image collection ID “users/mgray/wffr-preds” within GEE.

1 Introduction

Wildfire predictions for near-term operations versus long-term planning and research operate at different spatiotemporal scales, aiming either to understand the risk posed over the course of an individual fire or fire season or to understand the broadscale characteristics of fire regimes. For example, operational needs emphasize contemporary, on-the-ground conditions (Brillinger et al., 2003; Martell et al., 1989; Sullivan, 2009a, b) and largely ignore the longer-term controls on fire (e.g., occurring years to decades prior to a fire). By contrast, predictions across longer time frames and often larger spatial scales will omit the contemporary weather patterns that drive fire occurrence (Krawchuk and Moritz, 2014; Littell et al., 2009; Urbietta et al., 2015). While many models and datasets exist to support these needs, they also reflect different and non-overlapping scales. We sought to fill this gap by developing a dataset of the predicted conditional probability that an area on the landscape will burn in a large wildfire (i.e., > 405 ha) given an ignition or fire spread to that area, which integrates across spatiotemporal scales in an empirical framework. We developed the dataset at a high spatial resolution (250 m pixel) and moderate temporal resolution (updated weekly) across forests and woodlands in the contiguous western US. The resulting dataset is intended to meet multiple objectives of local to national research, management, and planning efforts.

The dataset that we describe in this paper is continually updated with near-term information, which we define as occurring over a period of days to months prior to and during a fire. A well-developed approach to similarly incorporate the dynamic near-term drivers of wildfires is to simulate the spread of individual fires over a landscape (Finney, 2004; Sullivan, 2009c; Tymstra et al., 2010). Modeling systems that perform these simulations, such as Farsite (Finney, 2004) and FSPro (Finney et al., 2011b), are used widely during wildfire incidents and in real time to understand the potential spread and behavior of burning fires. These tools can provide critical information for individual or localized fire probability in real time but are limited in their ability to elucidate regional and cross-regional fire risk at similar time frames and are dependent on fuels data, e.g., from the LANDFIRE project (Rollins, 2009), which are often not updated for years at a time. Although the work described herein does not attempt to model the risk posed by individual fires, it is meant to provide contemporary fire information across regional extents, drawing on continually updated fuel and weather data to predict conditional large fire probability at a high resolution. Therefore, it provides a needed, complementary dataset to existing models that operate on near-term timescales.

By simulating individual fires across time and space, the fire modeling systems described above can also scale up to predict the long-term, multiyear potential of fires at every point on a landscape (Finney et al., 2011a; Parisien et al., 2005). This approach is commonly used for the longer-term

planning of fuel treatments and other fire risk planning and assessments (Haas et al., 2013; Thompson et al., 2017). However, these landscape-scale simulations can be user and computationally intensive (Parisien et al., 2012a; Varner et al., 2009), constraining the ability of analysts and planners to update datasets at both broad spatial scales and decision-relevant timescales. For example, regional or national predictive datasets may need to be updated according to changes in fuel that occur within a fire season and on an interannual basis.

Alternative methods to predict fire occurrence relate empirical fire data to environmental predictors in statistical models (Gray et al., 2014; Preisler et al., 2016; Stavros et al., 2014). Data availability in this case, namely the spatiotemporal alignment of accurate and high-resolution fire, weather, and fuels data, also acts as a constraint on either the spatial or temporal scale of analysis (Taylor et al., 2013). However, such statistical methods are common in predicting fire occurrence on a macroscale because they can draw on coarse-scale data to overcome this constraint (Krawchuk et al., 2009; Moritz et al., 2012; Parisien et al., 2012b). Owing to the flexibility of model specification and data inputs as well as increasingly accurate and high-resolution observational data, statistically based empirical models can integrate both the contemporary, near-term drivers as well as the long-term controls on fire potential.

Indeed, recent studies have explicitly compared the role of the temporal scale in predicting fire occurrence and have shown that long-term normals and variability in climate and vegetation provide complementary predictive power (Abatzoglou and Kolden, 2011, 2013; Parisien et al., 2014; Riley et al., 2013). For example, the long-term climate exerts an influence on the flammability (e.g., due to biomass production, vegetation composition, and average fuel moisture) of a fuel bed, but weekly and sub-weekly weather will moderate fuel moisture in a site-specific way. Similarly, relatively recent disturbance events such as previous burns can regulate biomass production and the subsequent fire risk on interannual timescales (Parisien et al., 2014; Parks et al., 2015). It follows that predictive datasets of wildfire potential should strive to integrate across complex, dynamic interactions at near- and long-term timescales. Here, we describe a time series of the conditional probability of a large fire, continually updated on a weekly basis (with a 1 week lag) to integrate the near-term controls on fire occurrence, which also considers the longer-term influences of land use, disturbance, the climate, and topography. The complete dataset (2005–present) can also be considered a foundational dataset for understanding the long-term, probabilistic exposure of forests and woodlands to large fires.

2 Methods

2.1 Modeling

We modeled the conditional probability of large fire occurrence, which we define as the probability that an area on the landscape will burn in a large (i.e., > 405 ha) fire, conditional on either an ignition event or fire spreading to that area. While defining large fire size is somewhat arbitrary, 405 ha is commonly used to distinguish large from small fires in western US forests (e.g., Westerling, 2006), and fires > 405 ha accounted for approximately 95 % of the area burned in western forests and woodlands from 1992 to 2015 (Short, 2017). Additionally, our method focused only on the probability of a large fire, irrespective of ignition likelihood or sources. Ignitions are non-random events that adhere to spatial patterns tied to anthropogenic or lightning activity, which are not accounted for in this dataset.

We used a random forest (RF) classification algorithm (Breiman, 2001) to train predictive models of large fire probability. RF is a machine-learning technique that recursively partitions variables to classify an outcome of interest, in this case small or large fire events. Multiple classification trees are fit to bootstrapped samples of the training data, but at each node, only a fraction of randomly selected predictors are available for the binary partitioning. The randomized process of recursive partitioning uncovers hidden structures in the data without overfitting and yields strong predictive models (Prasad et al., 2006). This makes RF an ideal method to predict fire occurrence across broad and diverse ecoregions, where high dimensionality is needed to account for unforeseen interactions between the climate, fuels, and the landscape (Cutler et al., 2007).

The binary response variable in our RF models was a point on the landscape where there was an ignition event that resulted in a small fire (i.e., < 405 ha; “0” response) or that historically burned in a large fire (i.e., > 405 ha; “1” response). Therefore, model outputs (i.e., raster maps) can be interpreted as reflecting the probability that a given area on the landscape will burn in a large fire, conditional on either an ignition or spread of fire to that area. We sampled large fire points from the Moderate Resolution Imaging Spectroradiometer (MODIS) burned area (BA) dataset (MCD45A1 v6; Roy et al., 2008), which is a 500 m remote sensing product that contains the day of burn. We sampled small fire points from a database of reported fires in the United States (Short, 2014, 2017) that contains the day of discovery (Sect. 2.2). To avoid spatial autocorrelation within large fires, we drew at most 1 sample (a point location) within each large fire (see Sect. 2.2). We then matched these large fire samples with an equally sized random sample of small fires (see Sect. 2.3) to build a single RF model across the western US.

While spatial autocorrelation is invariably present within individual fires, burning conditions can also be quite heterogeneous over the course of a single large fire (Turner, 2010).

Therefore, we took a step further in capturing this heterogeneity. We repeated the above sampling and model building protocol using 10 different random samples of large and small fires, such that each of 10 RF models was not entirely independent but contributed slightly novel information to a mean prediction across those 10 models. This type of ensemble modeling provides a means of producing models that are more accurate than the individual models that make them up, while depicting the variance across predictions, which is critical for risk assessment (Dietterich, 2000; Palmer et al., 2005).

Using 10 trained RF models, we created spatial predictions of the mean and standard deviation of large fire probability at 250 m resolution across western US forests and woodlands. Daily spatial predictions were created at weekly intervals from 2005 through the present. See Sect. 4 below that describes the process by which new predictor data acquisitions are automatically and continually integrated into weekly predictions and uploaded to the cloud. Models were trained and spatial predictions created within Google Earth Engine (GEE; Gorelick et al., 2017), which is a cloud-based platform that makes terabyte-scale analysis available on an extensive catalog of satellite imagery and geospatial datasets.

2.2 Response variables

We sampled large fires by retaining MODIS BA pixels that were within 8 days of the reported burn date of neighboring burned pixels. This boosted our confidence in the likelihood that connected pixels were part of the same fire (Archibald and Roy, 2009), which we also required to be connected to ≥ 15 other pixels ($\cong 405$ ha). We then used the Monitoring Trends in Burn Severity (MTBS; Eidenshink et al., 2007) dataset to delineate the perimeters of annual large wildfires (excluding prescribed fires) and sampled daily MODIS burned area pixels in a given year from within these perimeters. We masked burned areas according to forest or woodland land cover types classified in the 2001 US National Land Cover Dataset (NLCD, 30 m resolution; Homer et al., 2007) before drawing 10 random samples across all large fires ($n \cong 900$ in each sample) from 2005 to 2014. Each individual large fire sample was taken as the centroid of a 500 m pixel (Fig. 2). We used the 2001 NLCD product because it represents the closest complete land cover prior to the fires selected for training data in this analysis.

We drew random samples of small wildfires from the US Fire Occurrence Dataset (FOD; Short, 2014, 2017), masked by NLCD forest and woodland cover. We did not draw small samples from the BA dataset because the estimated minimum detectable burn size is approximately 120 ha, which means that smaller fires are grossly underestimated (Giglio et al., 2009; Roy and Boschetti, 2009). Within each Environmental Protection Agency (EPA) level III ecoregion in the contiguous western US (Fig. 2), we paired an equally sized random sample of small fires with each of the 10 large fire samples,

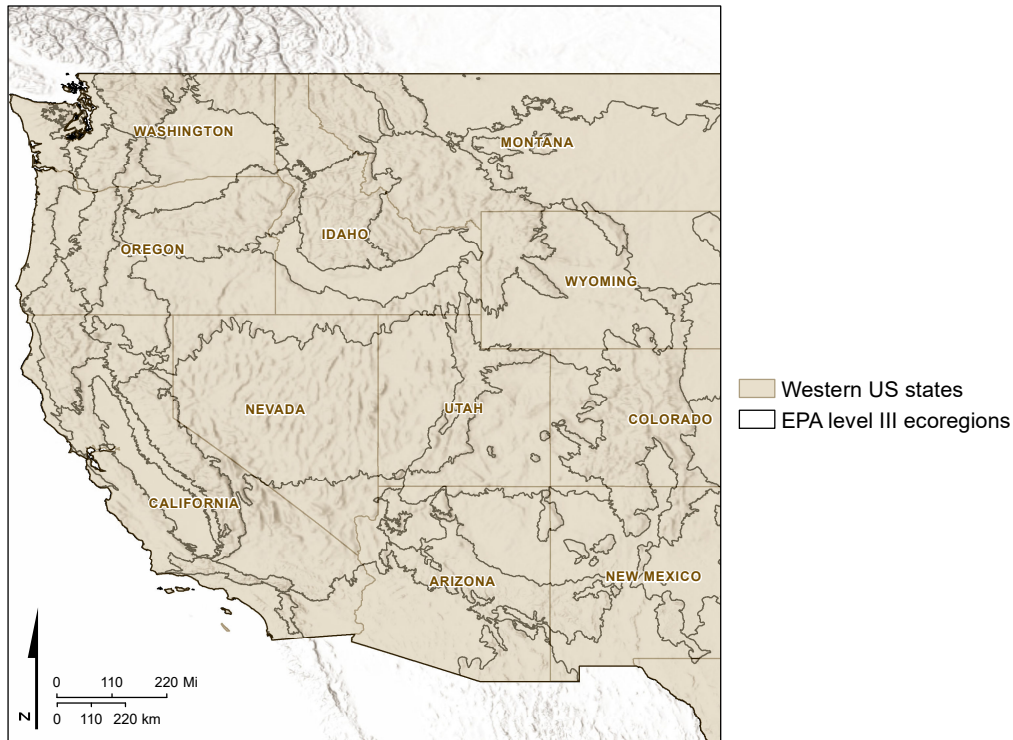


Figure 1. The dataset described in this paper predicts conditional large fire probability across forests and woodlands in the 11 contiguous western US states. Environmental Protection Agency (EPA) level III ecoregions were used to stratify sampling and create a spatially balanced 1 : 1 sample of small and large fires across diverse ecoregions, which was then used to train a single random forest model across the western US to predict large fire probability.

resulting in spatially balanced, 1 : 1 training datasets across diverse ecoregions. Although there are ecoregional differences in the individual drivers of large wildfires (e.g., Barbero et al., 2014), we used the spatially balanced response data and a myriad of predictor data (see below) to develop an RF model that covered all ecoregions. RF was an ideal method in this case because high dimensionality in the predictor data accounts for unforeseen interactions between the climate, fuels, and the landscape (Cutler et al., 2007), which likely drive ecoregional differences in fire response.

2.3 Predictor variables

We derived predictor variables that describe the land surface and climate over multiyear, long-term time frames. Similarly, we derived predictor variables that describe the land surface and weather over weekly, near-term time frames (Table 1). Specifically, an individual large or small fire sample was spatially related to long-term predictors derived over a multiyear period and near-term predictors derived over the week before and after a fire occurrence. The integration of predictors in this way resolves the dynamic probability of a large fire into long-term drivers of fire and near-term land surface and ambient conditions directly leading up to and following a fire event. To account for the difference in spatial scales between

a large fire and the native resolution of spatial predictors (i.e., ranging from 30 m to 4 km), we used a moving window to summarize predictors within a circular kernel with a radius of 1135 m. Predictor variables that were not in a native 250 m resolution were resampled using bilinear interpolation.

2.3.1 Long-term land-surface variables

To characterize long-term live fuel availability and water content per pixel, we used the enhanced vegetation index (EVI, 250 m resolution) from the MODIS MOD13Q1 v006 product (Didan, 2015) and the normalized difference water index (NDWI, 500 m resolution) derived from the MODIS MCD43A4 v006 product (Schaaf, 2015). MODIS EVI and the normalized difference vegetation index (NDVI) both provide proxies for total vegetation, but the EVI is more sensitive to canopy variations in densely vegetated areas (Huete et al., 2002). We used a multiyear time series of the EVI not only to capture the variability in overall biomass production across the western US, but also as a basis to capture variability in sub-pixel vegetation dynamics (e.g., Helman et al., 2015). We also included the EVI to capture longer-term changes in fuel abundance due to prior burns, based on findings that forested ecoregions have shown large to moderate

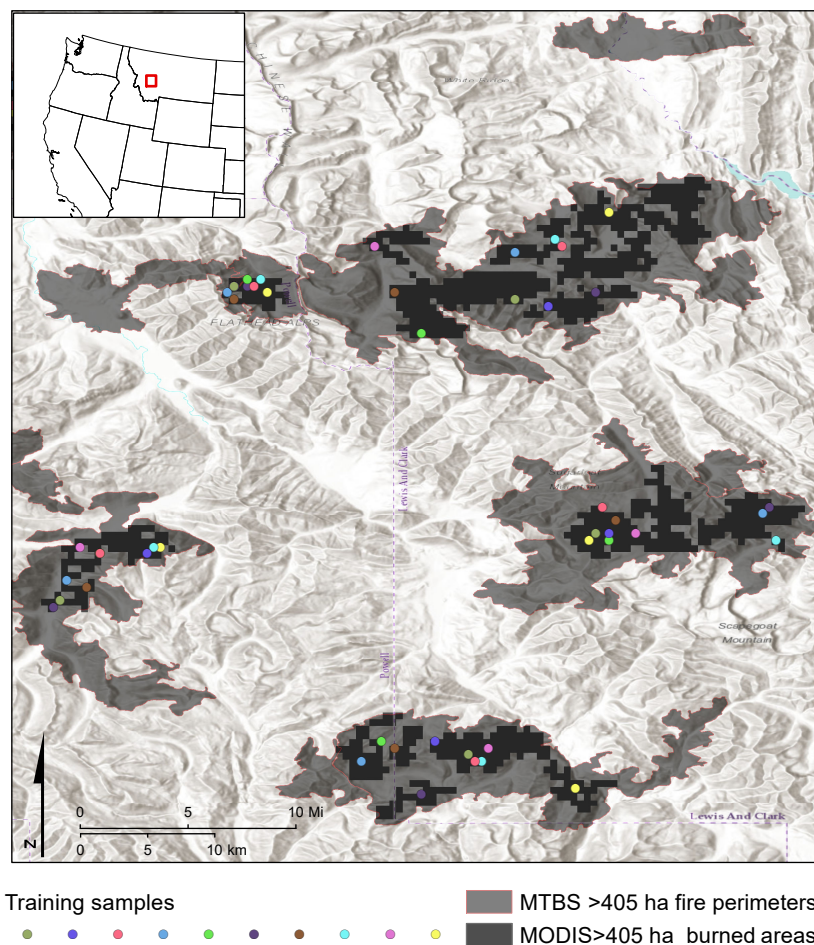


Figure 2. Example of how the Moderate Resolution Imaging Spectroradiometer (MODIS) burned area (BA) dataset was used to draw 10 random samples from within large fires. Each sample, taken across all large fires in 2005–2014, was used to train a random forest model to predict large fire probability. Fire perimeters from the Monitoring Trends in Burn Severity (MTBS) dataset are included because they were used to restrict BA sampling within individual wildfires (excluding prescribed fires).

post-fire reductions in MODIS NDVI over a 10-year period (Yang et al., 2017).

The NDWI was originally proposed as a complementary vegetation index to the NDVI and EVI to detect vegetation liquid water content (Gao, 1996), and has since been shown to relate strongly to the total water content per pixel (Cheng et al., 2006; Maki et al., 2004). Similar to the EVI, we included a multiyear time series of the NDWI to capture moisture gradients across space. The NDWI has also been successful in estimating vegetation moisture and fire hazard when coupled with an estimate of the total vegetation. Thus, the interaction between the EVI and NDWI may provide important information about pixel-wise fuel moisture (Maki et al., 2004).

Each of the NDWI and EVI products used in our analysis were 16-day composites computed from atmospherically corrected, bidirectional daily surface reflectance. MOD13Q1 contains pixel quality information and MCD43A4 contains

pixel and band quality information. For both products, we only retained observations that were free of ice and snow and that fell between the pixel-wise median date of the onset of greenness and the median date of the onset of senescence, determined from the MODIS Global Vegetation Phenology product (MCD12Q2 v005). We took the median greenness and senescence days of year from 2001 to 2004, corresponding to the beginning of MCD12Q2 availability to the start of our fire samples. In general, limiting observations to the growing season is more appropriate for land cover mapping (Hansen et al., 2013). We extracted 5 percentile values (10, 25, 50, 75 and 90 %) of the EVI and NDWI as well as the slope of linear regression of the EVI and NDWI versus image date from 2000 (the year MODIS was deployed) to the approximate date of each fire occurrence. These values provided at least 5 complete years of the observed EVI and NDWI prior to the occurrence of a given fire. We included these metrics to build a generic feature space to characterize

Table 1. Spatially explicit climate and land-surface predictors of conditional large fire probability, including the data source, spatial resolution, and description of how variables were derived from the source data. Grouping of predictor variables indicates whether they are derived over the near term (months or weeks preceding fire occurrence) or long term (multiyear).

Predictor	Source	Resolution	Description
(1) Long-term climate variables			
Annual precipitation, temperature seasonality, precipitation of the warmest month, mean temperature of the wettest month, mean temperature of the warmest month	PRISM	800 m	Derived from monthly normals from 1981 to 2010
(2) Long-term land-surface variables			
EVI	MODIS	250 m	10th, 25th, 50th, 75th, and 90th percentiles, and slope of linear regression with image date, from 2000 to the date of fire occurrence
NDWI	MODIS	500 m	
Human modification, distance to urban development	CSP 2016	30 m	Index or distance value at 2001 for fires pre 2011, and 2011 for fires post 2011
Elevation, slope, aspect, topographic roughness	USGS	30 m	
(3) Near-term land-surface variables			
EVI	MODIS	250 m	Index or temperature value immediately preceding fire occurrence
NDWI	MODIS	500 m	
LST	MODIS	1 km	
(4) Near-term weather variables			
100 h fuel moisture, 1000 h fuel moisture, burning index, energy release component, precipitation, temperature, relative humidity, specific humidity, potential evapotranspiration, solar radiation, wind speed, wind direction, PDSI	GRIDMET	4 km	Mean values in the 2 weeks surrounding fire occurrence

vegetation over at least 5 complete years, as they have been used in previous machine-learning applications to characterize regional-scale forest cover (Hansen et al., 2013).

To characterize the land surface as modified by humans over the long-term, we included indices of human modification for the years 2001 and 2011 (Conservation Science Partners Inc., 2016; 30 m resolution). This index quantifies the cumulative degree of modification of natural lands attributable directly to energy, residential, commercial, transportation, and agricultural development. Since they are less natural and generally more fragmented, we hypothesized that more developed landscapes are less likely to burn in large fires. We also used the associated residential and commercial development dataset (Conservation Science Partners Inc., 2016; 30 m resolution) to compute the Euclidean distance to urban development in 2001 and 2011. Urban development in this case was approximated by a “moderate” value of residential and commercial development, which is roughly equivalent to the “built-up moderate” class in the NLCD, except that it removes the exaggerated effects of roads. We assumed that suppression resources and mandates are more readily accessed closer to urban centers and thus constrains the likelihood of large fires. Lastly, we used the Shuttle Radar To-

pography Mission digital elevation data (Farr et al., 2007) to characterize topographic variables, namely, elevation, slope, aspect, and terrain roughness (standard deviation of elevation), each at a 30 m resolution.

2.3.2 Long-term climate variables

We incorporated predictors computed from monthly climatological normals of temperature and precipitation for the period 1981–2010, as derived from the Parameter-elevation Regressions on Independent Slopes Model (PRISM Norm81m vM2; 800 m resolution; Daly et al., 1994). We selected 5 metrics which summarized the long-term annual means, extremes, and seasonality of temperature and precipitation and have been used previously to capture the amount and dryness of biomass to predict fire occurrence (Krawchuk et al., 2009; Moritz et al., 2012). These metrics included the annual precipitation, precipitation of the warmest month, mean temperature of the wettest month, mean temperature of the warmest month, and temperature seasonality (i.e., the standard deviation of mean monthly temperatures; O’Donnell and Ignizio, 2012).

2.3.3 Near-term land-surface variables

We characterized the short-term live vegetation abundance and condition as well as pixel water content with the single EVI and NDWI observations in the month prior to fire occurrence. These near-term indices are meant to capture the vegetation abundance and condition immediately prior to burning. For instance, when coupled with the EVI, the NDWI has been shown to contribute to fire risk on sub-monthly timescales (Maki et al., 2004).

We used the MODIS MOD11A2 daytime Land Surface Temperature (LST) 8-day composites (1 km resolution; NASA LP DAAC, 2015), which represent average values of clear-sky LSTs, to similarly characterize the ground temperature immediately leading up to a fire occurrence. Due to feedback between LST and near-surface humidity, remotely sensed LST has been used to predict the vapor pressure deficit, which in itself is a good short-term predictor of fine dead fuel moisture and fire danger (Boer et al., 2017; Nolan et al., 2016). We included the value of LST from the 8 days prior to the fire.

2.3.4 Near-term weather variables

The standard meteorological variables known to influence the daily fire and fuel environment were taken from the GRIDMET gridded daily surface meteorological dataset (4 km resolution; Abatzoglou, 2013). We incorporated the total precipitation, mean minimum and maximum temperatures, mean minimum and maximum relative humidity, mean wind speed and direction and the mean Palmer drought severity index (PDSI) for the 2 weeks surrounding fire occurrence.

The standard weather variables have also been compiled into indices that more directly address the processes by which they effect fires and fuels, including the energy release component (ERC), the burning index (BI), and 100 and 1000 h dead fuel moisture (FM100 and FM1000). These indices are components of the US National Fire Danger Rating System (NFDRS) and are derived from models built on the combustion physics and moisture dynamics of the fuel environment, assuming a consistent fuel model “G” typified by short needle pine and heavy dead loads (Abatzoglou, 2013; Schlobohm and Brain, 2002). The FM100 and FM1000 indices represent the modeled moisture content of large dead fuels in the 2.5 to 7.6 cm diameter class and the 7.6 to 20.3 cm diameter class, respectively. ERC is a cumulative fuel moisture index reflecting the contribution of all live and dead fuel moisture on the potential heat release and is also an input into the BI, which additionally incorporates the potential rate of fire spread. GRIDMET assumes that the persistent fuel environment includes all size classes of dead fuels as well as herbaceous and woody live fuels, all contributing to the derived values of these indices. We incorporated the mean values of ERC, BI, FM100, and FM1000 in the 2 weeks surrounding fire occurrence.

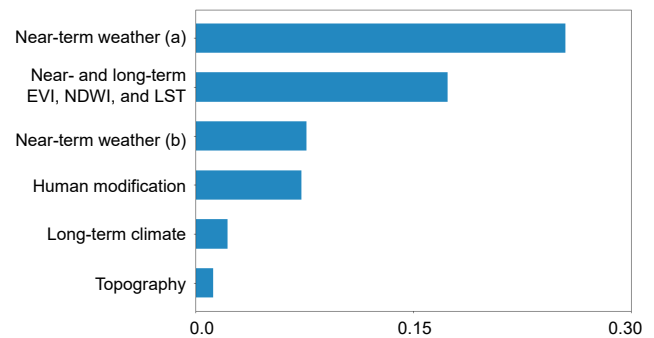


Figure 3. Graph of relative variable importance, based on the permutation importance measure, which directly measures importance by observing the effect on model accuracy by randomly permuting the values of each predictor variable. Because random forest “spreads” variable importance across collinear variables, non-independent variables were grouped together to determine their collective importance (see Table 1 for details on the variable groups). Near-term weather variables (a) include the energy release component, burning index, 100 and 1000 h fuel moisture, relative humidity, and precipitation. Near-term weather variables (b) include temperature, vapor pressure deficit, specific humidity, solar radiation, wind speed, and wind direction.

3 Dataset evaluation

Using all training data from 2005 to 2014 (i.e., no independent testing data), we compared models in R using the “caret” package (Kuhn, 2008), and extracted variable importance using the “rfpimp” package in Python (available at <https://github.com/parrt/random-forest-importances>, last access: 14 September 2018). We ranked predictor variable importance based on the permutation importance, which directly measures importance by observing the effect on model accuracy by randomly permuting the values of each predictor variable (Cutler et al., 2007). Since RF “spreads” variable importance across collinear variables (Cutler et al., 2007), we used a built-in function in the “rfpimp” package to permute collinear variables together and determine their relative and collective importance (Fig. 3). Across the 10 models, overall accuracy was consistently between 0.77 and 0.79 and area under the receiver operating curve (AUC) was consistently 0.83–0.86. Out of 46 total predictor variables, the most important variables were near-term weather variables that included the ERC, BI, FM100, FM1000, relative humidity, and precipitation, as well as the collective near- and long-term EVI, NDWI, and LST variables (Fig. 3).

To independently evaluate the model on data from 2015 to 2016, we used the MODIS BA and FOD datasets to draw a testing sample from within all large fires and an equally sized random sample of small fires (response value of “0” and “1”, respectively; $n \cong 400$ large fires). Again, large samples were taken as the centroid of 500 m pixels. Using weekly predictions (i.e., raster maps; Fig. 5) of large fire probability

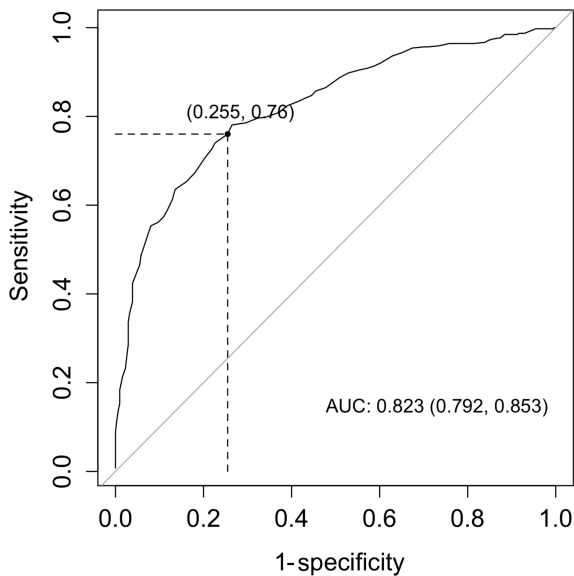


Figure 4. Receiver operating curve (ROC) for an independent testing dataset of small and large fires that occurred from 2015 to 2016. Sensitivity and (1-specificity) values are shown for the point where large fire probability values > 0.47 are classified as a large fire and values < 0.47 are classified as a small fire, since this value was found to simultaneously maximize sensitivity and specificity.

in 2015 and 2016, we extracted the predicted values at the time (i.e., the closest prediction in time prior to fire occurrence) and location of individual testing points. We used the R package “OptimalCutpoints” (López-Ratón et al., 2014) to determine an optimal cutoff between 0 and 1 that simultaneously maximized the sensitivity (true positive rate) and specificity (true negative rate) of predictions. In this case, using a probability cutoff of 0.47 to predict binary large (> 0.47) versus small (< 0.47) fires resulted in the greatest rate of true positives and negatives in our testing datasets. Based on an optimal cutoff of 0.47 and 2 years of independent data, the sensitivity of the dataset was 0.76, the specificity was 0.75, and the area under the receiver operating curve (ROC) curve was 0.82 (Fig. 4). We took another step to visualize model performance by mapping the rate of false positives and false negatives (i.e., the number of false positives or false negatives normalized by the number of testing samples) within each EPA level III ecoregion to examine any obvious biases in under or overprediction across ecoregions (Fig. 6). There was more of a tendency for the model to overpredict large fires in some of the drier ecoregions, such as the Colorado Plateau and the Central Basin and Range, and the inverse was true in some of the wetter ecoregions. In particular, the Cascades and Southern Rockies tended to underpredict large fires rather than overpredict (Fig. 6).

4 Continuous integration

We developed a continuous integration (CI) “pipeline” to generate new predictions as soon as the dynamic predictors upon which the model is conditioned become available in GEE. The refresh rate of each predictor varies based on the data sources. For example, GRIDMET assets are updated approximately every 2 days, whereas the MODIS products are updated approximately every 8 days. The pipeline, which tests for the availability of predictors against the requirements of the model, runs on a schedule, compiling each morning at 04:00 Pacific standard time. If all of the criteria are met, a new prediction is generated and appended to the existing collection. We used GitLab.com because GitLab offers continuous integration (CI) services at no cost. The builds are executed using a custom Docker image, which is a bare-bones Ubuntu image configured with the Google Earth Engine Python application program interface (API) client library and its dependencies.

5 Band descriptions

Each image in the dataset contains the following bands:

- Band 1 (“mean”) represents the mean probability of large fires across 10 trained models. Values range from 0 to 1.
- Band 2 (“stdDev”) represents the standard deviation of the probability of large fires across 10 trained models.
- Band 3 (“modis_QA”) indicates if one of the near-term predictors (i.e., MOD13Q1, MCD43A4, or MOD11A2 immediately preceding the prediction date) had unreliable quality. If this band value is equal to 0, all near-term MODIS pixels were processed and of good quality. If this band value is equal to 1, at least one near-term MODIS pixel was not processed or was of bad quality (note: for the MODIS products described above, only good quality pixels were retained for model training, but all pixels were retained when creating spatial predictions).

6 Code and data availability

Weekly large fire probability GeoTiff products from 2005 to 2017 are archived on the Figshare online digital repository with the DOI <https://doi.org/10.6084/m9.figshare.5765967> (Gray et al., 2018; available at <https://doi.org/10.6084/m9.figshare.5765967.v1>). GeoTiff products and the entire dataset from 2005 onwards are also continually uploaded to a Google Cloud Storage bucket at <https://console.cloud.google.com/storage/wffr-preds/V1> (last access: 14 September 2018) and are available free of charge with a Google account. Continually updated products and the long-term archive are also available to registered

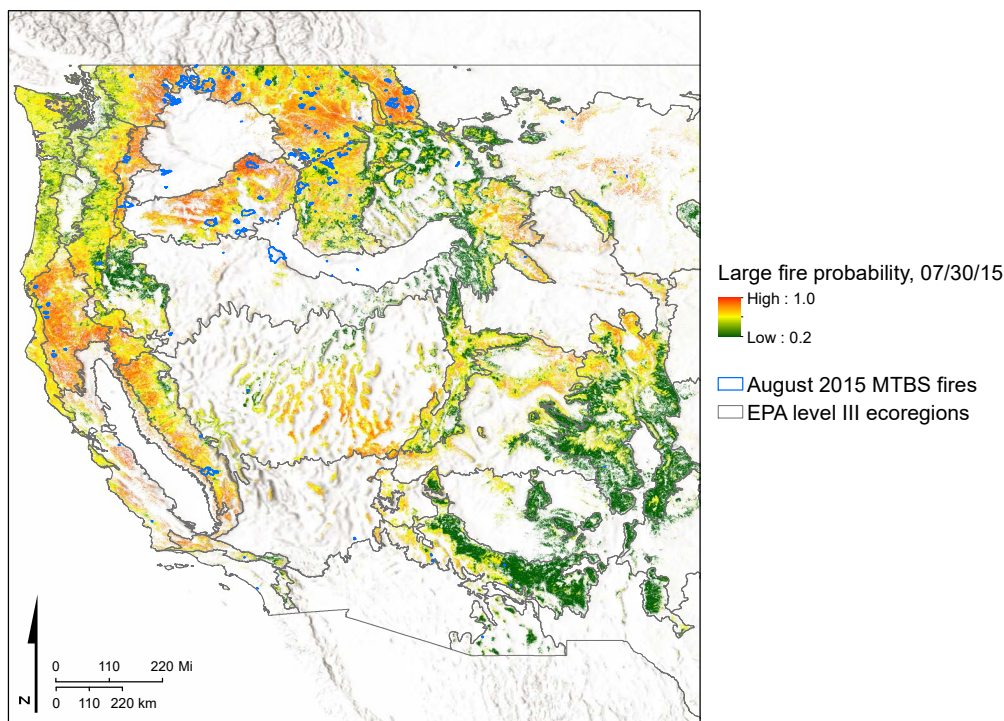


Figure 5. Predicted conditional large fire probability for the week of 30 July 2015. MTBS fires greater than 405 ha that started in August 2015 are overlaid on the map. White (non-colored) areas are non-forested.

GEE users as public GEE assets and can be accessed with the image collection ID “users/mgray/wffr-preds” within GEE. All source code is available at a GitLab repository (<https://gitlab.com/wffr>, last access: 14 September 2018; only accessible after free registration on GitLab).

7 Conclusions

The dataset we describe here of weekly predictions of the probability of large forest or woodland fires across the western US invokes interacting effects over multiple timescales that contribute to a site’s dynamic fire potential. By drawing on weather, climate, and land-surface dynamics at multiple timescales to predict individual fire occurrence at a high spatial and temporal resolution, this dataset fills a gap in existing datasets. The result is relevant to research, planning, and management objectives that span across the western US, ranging from short-term outlooks to long-term planning.

More strategic planning for fuels management is critically needed to adapt to an inevitable increase in wildfires in the western US in the coming decades (Schoennagel et al., 2017). For instance, fuels treatments as currently implemented are limited in their ability to mitigate the broad-scale effects of wildfires, because it is relatively rare that treatments actually encounter wildfires (Barnett et al., 2016). Strategically targeting areas for treatment based on large wildfire potential, coupled with estimates of burn severity,

will lead to more cost and ecologically effective decisions (Scott et al., 2016; Thompson et al., 2017). However, modeling systems currently used for this purpose are often computationally and user-intensive, constraining the ability to update results at both broad spatial scales and timescales concurrent with the changing fire environment. For example, the Wildland Fire Potential dataset is available for the entire US at 270 m resolution and describes the static fire potential as of 2007, 2012, and 2014 (Dillon et al., 2015). The dataset we describe here is automatically updated weekly (as reflected in fuel abundance and condition and fire weather) and annually (as reflected in the NDWI and EVI) to match higher-frequency dynamics of the fuel and fire environment, which change on these timescales and critically effect fuels-management decisions.

Another area where probabilistic fire exposure analysis can help with strategic fuels and fire planning is at the wildland–urban interface (WUI; e.g., Haas et al., 2013). WUI lands in the western US have expanded dramatically over the past few decades, and roughly 40% of these lands are predicted to experience moderate to large increases in the probability of wildfires in the next 20 years (Schoennagel et al., 2017). Considering also that a large percentage of potential WUI lands are still undeveloped, strategic planning for both fuels management and infrastructure development can make communities more resilient to wildfires. This dataset can help guide development plans on multiple scales (e.g.,

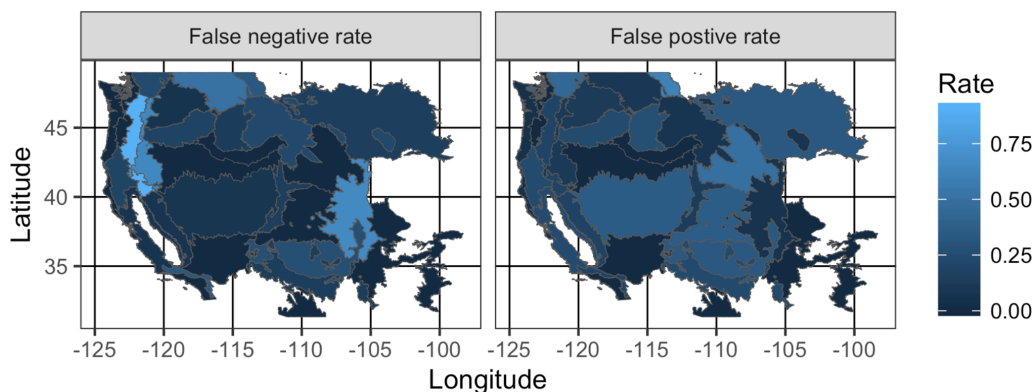


Figure 6. False positive (FP) and false negative (FN) rates of an independent testing dataset of small and large fires from 2015 to 2016, mapped across Environmental Protection Agency (EPA) level III ecoregions. No testing data were available for those ecoregions that are not displayed.

city, county, or state), drawing on a rich time series that gives analysts and planners access to the observed trends, means, and extremes of the potential for large wildfires over time. For example, planners may be interested in assessing the risk of new development within the WUI, recognizing that new development would potentially introduce more sources of ignition throughout the year. Therefore, planners might seek to understand interannual patterns in the timing and magnitude of the conditional probability of large fires, given an increase in the number of ignition sources.

In contrast to longer-term predictions, contemporary predictions of large fire potential provide operational fire managers with immediate, on-the-ground information to closely monitor how changing conditions affect active or impending fires and the likelihood that fire suppression will require outside resources. In the US, contemporary predictions are widely used during the peak fire season (Owen et al., 2012). Available products through the US Predictive Services program (<http://psgeodata.fs.fed.us/>, last access: 14 September 2018) and the Wildland Fire Assessment System (www.wfas.net, last access: 14 September 2018; Preisler et al., 2016) consider fuel and weather conditions that change on daily to weekly timescales while ignoring the longer-term climate and fuel variability that moderate a site's current fire potential. Modeling systems that perform simulations of fires as they are occurring, such as FARSITE and FSPRO, provide critical information for individual or localized fire probability and behavior but are limited in their ability to elucidate contemporary regional and cross-regional fire risk and are additionally dependent on fuels data (e.g., from LANDFIRE) that are not updated to the present. The dataset described here provides continually updated predictions across the western US while simultaneously accounting for dynamic fuel and landscape compositions that are shaped over the near and long term. Thus, the dataset is a needed addition to operational products of contemporary fire potential.

As the observational record grows longer to include more temporal variability and new normals, we can continue to re-train models on the same basis of predictors and update and evaluate this dataset. This will allow for any non-stationary relationships between wildfires, the climate, fuels, and the landscape to be easily integrated into predictions. For example, if underlying relationships such as the precipitation of the wettest month or average early May EVI change in the future, models would simply need to be retrained on updated datasets to integrate such non-stationarities. In future development, forecasted climate, weather, and fuels data may also be integrated into the analysis in order to create predictions of large fire probability into the future.

Author contributions. MEG developed the fire models and data products with critical contributions from both LJZ and BGD. LJZ developed the continuous integration algorithm. All authors contributed to the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This research was supported by the Wilburforce Foundation.

Edited by: David Carlson

Reviewed by: two anonymous referees

References

- Abatzoglou, J. T.: Development of gridded surface meteorological data for ecological applications and modelling, *Int. J. Climatol.*, 33, 121–131, <https://doi.org/10.1002/joc.3413>, 2013.
- Abatzoglou, J. T. and Kolden, C. A.: Relative importance of weather and climate on wildfire growth in interior Alaska, *Int. J. Wildland Fire*, 20, 479–486, <https://doi.org/10.1071/WF10046>, 2011.
- Abatzoglou, J. T. and Kolden, C. A.: Relationships between climate and macroscale area burned in the western United States, *Int. J. Wildland Fire*, 22, 1003–1020, 2013.
- Archibald, S. and Roy, D. P.: Identifying Individual Fires From Satellite-Derived Burned Area Data, in *International Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, Cape Town, South Africa, 160–163, 2009.
- Barbero, R., Abatzoglou, J. T., Steel, E. A., and Larkin, N. K.: Modeling very large–fire occurrences over the continental United States from weather and climate forcing, *Environ. Res. Lett.*, 9, 124009, <https://doi.org/10.1088/1748-9326/9/12/124009>, 2014.
- Barnett, K., Parks, S. A., Miller, C., and Naughton, H. T.: Beyond fuel treatment effectiveness: Characterizing interactions between fire and treatments in the US, *Forests*, 7, 237, <https://doi.org/10.3390/f7100237>, 2016.
- Boer, M. M., Nolan, R. H., Resco De Dios, V., Clarke, H., Price, O. F., and Bradstock, R. A.: Changing Weather Extremes Call for Early Warning of Potential for Catastrophic Fire, *Earth's Future*, 5, 1196–1202, <https://doi.org/10.1002/2017EF000657>, 2017.
- Breiman, L.: Random forests, *Mach. Learn.*, 45, 5–32, <https://doi.org/10.1023/A:1010933404324>, 2001.
- Brillinger, D. R., Preisler, H. K., and Benoit, J. W.: Risk assessment: a forest fire example, in: *Science and statistics: a festschrift for Terry Speed*, edited by: Goldstein, D., Institute of Mathematical Statistics, Beachwood, 177–196, 2003.
- Cheng, Y. B., Zarco-Tejada, P. J., Riaño, D., Rueda, C. A., and Ustin, S. L.: Estimating vegetation water content with hyperspectral data for different canopy scenarios: Relationships between AVIRIS and MODIS indexes, *Remote Sens. Environ.*, 105, 354–366, <https://doi.org/10.1016/j.rse.2006.07.005>, 2006.
- Conservation Science Partners Inc.: Human modification in the western United States, available at <https://databasin.org/datasets/d9d70bfc6e0b46789f1113c63f373c96> (last access: 14 September 2018), 2016.
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., and Lawler, J. J.: Random Forests for Classification in Ecology, *Ecology*, 88, 2783–2792, <https://doi.org/10.1890/07-0539.1>, 2007.
- Daly, C., Neilson, R. P., and Phillips, D. L.: A statistical–topographic model for mapping climatological precipitation over mountainous terrain, *J. Appl. Meteorol.*, 33, 140–158, <https://doi.org/10.1002/asl.228>, 1994.
- Didan, K.: MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006, <https://doi.org/10.5067/MODIS/MOD13Q1.006>, 2015.
- Dietterich, T. G.: Ensemble Methods in Machine Learning, *Lect. Notes Comput. Sc.*, 1857, 1–15, <https://doi.org/10.1007/3-540-45014-9>, 2000.
- Dillon, G. K., Menakis, J., and Fay, F.: Wildland Fire Potential?: A Tool for Assessing Wildfire Risk and Fuels Management Needs, *Proc. Large Wildl. Fires Conf.*, 60–76, 2015.
- Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.-L., Quayle, B., and Howard, S.: A Project for Monitoring Trends in Burn Severity, *Fire Ecol.*, 3, 3–21, <https://doi.org/10.4996/fireecology.0301003>, 2007.
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D.: The Shuttle Radar Topography Mission, *Rev. Geophys.*, 45, 2005RG000183, <https://doi.org/10.1029/2005RG000183>, 2007.
- Finney, M. A.: FARSITE: Fire Area Simulator – model development and evaluation, Res. Pap. RMRS-RP-4, Ogden, UT, U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 47 pp., revised 2004.
- Finney, M. A., Mchugh, C. W., Grenfell, I. C., Riley, K. L., and Short, K. C.: A simulation of probabilistic wildfire risk components for the continental United States, *Stoch. Env. Res. Risk A.*, 25, 973–1000, 2011a.
- Finney, M. A., Grenfell, I. C., Mchugh, C. W., Seli, R. C., Trethewey, D., Stratton, R. D., and Brittain, S.: A Method for Ensemble Wildland Fire Simulation, *Environ. Model. Assess.*, 16, 153–167, <https://doi.org/10.1007/s10666-010-9241-3>, 2011b.
- Gao, B. C.: NDWI – A normalized difference water index for remote sensing of vegetation liquid water from space, *Remote Sens. Environ.*, 58, 257–266, [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3), 1996.
- Giglio, L., Loboda, T., Roy, D. P., Quayle, B., and Justice, C. O.: An active-fire based burned area mapping algorithm for the MODIS sensor, *Remote Sens. Environ.*, 113, 408–420, <https://doi.org/10.1016/j.rse.2008.10.006>, 2009.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, *Remote Sens. Environ.*, 202, 18–27, <https://doi.org/10.1016/j.rse.2017.06.031>, 2017.
- Gray, M., Zachmann, L., and Dickson, B.: Weekly Large Wildfire Probability in Western US Forests and Woodlands, 2005–2017, <https://doi.org/10.6084/m9.figshare.5765967>, 2018.
- Gray, M. E., Dickson, B. G., and Zachmann, L. J.: Modeling and mapping dynamic variability in large fire probability in the lower Sonoran Desert of southwestern Arizona, *Int. J. Wildland Fire*, 23, 1108–1118, <https://doi.org/10.1071/WF13115>, 2014.
- Haas, J. R., Calkin, D. E., and Thompson, M. P.: A national approach for integrating wildfire simulation modeling into Wildland Urban Interface risk assessments within the United States, *Landscape Urban Plan.*, 119, 44–53, <https://doi.org/10.1016/j.landurbplan.2013.06.011>, 2013.
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., and Townshend, J. R. G.: High-Resolution Global Maps of 21st-Century Forest Cover Change, *Science*, 342, 850–853, 2013.
- Helman, D., Lensky, I. M., Tessler, N., and Osem, Y.: A phenology-based method for monitoring woody and herbaceous vegetation in mediterranean forests from NDVI time series, *Remote Sens.*, 7, 12314–12335, <https://doi.org/10.3390/rs70912314>, 2015.
- Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J. N., and Wickham, J.: Completion of the 2001 National Land Cover Database for the

- conterminous United States, *Photogramm. Eng. Rem. S.*, 73, 337–341, 2007.
- Huete, A., Didan, K., Miura, H., Rodriguez, E. P., Gao, X., and Ferreira, L. F.: Overview of the radiometric and biophysical performance of the MODIS vegetation indices, *Remote Sens. Environ.*, 83, 195–213, 2002.
- Krawchuk, M. A. and Moritz, M. A.: Burning issues: statistical analyses of global fire data to inform assessments of environmental change, *Environmetrics*, 25, 472–481, <https://doi.org/10.1002/env.2287>, 2014.
- Krawchuk, M. A., Moritz, M. A., Parisien, M. A., Van Dorn, J., and Hayhoe, K.: Global pyrogeography: The current and future distribution of wildfire, *PLoS One*, 4, e5102, <https://doi.org/10.1371/journal.pone.0005102>, 2009.
- Kuhn, M.: Building Predictive Models in R Using the caret Package, *J. Stat. Softw.*, 28, 1–26, <https://doi.org/10.1053/j.sodo.2009.03.002>, 2008.
- Littell, J. S., McKenzie, D., Peterson, D. L., and Westerling, A. L.: Climate and wildfire area burned in western U.S. ecoprovinces, 1916–2003, *Ecol. Appl.*, 19, 1003–1021, 2009.
- López-Ratón, M., Rodríguez-Álvarez, M. X., Suárez, C. C., and Sampedro, F. G.: OptimalCutpoints?: An R Package for Selecting Optimal Cutpoints in Diagnostic Tests, *J. Stat. Softw.*, 61, 1–36, <https://doi.org/10.18637/jss.v061.i08>, 2014.
- Maki, M., Ishihara, M., and Tamura, M.: Estimation of leaf water status to monitor the risk of forest fires by using remotely sensed data, *Remote Sens. Environ.*, 90, 441–450, <https://doi.org/10.1016/j.rse.2004.02.002>, 2004.
- Martell, D. L., Bevilacqua, E., and Stocks, B. J.: Modelling seasonal variation in daily people-caused forest fire occurrence, *Can. J. Forest Res.*, 19, 1555–1563, <https://doi.org/10.1017/CBO9781107415324.004>, 1989.
- Moritz, M. A., Parisien, M.-A., Batllori, E., Krawchuk, M. A., Van Dorn, J., Ganz, D. J., and Hayhoe, K.: Climate change and disruptions to global fire activity, *Ecosphere*, 3, 49, <https://doi.org/10.1890/ES11-00345.1>, 2012.
- Nolan, R. H., Boer, M. M., Resco De Dios, V., Caccamo, G., and Bradstock, R. A.: Large-scale, dynamic transformations in fuel moisture drive wildfire activity across southeastern Australia, *Geophys. Res. Lett.*, 43, 4229–4238, <https://doi.org/10.1002/2016GL068614>, 2016.
- O'Donnell, M. S. and Ignizio, D. A.: Bioclimatic Predictors for Supporting Ecological Applications in the Conterminous United States, *US Geol. Surv. Data Ser.*, 691, 10 pp., 2012.
- Owen, G., McLeod, J. D., Kolden, C. A., Ferguson, D. B., and Brown, T. J.: Wildfire Management and Forecasting Fire Potential: The Roles of Climate Information and Social Networks in the Southwest United States, *Weather. Clim. Soc.*, 4, 90–102, <https://doi.org/10.1175/WCAS-D-11-00038.1>, 2012.
- Palmer, T. N., Shutts, G. J., Hagedorn, R., Doblaser-Reyes, F. J., Jung, T., and Leutbecher, M.: Representing Model Uncertainty in Weather and Climate Prediction, *Annu. Rev. Earth Pl. Sc.*, 33, 163–193, <https://doi.org/10.1146/annurev.earth.33.092203.122552>, 2005.
- Parisien, M. A., Kafka, V. G., Hirsch, K. G., Todd, J. B., Lavoie, S. G., and Maczek, P. D.: Mapping wildfire susceptibility with the BURN-P3 simulation model, *Nat. Resour. Can., Can. For. Serv., North. For. Cent., Edmonton, Alberta. Inf. Rep. NOR-X-405*, 2005.
- Parisien, M.-A., Walker, G. R., Little, J. M., Simpson, B. N., Wang, X., and Perrakis, D. D. B.: Considerations for modeling burn probability across landscapes with steep environmental gradients: an example from the Columbia Mountains, Canada, *Nat. Hazards*, 66, 439–462, <https://doi.org/10.1007/s11069-012-0495-8>, 2012a.
- Parisien, M.-A., Snetsinger, S., Greenberg, J. a., Nelson, C. R., Schoennagel, T., Dobrowski, S. Z., and Moritz, M. A.: Spatial variability in wildfire probability across the western United States, *Int. J. Wildland Fire*, 21, 313–327, <https://doi.org/10.1071/WF11044>, 2012b.
- Parisien, M. A., Parks, S. A., Krawchuk, M. A., Little, J. M., Flannigan, M. D., Gowman, L. M., and Moritz, M. A.: An analysis of controls on fire activity in boreal Canada: Comparing models built with different temporal resolutions, *Ecol. Appl.*, 24, 1341–1356, <https://doi.org/10.1890/13-1477.1>, 2014.
- Parks, S. A., Holsinger, L. M., Miller, C., and Nelson, C. R.: Wildland fire as a self-regulating mechanism: The role of previous burns and weather in limiting fire progression, *Ecol. Appl.*, 25, 1478–1492, <https://doi.org/10.1890/14-1430.1>, 2015.
- Prasad, A. M., Iverson, L. R., and Liaw, A.: Newer classification and regression tree techniques: Bagging and random forests for ecological prediction, *Ecosystems*, 9, 181–199, <https://doi.org/10.1007/s10021-005-0054-1>, 2006.
- Preisler, H. K., Riley, K. L., Stonesifer, C. S., Calkin, D. E., and Jolly, W. M.: Near-term probabilistic forecast of significant wildfire events for the Western United States, *Int. J. Wildland Fire*, 25, 1169–1180, <https://doi.org/10.1071/WF16038>, 2016.
- Riley, K. L., Abatzoglou, J. T., Grenfell, I. C., Klene, A. E., and 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*, 22, 894–909, <https://doi.org/10.1071/WF12149>, 2013.
- Rollins, M. G.: LANDFIRE: A nationally consistent vegetation, wildland fire, and fuel assessment, *Int. J. Wildland Fire*, 18, 235–249, <https://doi.org/10.1071/WF08088>, 2009.
- Roy, D. P. and Boschetti, L.: Southern Africa validation of the MODIS, L3JRC, and GlobCarbon burned-area products, *IEEE T. Geosci. Remote Sens.*, 47, 1032–1044, <https://doi.org/10.1109/TGRS.2008.2009000>, 2009.
- Roy, D. P., Boschetti, L., Justice, C. O., and Ju, J.: The collection 5 MODIS burned area product – Global evaluation by comparison with the MODIS active fire product, *Remote Sens. Environ.*, 112, 3690–3707, <https://doi.org/10.1016/j.rse.2008.05.013>, 2008.
- Schaaf, C.: MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted RefDaily L3 Global – 500m V006, <https://doi.org/10.5067/MODIS/MCD43A4.006>, 2015.
- Schlobohm, P. and Brain, J.: Gaining an understanding of the National Fire Danger Rating System, A Publication of the National Wildfire Coordinating Group, PMS 932, NFES 2665, 2002.
- Schoennagel, T., Balch, J. K., Brenkert-Smith, H., Dennison, P. E., Harvey, B. J., Krawchuk, M. A., Miettinen, N., Morgan, P., Moritz, M. A., Rasker, R., Turner, M. G., and Whitlock, C.: Adapt to more wildfire in western North American forests as climate changes, *P. Natl. Acad. Sci. USA*, 114, 4582–4590, <https://doi.org/10.1073/pnas.1617464114>, 2017.
- Scott, J. H., Thompson, M. P., and Gilbertson-Day, J. W.: Examining alternative fuel management strategies and the relative contribution of National Forest System land to wildfire

- risk to adjacent homes – A pilot assessment on the Sierra National Forest, California, USA, *Forest Ecol. Manag.*, 362, 29–37, <https://doi.org/10.1016/j.foreco.2015.11.038>, 2016.
- Short, K. C.: A spatial database of wildfires in the United States, 1992–2011, *Earth Syst. Sci. Data*, 6, 1–27, <https://doi.org/10.5194/essd-6-1-2014>, 2014.
- Short, K. C.: Spatial wildfire occurrence data for the United States, 1992–2015 [FPA_FOD_20170508], 4th Edn., Fort Collins, CO, Forest Service Research Data Archive, 2017.
- Stavros, E. N., Abatzoglou, J., Larkin, N. K., Mckenzie, D., and Steel, E. A.: Climate and very large wildland fires in the contiguous western USA, *Int. J. Wildland Fire*, 23, 899–914, <https://doi.org/10.1071/WF13169>, 2014.
- Sullivan, A. L.: Wildland surface fire spread modelling, 1990–2007, 1: Physical and quasi-physical models, *Int. J. Wildland Fire*, 18, 369–386, <https://doi.org/10.1071/WF06144>, 2009a.
- Sullivan, A. L.: Wildland surface fire spread modelling, 1990–2007, 2: Empirical and quasi-empirical models, *Int. J. Wildland Fire*, 18, 369–386, <https://doi.org/10.1071/WF06142>, 2009b.
- Sullivan, A. L.: Wildland surface fire spread modelling, 1990–2007, 3: Simulation and mathematical analogue models, *Int. J. Wildland Fire*, 18, 387–403, <https://doi.org/10.1071/WF06144>, 2009c.
- Taylor, S. W., Woolford, D. G., Dean, C. B., and Martell, D. L.: Wildfire Prediction to Inform Fire Management: Statistical Science Challenges, *Stat. Sci.*, 28, 586–615, <https://doi.org/10.1214/13-STS451>, 2013.
- Thompson, M. P., Riley, K. L., Loeffler, D., and Haas, J. R.: Modeling fuel treatment leverage: Encounter rates, risk reduction, and suppression cost impacts, *Forests*, 8, 1–26, <https://doi.org/10.3390/f8120469>, 2017.
- Turner, M. G.: Disturbance and landscape dynamics in a changing world, *Ecology*, 91, 2833–2849, 2010.
- Tymstra, C., Bryce, R. W., Wotton, B. M., Taylor, S. W., and Armitage, O. B.: Development and structure of Prometheus: the Canadian Wildland Fire Growth Simulation Model, *Nat. Resour. Can., Can. For. Serv., North. For. Cent., Edmonton, AB. Inf. Rep. NOR-X-417*, 2010.
- Urbieto, I. R., Zavala, G., Bedía, J., Gutiérrez, J. M., San Miguel-Ayanz, J., Camia, A., Keeley, J. E., and Moreno, J. M.: Fire activity as a function of fire–weather seasonal severity and antecedent climate across spatial scales in southern Europe and Pacific western USA, *Environ. Res. Lett.*, 10, 114013, <https://doi.org/10.1088/1748-9326/10/11/114013>, 2015.
- Varner, J. M., Keyes, C. R., States, U., Simulator, E. P., Analyst, F. M., and Initiation, C. F.: Fuels treatments and fire models: errors and corrections, *Fire Management Today*, 69, 47–50, 2009.
- Wan, Z., Hook, S., and Hulley, G.: MOD11A2 MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1km SIN Grid V006 [Data set], NASA EOSDIS LP DAAC, <https://doi.org/10.5067/MODIS/MOD11A2.006>, 2015.
- Westerling, A. L.: Warming and Earlier Spring Increase Western U.S. Forest Wildfire Activity, *Science*, 80, 940–943, <https://doi.org/10.1126/science.1128834>, 2006.
- Yang, J., Pan, S., Dangal, S., Zhang, B., Wang, S., and Tian, H.: Continental-scale quantification of post-fire vegetation greenness recovery in temperate and boreal North America, *Remote Sens. Environ.*, 199, 277–290, <https://doi.org/10.1016/j.rse.2017.07.022>, 2017.