

JGR Atmospheres



RESEARCH ARTICLE

10.1029/2023JD040514

Key Points:

- Random forest models trained on weather, fuel, and firefighting data surpass persistence and weather-based methods to predict fire energy
- Random forest models beat persistence across states, for most days in the 2020 fire season, and across levels of fire severity
- Fire weather and latest fire energy predictors add most skill to the random forest; models using agency weather data perform similarly

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

L. H. Thapa,
lthapa@ucla.edu

Citation:

Thapa, L. H., Saide, P. E., Bortnik, J., Berman, M. T., da Silva, A., Peterson, D. A., et al. (2024). Forecasting daily fire radiative energy using data driven methods and machine learning techniques. *Journal of Geophysical Research: Atmospheres*, 129, e2023JD040514. <https://doi.org/10.1029/2023JD040514>

Received 1 DEC 2023

Accepted 7 JUL 2024

Author Contributions:

Conceptualization: Laura H. Thapa,

Pablo E. Saide, Jacob Bortnik, Arlindo da Silva, David A. Peterson

Data curation: Melinda T. Berman,

Arlindo da Silva, Fangjun Li, Shobha Kondragunta, Ravan Ahmadov, Eric James, Johana Romero-Alvarez, Amber Soja, Elizabeth Wiggins, Emily Gargulinski











Formal analysis: Laura H. Thapa

Funding acquisition: Laura H. Thapa, Pablo E. Saide, Arlindo da Silva

Investigation: Laura H. Thapa

Methodology: Laura H. Thapa, Pablo E. Saide, Jacob Bortnik, Ravan Ahmadov, Eric James, Johana Romero-Alvarez

Forecasting Daily Fire Radiative Energy Using Data Driven Methods and Machine Learning Techniques

Laura H. Thapa¹ , Pablo E. Saide^{1,2} , Jacob Bortnik¹ , Melinda T. Berman³ , Arlindo da Silva⁴ , David A. Peterson⁵, Fangjun Li⁶ , Shobha Kondragunta⁷ , Ravan Ahmadov⁸, Eric James^{8,9}, Johana Romero-Alvarez^{8,9}, Xinxin Ye¹⁰, Amber Soja^{11,12} , Elizabeth Wiggins¹¹ , and Emily Gargulinski¹² 

¹Department of Atmospheric and Oceanic Sciences, University of California Los Angeles, Los Angeles, CA, USA,

²Institute of the Environment and Sustainability, University of California Los Angeles, Los Angeles, CA, USA,

³Department of Climate, Meteorology and Atmospheric Sciences, University of Illinois Urbana-Champaign, Urbana, IL, USA, ⁴NASA Global Modeling and Assimilation Office, GSFC, Greenbelt, MD, USA, ⁵Naval Research Laboratory,

Monterey, CA, USA, ⁶Department of Geography and Geospatial Sciences, South Dakota State University, Brookings, SD, USA, ⁷NOAA NESDIS Center for Satellite Applications and Research, College Park, MD, USA, ⁸NOAA Global Systems

Laboratory, Boulder, CO, USA, ⁹Cooperative Institute for Research in Environmental Sciences, University of Colorado

Boulder, Boulder, CO, USA, ¹⁰Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, China, ¹¹NASA

Langley Research Center, Hampton, VA, USA, ¹²National Institute of Aerospace, Hampton, VA, USA

Abstract Increasing impacts of wildfires on Western US air quality highlights the need for forecasts of smoke emissions based on dynamic modeled wildfires. This work utilizes knowledge of weather, fuels, topography, and firefighting, combined with machine learning and other statistical methods, to generate 1- and 2-day forecasts of fire radiative energy (FRE). The models are trained on data covering 2019 and 2021 and evaluated on data for 2020. For the 1-day (2-day) forecasts, the random forest model shows the most skill, explaining 48% (25%) of the variance in observed daily FRE when trained on all available predictors compared to the 2% (<0%) of variance explained by persistence for the extreme fire year of 2020. The random forest model also shows improved skill in forecasting day-to-day increases and decreases in FRE, with 28% (39%) of observed increase (decrease) days predicted, and increase (decrease) days are identified with 62% (60%) accuracy. Error in the random forest increases with FRE, and the random forest tends toward persistence under severe fire weather. Sensitivity analysis shows that near-surface weather and the latest observed FRE contribute the most to the skill of the model. When the random forest model was trained on subsets of the training data produced by agencies (e.g., the Canadian or US Forest Services), comparable if not better performance was achieved (1-day $R^2 = 0.39$ – 0.48 , 2-day $R^2 = 0.13$ – 0.34). FRE is used to compute emissions, so these results demonstrate potential for improved fire emissions forecasts for air quality models.

Plain Language Summary Increasing wildfire smoke is undoing decades of air quality progress, yet air quality forecasts often miss the most intense smoke events. This is because forecasted smoke is released at constant rates whereas the rate of smoke release from real wildfires varies in time. In this work we teach a machine learning algorithm to predict the daily change in fire heat output, a quantity that is used to calculate wildfire emissions. The machine learning algorithm uses information regarding weather, fuel moisture and amount, and firefighting efforts to make its predictions. We also tried to predict the daily change in fire heat output using only weather information but found the machine learning method to be more successful. Many federal agencies have their own ways of tracking fire weather and fuel moisture, and in this paper, we show that we can apply machine learning to the data from any of several agencies and get the same level of forecasting skill.

1. Introduction

The worsening state of wildfires in the Western United States and Canada is causing increased exposure to harmful particulate matter and degrading air quality (Burke et al., 2021; Munoz-Alpizar et al., 2017), and our warming climate is known to exacerbate these effects (Westerling et al., 2006; Zhuang et al., 2021). Chemical transport models are vital for warning the public about episodes of poor air quality and for quantifying the climate impacts of wildfire smoke, but uncertainty in smoke emissions can propagate through these predictions and cause large errors (Carter et al., 2020; Ye et al., 2021). In chemical transport models, fire emissions are an essential model input and

© 2024. The Author(s).

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Project administration: Pablo E. Saide

Resources: Pablo E. Saide, Melinda T. Berman, Arlindo da Silva, Fangjun Li, Shobha Kondragunta, Ravan Ahmadov, Eric James, Johana Romero-Alvarez, Xinxin Ye, Amber Soja, Elizabeth Wiggins, Emily Gargulinski

Software: Laura H. Thapa, Melinda T. Berman, Johana Romero-Alvarez, Xinxin Ye

Supervision: Pablo E. Saide, Jacob Bortnik, Ravan Ahmadov, Eric James, Johana Romero-Alvarez

Validation: Laura H. Thapa, Pablo E. Saide

Visualization: Laura H. Thapa

Writing – original draft: Laura H. Thapa

Writing – review & editing: Pablo E. Saide, Melinda T. Berman, David A. Peterson, Fangjun Li,

Shobha Kondragunta, Ravan Ahmadov, Eric James, Xinxin Ye, Amber Soja, Emily Gargulinski

are commonly estimated using the burned area-based conventional approach (Seiler & Crutzen, 1980) or the Fire Radiative Power (FRP)-based method (Ichoku et al., 2008; Wooster et al., 2005). In forecasting mode, many air quality models copy the most recent daily satellite-derived emissions throughout the forecast period, an assumption referred to as persistence. Sub daily emissions are forecasted using a prescribed climatological diurnal cycle, which assumes a peak in fire intensity in the local afternoon (e.g., Ahmadov et al., 2017). A recent study (Ye et al., 2021) has shown that persistence can result in over- or underpredicted emissions in wildfires when daily changes in fire activity are large using 12 forecast systems deployed in the Fire Influence on Regional to Global Environments and Air Quality (FIREX-AQ) field campaign. Persistence may delay the onset or extend the lifetime of fires within the forecast window (Di Giuseppe et al., 2018; Thapa et al., 2022). In spite of advances such as inverse modeling of emissions (e.g., Saide et al., 2015) and statistical methods to forecast fire evolution (Di Giuseppe et al., 2018; Graff et al., 2020; Huot et al., 2022; D. Peterson et al., 2013; Preisler & Westerling, 2007; Wang et al., 2022) operational smoke forecasts rely primarily on persistence to forecast daily emissions totals. Precipitation is simplistically included in some of the models (Ahmadov et al., 2017).

Wildfire spread and intensity are governed by the interactions between three sides of “the fire behavior triangle”: weather, topography, and fuel (Countryman, 1972). First, the atmosphere can drive fire via low humidity, high temperature and wind speed, and vertical instability (Rothermel, 1991; Srock et al., 2018). Second, fires spread more easily in an uphill direction (Rothermel, 1972), as a steep slope allows for preheating of fuels, but complex flows in mountainous regions may confound this relationship (Viegas, 2004). Finally, fuel type, moisture, and loading are key components of determining the intensity of a fire (Rothermel, 1972). The moisture of dead fuels and organic soils respond over the course of a day to changing weather conditions (Wagner, 1987), while the moisture content of living fuels and organic soils can hold a 1–2 month memory of previous moisture conditions (Krueger et al., 2022). The latter has been shown to be a good indicator of fire intensity (Sazib et al., 2022).

Currently, wildfire indices are used to inform wildfire behavior predictions, plan firefighting strategies, and forecast fire intensity. These indices may represent one aspect of the fire behavior triangle, such as the hot-dry-windy index (HDW) or the continuous Haines index (CHI) which represent the effect of meteorology on wildfire (Pinto et al., 2020; Srock et al., 2018). These indices may represent multiple aspects, such as the Canadian Forest Fire Danger Rating System (CFFDRS), the National Fire Danger Rating System (NFDRS), or the National Oceanic and Atmospheric Administration (NOAA) hourly wildfire potential (HWP) which represents fuel and weather effects on wildfire (Bradshaw et al., 1984; Wagner, 1987; Text S1 in Supporting Information S1). To the best of our knowledge, wildfire indices which include all aspects of the fire triangle are rare. However, wildfire indices such as the CFFDRS fire weather index (FWI) (Di Giuseppe et al., 2017, 2018) and vapor pressure deficit (VPD), and/or wind speed have been used to successfully modulate persistence-based forecasts (Graff et al., 2020) forecasts or to derive conditional probability estimates of FRP exceedances (Hernandez et al., 2015). Additionally, components of the NFDRS, notably the energy release component (ERC), have served as inputs to machine learning frameworks which have predicted monthly PM_{2.5} and daily fire perimeters (Huot et al., 2022; Wang et al., 2022). Other fire weather indices not covered in this paper, such as the Grassland Fire Danger Index, have shown promise for modulating climatological predictions of fire occurrence in regions such as sub-Saharan Africa which have a regular burning season (Partanen & Sofiev, 2022).

In addition to the natural fire behavior triangle, firefighting may also shape the evolution of a wildfire. Naturally occurring wildfires may be managed or suppressed to thin fuels or protect life and property (Young et al., 2019), and this may affect the final burned area (Hu & Ntamo, 2009). Since burned area is correlated with both FRP and emissions (e.g., Thapa et al., 2022; Wiedinmyer et al., 2011), we would expect containment to impact these quantities as well. However, much of the study of firefighting as it pertains to intensity and burned area has focused on including containment in fire spread modeling (Hu & Ntamo, 2009), planning the most effective containment strategies (Rodrigues et al., 2020), and understanding containment successes and failures (Arienti et al., 2006). Only recent studies have connected containment efforts, fire spread, and FRP (Jolly et al., 2019; Turney et al., 2023). While these studies underscore the importance of including containment in fire spread forecasts, they also highlight that containment efforts can be hindered by severe fire weather and the inaccessibility of some fires (Arienti et al., 2006; Rodrigues et al., 2020).

Machine learning techniques are a subset of statistical models which can learn underlying relationships in systems with many predictive variables. The random forest machine algorithm has been applied successfully to multiple areas within the wildfire problem including understanding drivers of fire severity in mediterranean pine forests

(García-Llamas et al., 2019), modeling containment probability using land surface variables (Rodrigues et al., 2020), and detecting fires from geostationary satellites over South Korea (Jang et al., 2019). More broadly, statistical models have been applied to predict fire evolution using meteorological quantities, fire weather and moisture indices, and land cover data sets. In Graff et al. (2020), the authors train a poisson regression on reanalysis data from the European Center for Medium-range Weather Forecasts (ERA-Interim) and MODIS land cover and are able to predict fire counts with more skill than persistence out to 5 days. D. Peterson et al. (2013) used a maximum likelihood classification method to predict the daily change in fire counts based on CFFDRS fire indices, convective available potential energy, moisture, and synoptic weather, and achieved better results than persistence. Finally, Pinto et al. (2020) compare the utility of the CFFDRS FWI alone and the CFFDRS FWI and the CHI, and they find that using both surface and stability information allows for a more skillful forecast of FRP exceedance probabilities. Generally statistical forecasts which have beaten persistence have tended to focus on two of the three legs of the fire behavior triangle. However, Huot et al. (2022) and Wang et al. (2022) successfully trained various machine learning algorithms on weather, topography, and fuel to generate maps of burned area and monthly PM_{2.5}. In these studies, some human influence variables were included (population density, GDP), but predictors relating to firefighting were not included.

In this work, we use weather indices and random forest machine learning as tools to forecast daily changes in fire radiative energy (FRE, the time-integrated form of FRP) in the Western US with the goal of improving upon the persistence assumption and maintaining computational efficiency. Unlike previous work (e.g., Huot et al., 2022; Wang et al., 2022) this work is directly applicable for emissions inventories which use FRP to compute emissions. Although we evaluate FRE only in this work, we can assume since emissions are proportional to FRE, any changes in FRE dictated by our models would translate to the same change in emissions. For the machine learning approaches, the training data set consists of variables describing near-surface weather, atmospheric stability, fuel density and water stress, fuel moisture, and firefighting efforts, and our target variable is based on a blended geostationary/polar orbiting emissions product (Li et al., 2022). All algorithms were trained on data from the 2019 and 2021 fire seasons and tested on comparable data from the 2020 fire season. In this work we look at July–December, as these were the months which contained significant fire activity from both summertime and wind-driven fire events. The methods tested represent varying levels of computational complexity, from daily changes in weather being assumed to be proportional to daily changes in fire intensity, to random forest machine learning.

2. Data

This study focuses on wildfires which occurred between July and December in 2019–2021 in the Western Continental US (Lat: 31–49°N, Lon: 125–101°W). Features were extracted (Section 3.2) for each day of each fire, totaling 1,238 training samples (2019 and 2021) and 1,293 1-day (1,160 2-day) testing samples (2020). Table 1 lists the availability, resolution, aggregation method, and primary reference for each data set used in this study. GFWED, NCAR Fuel Moisture, and final Monitoring Trends in Burn Severity (MTBS) fire perimeters, are only available for portions of the 1 January 2019–31 December, 2021 period. RAVE data are not available before 1 January, 2019. At the time of analysis, MTBS perimeters for the 2021 fires were considered provisional. For further description of the data sets used, see Text S1 in Supporting Information S1.

3. Methods

3.1. Building Polygons

Previous work has tracked sub-daily to daily wildfire movement by aggregating VIIRS 375 m fire detections into polygons (Berman et al., 2023; McClure et al., 2023). In this work, a similar approach is taken:

1. For each overpass and each active fire, all VIIRS detections potentially corresponding to each fire are selected. Corresponding points are identified as points within a 0.25-degree radius of the previous day's polygon (gray shaded region in Figure 1a, Step 2). If it's the first day of the fire, the “previous day's polygon” is the ignition location from the SIT-209. The radius is increased by 0.25° incrementally and the search repeats recursively until no new points are added, with the 0.25-degree buffer chosen to reduce computing time. This process is shown graphically in Figure 1a for the first 2 days of the August 2020 Complex fire.

Table 1
Summary of Data Used in This Work, Including Temporal Coverage, Temporal and Spatial Resolution, Aggregation Method, and Primary Reference

Data set (Type of data)	Available dates in 2019–2020	Resolution	Variables extracted	Aggregation method	Reference (date of access)
VIIIRS (fire location)	SNPP: 1/1/2019–12/31/2021 NOAA-20: 1/1/2020–12/31/2021	375 m, twice daily	Lat/lon of detection	Aggregate detections into polygons (Section 3.3.1)	Schroeder et al. (2014) (accessed 9/13/2021, 10/20/2022, 4/5/2023)
MTBS (fire location)	All 2019–2020, partial 2021. 2021 Filled with MTBS provisional perimeters.	Total burn perimeter, one per fire	Final fire perimeter	N/A	Eidenshink et al. (2007a), Eidenshink et al. (2007b) (MTBS April, 2023 and Provisional May 2023)
RAVE (fire intensity)	1/1/2019–12/31/2021	3 km hourly	Fire Radiative Energy (FRE)	Summed over day and polygon	Li et al. (2022) (March 2023)
GridMET (fire weather, fuel moisture)	1/1/2019–12/31/2021	4 km daily	Vapor pressure deficit (VPD); wind speed (WS); max and min temperature (T_{max} , T_{min}); NFDERS Burning Index (BI) and Energy Release Component (ERC); 100 and 1,000 hr fuel moisture	Averaged over day and polygon	Abatzoglou (2013a) (August, 2023)
HRRR (fire weather, soil moisture)	1/1/2019–12/31/2021	3 km, hourly	VPD, WS, hot-dry, windy index (HDWI); hourly wildfire potential (HWP); continuous Haines index (CHI); pyroCb firepower threshold (PFT); surface, 1, 4, 10, 30 cm soil moisture	Averaged over day and polygon	Dowell et al. (2022a), Dowell et al. (2022b) (March–April, 2023)
GFVED (fire weather)	1/1/2019–9/30/2021	0.1°, daily	GFVED fire weather index (FWI); buildup index (BUI); fine fuel moisture code (FFMC)	Averaged over day and polygon	Field (2015), Field et al. (2015) (April, 2023)
NOAA Soil Moisture (moisture)	1/1/2019–12/31/2021	0.25°, daily	Surface soil moisture (Blended_SM)	Averaged over day and polygon	Liu et al. (2016a), Liu et al. (2016b) (March–April, 2023)
NCAR Fuel Moisture (moisture)	8/1/2019–11/26/2021	1 km, daily	Dead fuel moisture (FMCG2D); live fuel moisture (FMCG1H2D)	Averaged over day and polygon	McCandless et al. (2020) (September–October, 2021)
Evaporative Stress Index (ESI) (moisture)	1/1/2019–12/31/2021	5 km, weekly	ESI	Averaged over day and polygon	Anderson et al. (2011) (April, 2023)
Plant Water Sensitivity (PWS) (moisture)	Assumed static for 2019–2021 period	4 km, static	PWS	Static map copied to daily, averaged over day and polygon	Rao et al. (2022) (August 2023)
Fuel Loading	Assumed static for 2019–2021 period	900 m, static	Low_N, Moderate_N, High_N, VeryHigh_N, Extreme_N	Static maps copied to daily, averaged over day and polygon	D. A. Peterson et al. (2022); Soja et al. (2004) (February, 2023)
Terrain	Assumed static for 2019–2021 period	900 m, static	Slope, elevation (mean and sidev)	Static map copied to daily, averaged over day and polygon	Landfire, https://www.landfire.gov/version_download.php (August, 2023)
SIT-209 (firefighting)	1/1/2019–12/31/2021	N/A, ~daily	Containment (%), structures, crews, engines, aircraft, construction, overhead	N/A	Jamieson, (2005a, 2005b) (2022)
Population Density	Assumed static for 2019–2021 period	2.5 arcmin, static	Population density	Static map copied to daily, averaged over day and polygon	CIESIN (2018a)

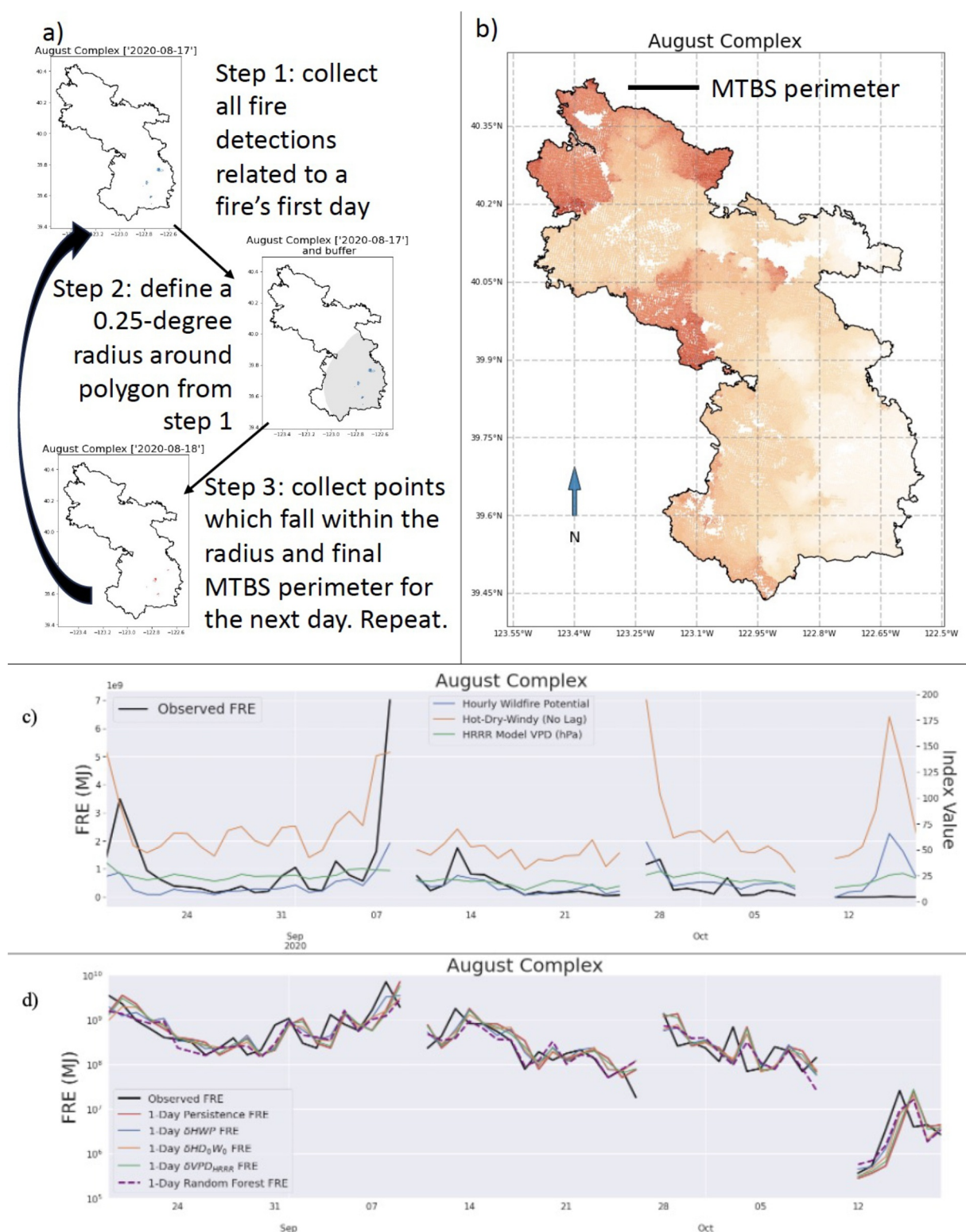


Figure 1. Workflow example for the August 2020 Complex Fire. (a) the process used to build 2 days of fire polygons for August Complex and (b) the final August Complex polygon at daily resolution (light orange for mid-August 2020 through dark red for mid-October 2020). Time series of (c) fire radiative energy (FRE) (black) and near surface fire weather indices (HWP—blue, HDW—orange, HRRR VPD—green) extracted using the daily polygons, and (d) observed (black) and forecasted FRE using persistence (red) scaled weather (HWP—blue, HDW—orange, HRRR VPD—green) and random forest model trained on all predictors (magenta dashed).

2. The per-overpass polygon is built by aggregating all fire detections, each of which have been buffered to a circle with a diameter matching the area of VIIRS pixels at nadir (375 m × 375 m), using the Python Shapely module's multipolygon datatype (Gillies et al., 2007).
3. Sub-daily polygons are aggregated to daily resolution and clipped using the final burn perimeters from MTBS. Clipping with final MTBS perimeters confirms that the sub-daily polygons generated in Step 2 are associated with the correct fire event. Figure 1b shows the final polygon for August Complex.

This method evaluates well against the method developed by McClure et al. (2023) for the August Complex Fire (not shown).

3.2. Extracting Data Using Polygons

Due to the variable resolutions of training variable data sets, the polygons discussed above are used to extract only the portions of the training data which directly impact each fire on each day (Figure 1c, Figure S1 in Supporting Information S1). This method has several advantages. First, it allows us to track the movement of the fire as a whole and the factors affecting the fire (including firefighting, which is reported per-fire) on a daily basis. Second, it is flexible for fires of different sizes. Finally, this method is capable of preserving high resolution information in cases where the feature grid may be finer resolution than the polygon size (Figure S2 in Supporting Information S1). A polygon representing a fire on a single day (e.g., Figure 1a) may be used to extract data as follows:

1. Grid cells falling in a bounding box defined by the polygon are determined, and each grid cell is transformed into a polygon using the shapely python library (Gillies et al., 2007).
2. The intersection area of each grid cell with the polygon is calculated (Gillies et al., 2007), and the fraction of the fire polygon which intersects each grid cell is saved.
3. Most variables are extracted as a weighted average over all the grid cells which intersect the polygon, with the weights being the results of step 2. FRE is an exception as it is extracted as the unweighted sum of all grid cells which intersect the polygon.
4. Values of the variables are aggregated (averaged or summed, Table 1) to daily resolution, with the day beginning at 12Z. 12Z was chosen to preserve the midafternoon peak in fire intensity, which occurs around 0Z.

3.3. Predicting FRE

Daily FRE values span around six orders of magnitude (Figures S3a and S3b in Supporting Information S1), owing to the extreme nature of the chosen wildfire seasons. Therefore, we introduce the scaling operator δ (Equation 1) which represents the daily change in a variable X over a fire polygon from 1 day to the next (Figure 1d).

$$\delta X = \frac{X_t}{X_{t-1}}, \quad (1)$$

where t is the valid time of the forecast, and $t-1$ is the previous day.

Applying the scaling operator to FRE (δ FRE) gives the right-skewed distribution in Figure S3c of Supporting Information S1. To deal with skew, the machine learning models are trained to predict the log of δ FRE (Figure S3d in Supporting Information S1) so as to give equal weight to large increases (δ FRE $\gg 1$) and large decreases (δ FRE $\ll 1$). Applying the scaling operator as well as the log transformation of FRE helps the model to deal with how different aspects of the fire behavior triangle affect the evolution of wildfires at any size. Working with a scaling factor allows us to further simplify the persistence model, as persistence would predict a scaling factor of 1 for all cases. We removed scaling factors which lie more than 1.5 interquartile ranges from the median, since our goal is to predict the evolution of existing wildfires, rather than how fires start or end. To minimize the impact of temporal covariance, 2019 and 2021 were chosen as the training periods and 2020 was chosen as the testing period. Scaled methods are evaluated for 2020 only so that the comparison is consistent.

In a strategy similar to that presented in Graff et al. (2020), Equation 1 can be applied to FRE and rearranged to obtain an FRE prediction (Equation 2, Figure 1d). The latest FRE observation is derived from RAVE as a daily (12Z–12Z) sum of hourly FRE (Section 3.2). Equation 2 can also be used to perform multiple steps of integration in time (Equation 3). The following sections (Sections 3.3.1 and 3.3.2) discuss in detail how δ_{pred} is derived.

$$\text{FRE}_{\text{pred},t} = \delta_{\text{pred},t} * \text{FRE}_{\text{obs},t-1} \quad (2)$$

$$\text{FRE}_{\text{pred},t+1} = \delta_{\text{pred},t+1} * \text{FRE}_{\text{pred},t} \quad (3)$$

3.3.1. Scaled Method

The scaled fire index (SFI) method applies the scaling operator (Equation 1) to daily and polygon-averaged values of fire indices to generate a value with which to modulate persistence. To do this, HRRR wind speed and VPD, GridMET wind speed and VPD, HDW, HWP, FWI, and BI (Text S1 in Supporting Information S1) are plugged into Equation 1. The lagged version of the HDW (HDmW0, Text S1 in Supporting Information S1) was also tested as a scaled variable. In subsequent discussion and figures, SFIs are denoted as δ Index, for instance δ HWP for the scaling operator applied to the HWP. SFIs are included to be a simpler alternative to the random forest that does not require training a machine learning algorithm.

We exclude cases where a SFI gives outliers that are further than 1.5 times the interquartile range away from the median. In total 9 (51) 1-day (2-day) forecast fire days contained a scaled index which was considered an outlier. 38 of the 2-day SFIs were zero values, primarily from zeros in the NFDRS BI and CFFDRS FWI. The remaining outliers were from δ HWP and δ BI, which contained unphysically large or small values (Tables S1–S4 in Supporting Information S1). This filtering was done in addition to filtering outliers from the observed δ FRE (Section 3.3, Figure 2d). This filtering was applied before the random forest was trained, resulting in 1,296 training fire-days for 2019 and 2021 (382 decreases, 435 slight changes, and 479 increases) and 1,293 (1,160) 1-day (2-day) fire-days to evaluate for 2020.

3.3.2. Random Forests

Random forests (RF) are a machine learning technique that use ensembles of decision trees which split the training data set into homogeneous subgroups and average predictions across trees in the forest (Breiman, 2001). During exploratory research, we tested the RF, support vector machine, multilayer perceptron neural network, and XGBoost, and the RF model performed the best and had a relatively shorter computing time. In this work, the scikit-learn implementation is used, with the default values used for all but two hyperparameters. We found that `min_samples_leaf` = 20, the parameter controlling the number of sample points needed for a leaf node to be created, and `max_depth` = 15, the parameter for controlling the depth of the trees in the forest, were best for minimizing overfitting. The `random_state` variable, which controls the bootstrapped subsampling of the feature sets used to train the trees in the forest, was set to 42 to ensure that models were trained on the same data points across feature configurations.

Our RF is trained on the full set of training variables (Table 1), including SFIs (Section 3.3.1) as well as subsets (Section 4.2–4.3). The RF was trained to predict the log of δ FRE to allow the model to learn both the large increases and the large decreases (Figure S3d in Supporting Information S1). The RF was trained on data extracted over observed polygons, assuming that, since these fires were past, conditions on the day of the prediction could be known. On the other hand, the random forest was tested using data extracted over the latest observed polygons to model real case scenarios where the location of the fire into the forecast window is unknown. Forecasted weather variables were used to make the 1- and 2-day predictions of fire growth and decay, and for other variables (fuel, topography, firefighting, moisture, FRE) the day 0 values over the latest observed polygon were persisted.

The 1-day RF prediction was made following Equation 2, where $\delta_{\text{pred},t}$ was derived from the RF which was given day 1 and day 0 weather inputs and day 0 fuel, topography, firefighting, moisture, and FRE inputs. The 2-day RF prediction was derived following Equation 3 where $\delta_{\text{pred},t+1}$ was derived from the RF which was given day 2 and day 1 weather inputs and day 0 fuel, topography, firefighting, moisture inputs, and predicted day 1 FRE values.

3.4. Evaluation Metrics

The SFIs and the RF model were evaluated on their ability to represent the daily change in fire behavior (increase, decrease, slight change) and FRE magnitude. Predicted scaling factors and FRE values (δ_{pred} and FRE_{pred}) were compared with observed scaling factors and FRE values (δ_{obs} and FRE_{obs}) for the testing period.

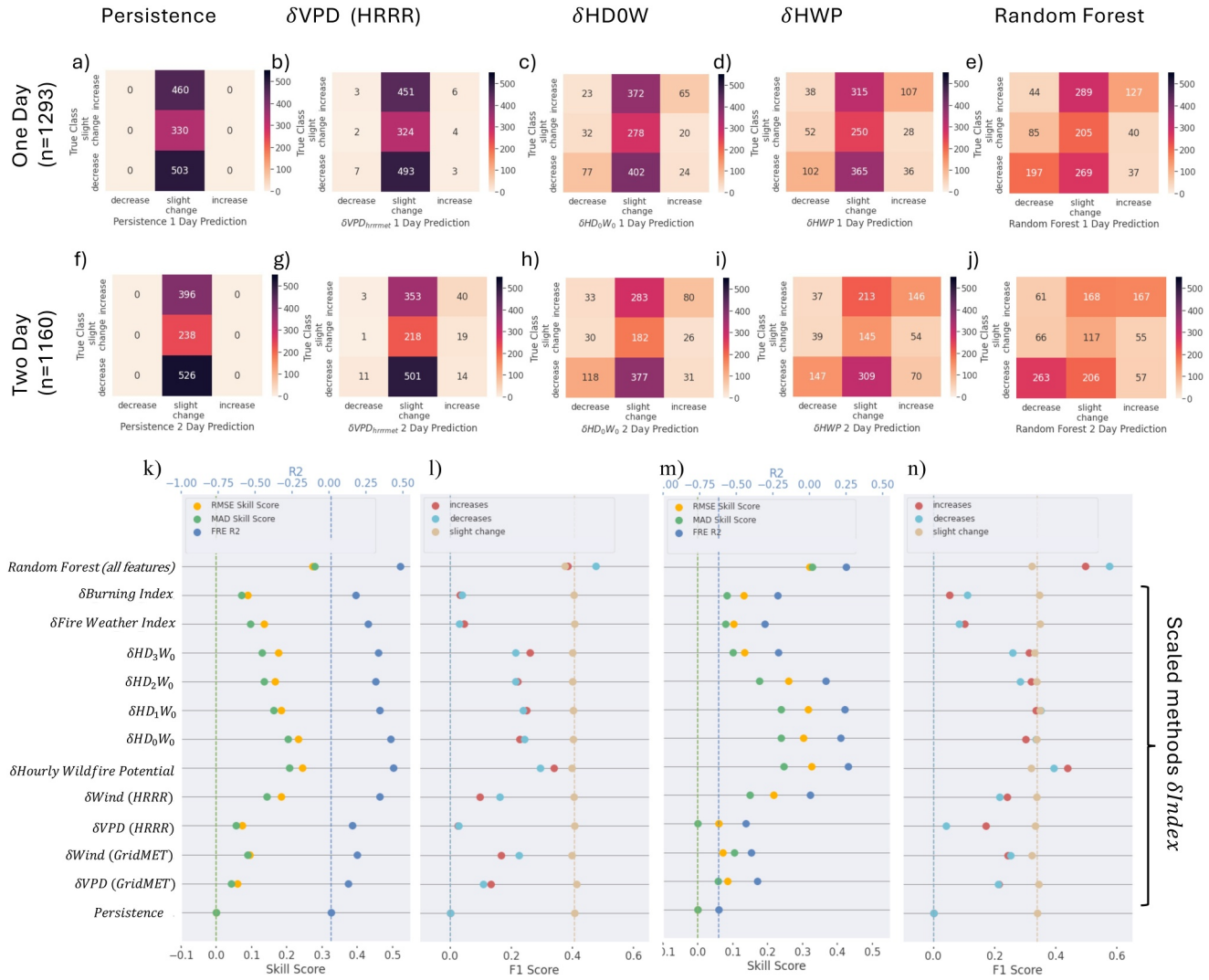


Figure 2. Comparison of persistence, scaled, and random forest methods. Performance of the (a and f) persistence (b and g) δ VPD (HRRR), (c and h) δ HD0W0, (d and i) δ HWP, (e and j) random forest method on their ability to forecast fire growth category, $F = 0.67$ – 1.5 is slight change and greater (less) than that value is an increase (decrease). R^2 and SS_{RMSE} and SS_{MAD} are shown for fire radiative energy for all 1-day (k) and 2-day (m) predictions in blue, yellow, and green dots respectively. F1 for all 1-day (l) and 2-day (n) predictions with increases in red, decreases in blue and slight changes in tan. Rows are labeled according to the forecasting method, with scaled methods labeled with a “ δ ” preceding the name of the scaled variable. Vertical dashed lines denote performance of persistence.

3.4.1. Evaluating Performance to Predict Categories

Whether FRE for a given fire increased, decreased, or remained constant on a day-to-day basis is based on thresholds of the δ FRE, where δ FRE = 0.67 – 1.5 (33% decrease–50% increase) represents slight change and anything above (below) that range is an increase (decrease). The ability of the model to forecast the daily change category is assessed using 3-by-3 tables and the associated F1 statistic (Figure 2). The F1 statistic (Equation 6) is derived from the harmonic mean of the precision (Equation 4) and recall (Equation 5), and higher values of F1 indicate both a higher proportion of true category instances modeled (recall) and a higher proportion of correct predictions of a given category (precision, Powers, 2008).

$$P_k = \frac{\text{True predictions of category } k}{\text{All predictions of category } k} \quad (4)$$

where k = increase, decrease, and slight change

$$R_k = \frac{\text{True predictions of category } k}{\text{All observed occurrences of category } k} \quad (5)$$

where k = increase, decrease, and slight change

$$F1_k = 2 \frac{P_k * R_k}{P_k + R_k} \quad (6)$$

where k = increase, decrease, and slight change.

3.4.2. Evaluating Scaling Factor and FRE Performance

Predicted scaling factor and FRE magnitude are assessed using R^2 , mean absolute deviation (MAD), and root-mean-squared error (RMSE). RMSE and MAD are defined as in Willmott and Matsuura (2005), and R^2 is the coefficient of determination. The adjusted R^2 (Equation 7) is also computed and can be interpreted as the percent of the variance explained by the model, adjusted for the number of predictors used to train the model. Adjusted R^2 and R^2 yield very similar results, indicating that more than additional model features are responsible for improved model performance (Table S5 in Supporting Information S1).

$$\text{Adjusted } R^2 = 1 - \frac{(1 - R^2) * (n - 1)}{(n - k - 1)} \quad (7)$$

where R is the coefficient of determination, n is the number of observations, and k is the number of predictors.

We report the unitless skill score (SS), the relative change in performance metric for a model with respect to a reference model. Skill scores are calculated for RMSE and MAD (SS_{RMSE} and SS_{MAD}) using persistence as the reference model (Graff et al., 2020; D. Peterson et al., 2013). R^2 and SS_{RMSE} are sensitive to large values, so these statistics will see larger improvements when large increases are correctly forecasted.

4. Results

4.1. Comparison of Two Modeling Approaches With Persistence

Figure 3 compares the performance of the persistence model, the SFI method (δ Index, Section 3.3.1), and the RF trained on the entire set of training variables (Table 1) over 2 days for the 2020 fire season. Figures 3a–3e and 3f–3j show modeled and observed growth categories (Section 3.4.1) for selected forecasts of δ FRE. Figures 3k–3m summarize error metrics for all methods, with Figures 3k and 3m showing performance of the FRE forecasts (R^2 , SS_{RMSE} , SS_{MAD}) and Figures 3l and 3n showing performance of the forecasted categories ($F1_{\text{decrease}}$ and $F1_{\text{increase}}$).

Overall, Figure 2 shows that the random forest is the top method tested and also that there is utility in using SFIs to predict FRE. In both the 1- and 2-day forecasts, the RF is the most skillful method (Figures 2e, 2j and 2k–2n) except for in R^2 in the 2-day forecast (Figure 2m), where it is beat only by scaled HWP. In the 1-day forecast, skill at predicting FRE or growth category increases as the SFIs get more complex or the RF is used (Figure 2k). For the 2-day forecasts (Figures 2m and 2n) the SFIs using simpler fire weather indices or raw meteorological variables (HDW, HWP, VPD) can show comparable FRE prediction skill to the RF. All methods tend to do better than persistence at forecasting FRE, and they all forecast increases and decreases at roughly the same rate.

Certain models (persistence, δ VPD (HRRR), δ HD0W0, δ HWP, and random forest) are highlighted to illustrate common error types (Figures 2a–2f, Figures S4a–S4j and S5a–S5j in Supporting Information S1). Figures 2a and 2f show, as expected, that the persistence model is only capable of forecasting slight changes in FRE. The “persistence” rows of Figures 2l and 2n show the F1 score of the increases and decreases is zero, indicating an inability for persistence to capture increases (decreases) of >50% (<33%) over the latest FRE measurement. In spite of this limitation, persistence gives a slightly capable 1-day FRE forecast (Figure 2k, $R^2 = 0.02$) and a severely limited 2-day FRE forecast (Figure 2m, $R^2 = -0.62$) for 2020 fires.

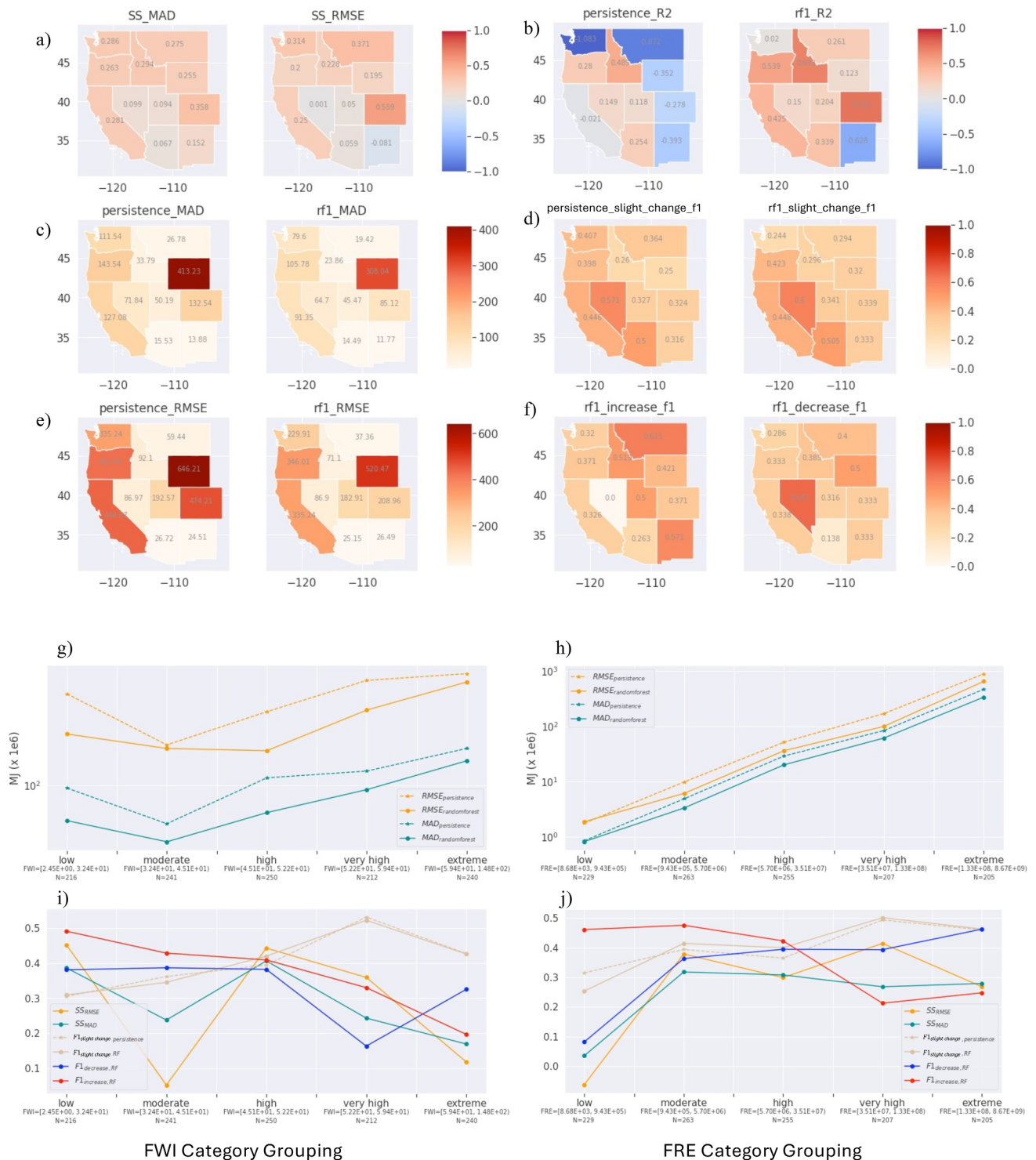


Figure 3. Random forest evaluation metrics by state (a–f) and by fire severity (g–j) for the 1-day RF and persistence predictions. (a) SS_{MAD} and SS_{RMSE} for the RF model; (b) R² for the persistence and RF models; (c) mean absolute deviation (MAD) for the persistence and RF models; (d) F1_{slight_change} for the persistence and RF models; (e) RMSE for the persistence and RF models; (f) F1_{increase} and F1_{decrease} for the RF model; (g) RMSE (yellow) and MAD (teal) for the persistence (dotted) and RF (solid) models split by 20th-percentiles of FWI; (h) as in f but split by 20th-percentiles of fire radiative energy (FRE). (i) SS_{MAD} (teal), SS_{RMSE} (yellow), F1_{increase} (red), F1_{decrease} (blue), and F1_{slight_change} (tan) for the RF models with F1_{slight_change} (tan dotted) for the persistence model split by 20th percentiles of fire weather index. (j) as in i but split by 20th percentiles of FRE.

Figures 2b–2c and 2g–2h show that assuming that daily changes in FRE are driven by daily changes in VPD (δ VPD (HRRR)), HD0W0 (δ HD0W0), and HWP (δ HWP) adds skill when compared with the persistence forecast. These forecasts are relatively low risk in that they boost performance on FRE ($R^2 = 0.345, 0.437$, and 0.418 for the three methods respectively vs. 0.02 for persistence, Figures 2k and 2l), while not straying too far from the categories persistence would predict (Figures 2i and 2n). Across both the 1-day and 2-day forecasts, derived indices δ HD0W0 and δ HWP are more skilled than raw variables δ VPD and δ WS (Figures 2k–2n). However, some of the more complex SFIs (δ BI and δ FWI) struggle to reliably capture fire growth or decay (Figures 2i and 2n) so perhaps index complexity is not always an indicator of ability to capture fire growth.

Figures 2e and 2j show that the random forest trained on the entire set of training variables is the top performer of all methods tested. The 1-day (2-day) random forest has $R^2 = 0.48$ ($R^2 = 0.25$) and gives $SS_{RMSE} = 0.27$ and $SS_{MAD} = 0.28$ ($SS_{RMSE} = 0.32$ and $SS_{MAD} = 0.33$) with respect to persistence (Figures 2k and 2m). The random forest also shows the strongest ability to forecast increases and decreases with little bias, with the F1 score being 0.38 and 0.48 (0.50 and 0.57), respectively (Figures 2i and 2n). The following results sections will focus exclusively on the random forest model, with Section 4.2 evaluating the random forest over space, time, and fire severity levels, 4.3 determining which predictors add the most skill to the model, and Section 4.4 and 4.5 discussing which subsets of the feature space lead to optimal model performance.

4.2. Evaluation of the Random Forest in Space, Across Fire Severity Levels, and in Time

We also evaluate the performance of the random forest by state, by two indicators of fire severity (FRE and FWI), and by day of the fire season. Across all states, using the RF model results in a more skilled or a very slightly less skilled FRE forecast than persistence (Figure 3a, Figure S6a in Supporting Information S1), as shown by the SS_{RMSE} and SS_{MAD} being positive or slightly negative. Generally speaking, the SS_{RMSE} and SS_{MAD} tend to be higher for the 2-day forecast than for the 1-day forecast (Figure 3a, Figure S6a in Supporting Information S1), except in UT and CO. More specifically, the RF model gives the largest improvements in CO for the 1-day forecast and the largest improvements in NV for the 2-day forecast. In WY, MAD and RMSE remain high on both days (Figures 3c and 3e, Figures S6c and S6e in Supporting Information S1). Using the RF also improves R^2 in most states in the 1- and 2-day forecasts. Except for NM (Figure 3b), all states show R^2 either changing sign or becoming more positive/less negative (Figure 3b, Figure S6b in Supporting Information S1). In the 1- and 2-day forecasts, the pattern of FRE error (RMSE, MAD) is driven mainly by FRE, with higher FRE states having higher error (Figures S7–S10 in Supporting Information S1, panels c–d). The remaining FRE error metrics (SS_{RMSE} , SS_{MAD} , and R^2) are driven by weather and moisture variables on day 1, with model skill reducing under more extreme fire weather conditions and increasing for more moist fuels (Figures S7–S8 panels a, b, e in Supporting Information S1). On Day 2, the SS_{RMSE} , SS_{MAD} , and R^2 spatial patterns are consistent with patterns in firefighting resources and moisture, with higher skill being achieved for fires in drier fuels with more resources applied (Figures S9 and S10 panels a, b, e in Supporting Information S1).

Across states, the RF maintains the ability to predict days with little change (Figure 3d, Figure S6d in Supporting Information S1) while also forecasting increases and decreases with skill (Figure 3f, Figure S6f in Supporting Information S1). In the 1- and 2-day forecasts, there is a slight decrease (never more than 0.2) in the F1_{slight change} score between the persistence and RF forecasts (Figure 3d, Figure S6d in Supporting Information S1). In certain states (e.g., MT, ID) increases are forecasted more skillfully than decreases and in other states (e.g., NV) decreases are forecasted more skillfully than increases. Generally speaking, F1 scores for increases and decreases are higher on day 2 than day 1 (Figure 3f, Figure S6f in Supporting Information S1). On day 1, the spatial pattern in the F1 scores is driven by fuel moisture, terrain, and weather variables (Figures S7 and S8 panels f–h in Supporting Information S1) and on day 2 the spatial pattern in the F1 scores is driven by fuel moisture, firefighting, and weather variables (Figures S9 and S10 panels f–h in Supporting Information S1).

Figures 3g–3j, Figures S6g–S6j in Supporting Information S1 shows error for the 1-day forecasts grouped into 20th-percentile bins of FWI (Figures 3g and 3i; Figures S6g and S6i in Supporting Information S1; a RF input) and FRE (Figures 3h and 3j, Figures S6g and S6i in Supporting Information S1; a RF output). In both the 1- and 2-day forecasts, the RF predictions tend toward persistence at extreme FWI (Figure 3g, Figure S6g in Supporting Information S1). Extreme FWI may correspond to a large range of fire behaviors (including cases of no fire, e.g. Field, 2020a, 2020b). Therefore, it is possible that the RF model cannot make these distinctions (see F1 scores in the Extreme class, Figure 3i, Figure S6i in Supporting Information S1). We also see error increasing with FRE in

both the 1- and 2-day forecasts (Figure 3h, Figure S6h in Supporting Information S1). However, for the 1-day forecast (Figure 3j) the SS_{MAD} and SS_{RMSE} tend to increase then plateau after the moderate FRE category, and for the 2-day forecast (Figure S6j in Supporting Information S1), the SS_{MAD} and SS_{RMSE} tend to increase as the FRE categories intensify. This implies that we are helping more with the upper end of FRE in the 2-day forecast and making improvements across most categories in the 1-day forecast. This is consistent with the spatial analysis, which showed that the model tends to under-perform for more extreme fires and more severe fire weather (Figures S7–S10 in Supporting Information S1).

Throughout the 2020 fire season, there are some days where the RF performs worse than persistence and some days where it performs better (Figures S11 and S12 in Supporting Information S1). The days when the RF is worse than persistence tend to be at the beginning and end of the fire season, but on these days the sample size is very low (<5 active fires per day). In spite of this, the RF beats persistence for a majority of the 1-day (2-day) forecasted days in the 2020 fire season, with SS_{RMSE} exceeding zero in 61% (63%) of days and SS_{MAD} exceeding zero in 67% (69%) of the 114 (113) days containing active fires. If the evaluation is restricted to days where there were at least 10 fires burning, SS_{RMSE} exceeds zero in 75% (73%) of days and SS_{MAD} exceeds zero in 80% (78%) of the 51 (48) days containing active fires.

4.3. Sensitivity Results

We performed sensitivity tests to evaluate each training variable's contribution to the overall forecast skill of the random forest (Figure 4). Many variables represent similar wildfire drivers (e.g., HWP, HD0W0, FWI, and BI all represent near surface meteorology) and so were grouped together (Figure 4, thick black lines and text). To test the random forest sensitivity to different feature groups, all variables in a certain feature group were dropped (Figure 4 black outlined points) as done in previous work (Huot et al., 2022). Additional sensitivities are performed where one variable in the group was used at a time, allowing for testing, re-training, and re-testing of the random forest model at each step (Figure 4, non-outlined points). This is slightly different from traditional random forest feature importance (Breiman, 2001), which is known to misrepresent the importance of correlated features (Nicodemus & Malley, 2009).

Figure 4 shows that groups of features generally fall into two tiers, one that includes feature groups that strongly influence the random forest, and one that includes the rest of the feature groups with less influence. The amount of influence a group of features exerts on the random forest is defined by the change in performance of the random forest when that group of features is excluded from the training data set (Figure 4, black outlined points) with respect to the performance of the random forest trained on all available features (Figure 4, vertical lines). For reference, the metrics for the full random forest are $R^2 = 0.48$, $SS_{MAD} = 0.28$, $SS_{RMSE} = 0.27$, $F1_{increase} = 0.38$, and $F1_{decrease} = 0.48$. The feature groups fall into the tiers as follows:

- Tier 1: This tier of features contributes a large amount of skill to the random forest across metrics. Two feature groups fall into this category, last day FRE and near-surface weather. Without these feature groups we see large drops in performance across all metrics ($R^2 = 0.30$ without surface weather and $R^2 = 0.45$ without last day FRE). The model trained including a single one of the simpler fire indices (HDW, HWP) are the top performers, sometimes outperforming the full RF in terms of F1 score and recovering much lost FRE skill when removing surface weather. Finally, the model trained using HWP as the only weather variable (Figure 4 features_only_hwp) indices has comparable R^2 , SS_{RMSE} , and SS_{MAD} to the model trained on all variables (Figure 4 features_all row).
- Tier 2: This tier of features includes those feature groups that moderately influenced the random forest model compared to Tier 1. Tier 2's feature groups include stability, living and dead fuel moisture, plant water sensitivity (PWS), evaporative stress index (ESI), terrain, fuel loading, and human variables. When feature groups in this tier are left out, performance metrics change less and may be better or worse with respect to the full random forest's performance metrics (e.g., $R^2 = 0.47$ – 0.48 , $F1_{increase} = 0.37$ – 0.40 , $F1_{decrease} = 0.41$ – 0.48).

Often, variables within the feature groups of Figure 4 are highly correlated. These correlations may be strongly positive, as in the case of different fire weather indices developed by different agencies capturing the same periods of favorable fire weather (Figure S13 in Supporting Information S1), fuel loading under different fire weather scenarios which is based on a fraction of the total biomass (Figure S14b in Supporting Information S1) or terrain variables where steeper slopes tend to occur in more variable elevations (Figure S14c in Supporting Information S1). Sometimes strong negative correlations exist between variables in a feature group. One such case is fuel

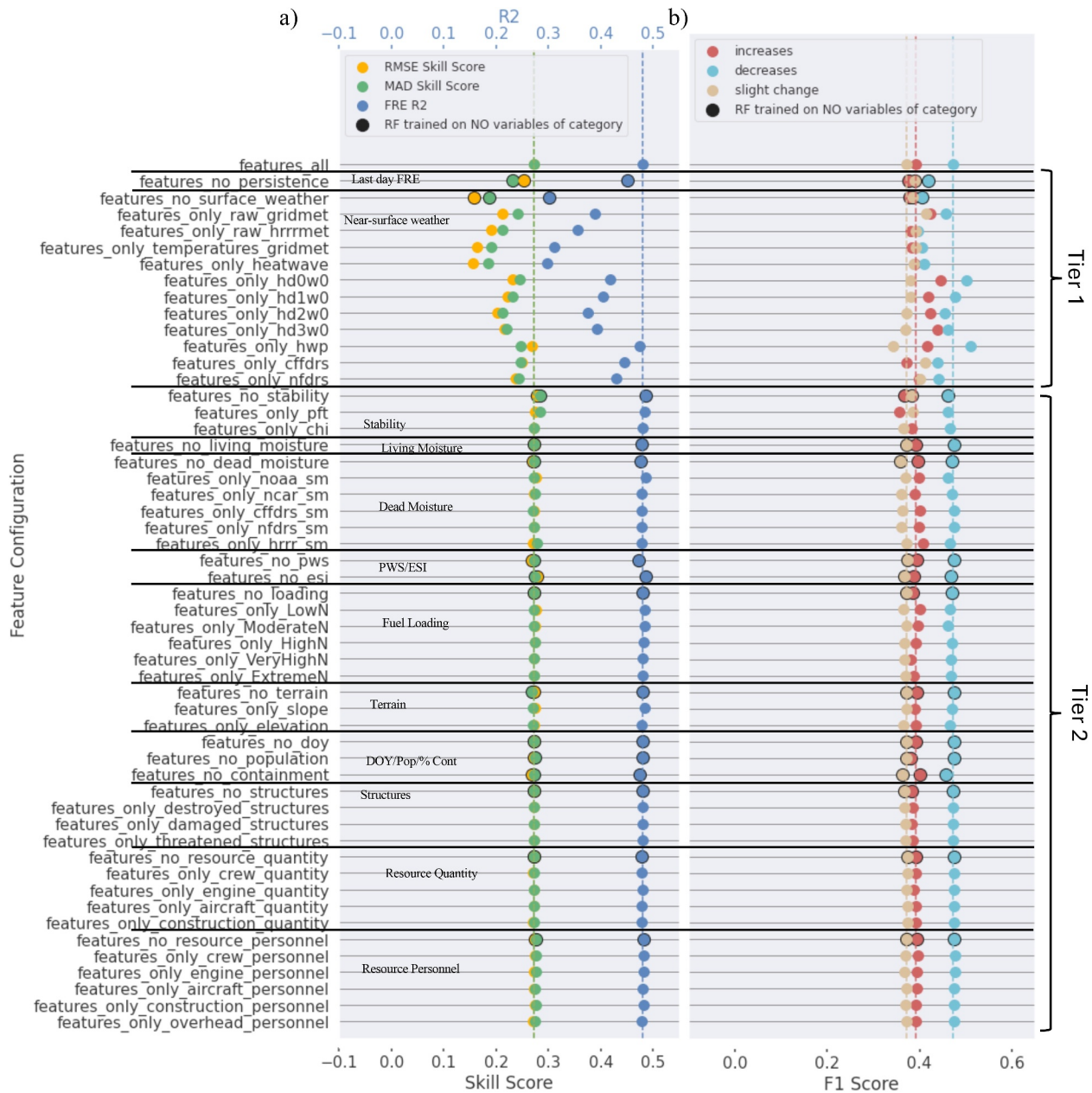


Figure 4. Performance of the random forest model for different sensitivity tests (row names). (a) shows SS_{RMSE} in yellow, SS_{MAD} in green, and R^2 in blue, and (b) shows $F1_{decrease}$ in light blue, $F1_{increase}$ in red, and $F1_{slight_change}$ in tan. Black horizontal lines delineate different groups of features for which sensitivity is tested, and these categories are also labeled in text to the left. Feature subsets for which no variables of the given category are used have their points outlined in black. The variable following "only" in the variable name is the sole variable of the given category used in training.

moisture, which is defined by some methodologies as codes which increase with increasing dryness (CFFDRS) and by others as fractional water content which decreases with increasing dryness (Figure S15 in Supporting Information S1). The high degree of correlations between variables within feature groups motivates training different versions of the random forest model subsets of the data to derive simpler models that could behave similarly to the model using all variables.

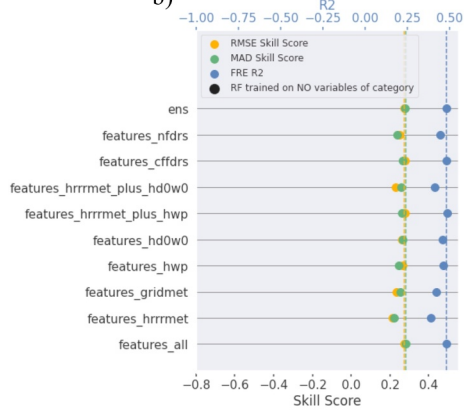
4.4. Selecting Optimal Subsets of Features

Random forest models were trained on subsets of the data which kept agency data (i.e., near-surface weather and dead fuel moisture from HRRR, GridMET, CFFDRS and NFDRS) separate where applicable, considered features

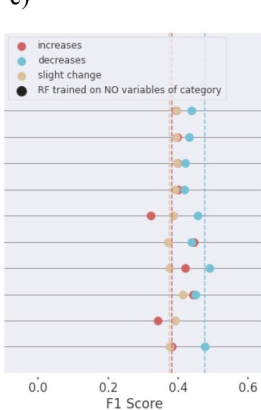
a)

Subset Name (based on source of meteorology)	US Fire Danger Rating System (features_nfrs)	Canadian Forest Fire Danger Rating System (features_cffdrs)	HRRR-Raw Meteorology + HWP (features_hrrrme t_plus_hwp)	HRRR-Raw Meteorology+ HD0W0 (features_hrrrme t_plus_hd0w0)	HRRR-HDW (features_hd0w0)	HRRR-HWP (features_hwp)	HRRR-Raw Meteorology (features_hrrrme t)	GridMET-Raw Meteorology (features_gridme t)
Persistence	FRE, one day old	FRE, one day old	FRE, one day old	FRE, one day old	FRE, one day old	FRE, one day old	FRE, one day old	FRE, one day old
Near-Surface Weather	NFDRS: BI, ERC	CFFDRS: FWI, BUI	HRRR: VPD, WS, HWP	HRRR: VPD, WS, HD0W0	HRRR: HWP	HRRR: HD0W0	HRRR: VPS, WS	
Stability	PFT	PFT	PFT	PFT	PFT	PFT	PFT	PFT
Dead Moisture	NFDRS: 100-hr, 1000-hr	CFFDRS: DC, DMC, FPMC	HRRR: 30cm moisture	HRRR: 30cm moisture	HRRR: 30cm moisture	HRRR: 30cm moisture	HRRR: 30cm moisture	NFDRS: 100-hr, 1000-hr
PWS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fuel Loading	Extreme	Extreme	Extreme	Extreme	Extreme	Extreme	Extreme	Extreme
Terrain	Elevation	Elevation	Elevation	Elevation	Elevation	Elevation	Elevation	Elevation
Human	Threatened structures, aircraft quantity	Threatened structures, aircraft quantity	Threatened structures, aircraft quantity	Threatened structures, aircraft quantity	Threatened structures, aircraft quantity	Threatened structures, aircraft quantity	Threatened structures, aircraft quantity	Threatened structures, aircraft quantity
Excluded Features	Living fuel moisture (NCAR), ESI, DOY, Containment, population	Living fuel moisture (NCAR), ESI, DOY, Containment, population	Living fuel moisture (NCAR), ESI, DOY, Containment, population	Living fuel moisture (NCAR), ESI, DOY, Containment, population	Living fuel moisture (NCAR), ESI, DOY, Containment, population	Living fuel moisture (NCAR), ESI, DOY, Containment, population	Living fuel moisture (NCAR), ESI, DOY, Containment, population	Living fuel moisture (NCAR), ESI, DOY, Containment, population

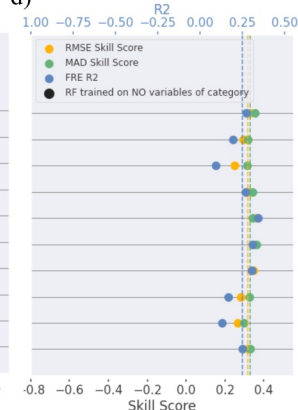
b)



c)



d)



e)

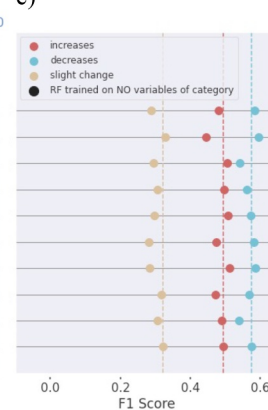


Figure 5. Performance of the random forest trained on subsets of the data. (a) shows subsets of features selected to optimize performance. R^2 and SS_{RMSE} and SS_{MAD} are shown for fire radiative energy for all 1-day (b) and 2-day (d) predictions in blue, yellow, and green dots respectively. F1 scores for all for all 1-day (c) and 2-day (e) predictions with increases in red, decreases in blue and slight changes in tan.

which can be operationally forecasted, and chose at most one feature or weather index system from each group to maximize FRE performance and F1 scores (Figure 4, non-outlined points). Figure 5a summarizes the subset of features that were tested. The subsets are as follows: “features_nfrs” is based on NFDRS near-surface weather and moisture indices; “features_cffdrs” is based on CFFDRS near-surface weather and moisture indices; “features_gridmet” is based on GridMET VPD and WS and NFDRS moisture indices; “features_hrrrmet” is based on HRRR VPD, WS, and soil moisture predictions; “features_hwp” is based on HRRR HWP and soil moisture predictions; and “features_hd0w0” is based on HRRR HD0W0 (Text S1.3 in Supporting Information S1) and soil moisture predictions. There are two additional subsets “features_hrrrmet_plus_hwp” and “features_hrrrmet_plus_hd0w0”, which are trained on HRRR VPD, WS, and HWP and HD0W0 respectively. These subsets were included to exploit potential nonlinear interactions between the raw (VPD, WS) and derived (HWP, HD0W0) weather variables that are available from HRRR. In this study we used the HRRR model as a reanalysis product and CFFDRS values from the GFWED version based on MERRA-2 (Text S1 in Supporting Information S1), but the HRRR and GEOS-5 models also produce forecasts of these same variables operationally. On the other

hand, NFDRS indices, VPD, and WS from GridMET (Text S1 in Supporting Information S1) are based on near real time observations only and do not have a forecasting component in GridMET. Therefore, any configuration based on GFWD or HRRR may be used on both forecast- and reanalysis-based meteorology to predict FRE.

Figure 5a also shows that the feature categories including population, containment percentage, day of year, NCAR live fuel moisture, and ESI were excluded from the subsets. These categories contained only one feature, and when this feature was dropped from the random forest, FRE scores increased slightly. Alternatively, persistence FRE and PWS are feature categories containing one variable which benefits model performance. When these variables were dropped, FRE and F1 skill decreased from $R^2 = 0.48$ in the full model to $R^2 = 0.45$ – 0.47 . Therefore, persistence FRE and PWS are important contributors to good model performance, while population, containment percentage, day of year, NCAR live fuel moisture, and ESI have little impact or slightly hinder model skill.

The variables listed for the stability (Figure S14a in Supporting Information S1), dead moisture (Figure S15 in Supporting Information S1), fuel loading (Figure S14b in Supporting Information S1), terrain (Figure S14c in Supporting Information S1), and human (Figure S16 in Supporting Information S1) categories in Figure 5a may co-vary within these categories and were all chosen to maximize FRE performance and minimize the impact of correlated features. When certain feature groups (i.e., dead fuel moisture, Figure 4, “features_no_dead_moisture”) are left out, FRE performance metrics do not change. We attribute this behavior to the random forest learning compensating errors between less correlated features (e.g., Figure S15 in Supporting Information S1, NCAR dead moisture NFDRS 1,000 hr moisture) within the feature groups or by interaction terms between feature groups.

4.5. Performance of Optimal Subsets of Features

Figures 5b–5d show the performance of random forest models which were trained using the subsets of features in Figure 5a. This analysis shows that subsets of the full feature set can be used to train random forest models and obtain performance comparable to or better than the random forest model trained on the full feature set (e.g., for the 1-day forecast $R^2 = 0.39$ – 0.48 , $SS_{RMSE} = 0.21$ – 0.28 , $SS_{MAD} = 0.22$ – 0.28 , $F1_{increase} = 0.39$ – 0.49 and $F1_{decrease} = 0.39$ – 0.49). Both the 1- and 2-day forecasts are similar to the SFI methods, in that RFs trained on more complex fire weather indices (Figures 5b and 5d “features_cffdrs”) have better skill at forecasting FRE than RFs trained on raw meteorological inputs (Figures 5b and 5d “features_hrrrmet”). Also like the SFI methods, the RFs trained on the simpler fire indices (Figures 5b and 5d “features_hwp”) show better performance than the RFs trained on more complex fire indices (Figures 5b and 5d “features_cffdrs”). Further, the models trained on derived indices (HWP, HD0W0) and raw weather variables (VPD, WS) perform better across most metrics than the models trained on either raw weather variables or derived indices (Figures 5b and 5d “features_hrrrmet”, “features_hwp”, “features_hrrrmet_plus_hwp”). An ensemble forecasting method, defined as the average of the predictions from the 8 RF models in 5a also shows comparable skill. All random forest-based models shown in Figure 5 have greater forecast skill than the scaled and persistence methods shown in Figure 3.

Overall, these results illustrate that the fire weather and fuel moisture indices developed by each agency independently represent important predictors of wildfire evolution and can be combined using machine learning methods to create forecasts that outperform persistence. Random forests trained on subsets of the predictors show similar behavior to both the ensemble method and the random forest trained on the full set of predictors, further highlighting the redundancy of including all available fire weather and fuel moisture indices in a single random forest model.

5. Discussion

5.1. Comparison With Previous Studies

The skill gained with respect to persistence is consistent with previous work which has beaten persistence. Graff et al. (2020) and D. Peterson et al. (2013) (G20 and P13 respectively) paired weather data with maximum likelihood classification and Poisson regression, respectively, to predict fire counts in Alaska. SS_{RMSE} ranging from 5% to 25% are reported in both papers, with the best performance occurring for decreases in fire count (P13) and 1- and 2-day forecasts with Poisson regression (G20). It is notable that, like this work, G20 also scales weather indices to forecast fire counts. In the 1-day forecast, the G20 scaled method is a less skillful forecast than persistence and the Poisson hurdle model, and in the 2-day forecast the G20 scaled method is more skillful than

persistence but less skillful than the poison hurdle model. The results of this study are also consistent with work by García-Llamas (2019), which explains up to 42% of variance in fire severity using random forest models trained on vegetation indices, rainfall, and time since last burn for Mediterranean pine forests in Spain. The most important variables reported in this study (i.e., near-surface fire weather) are also consistent with those reported in Gray et al. (2018).

This study can also be compared to Huot et al. (2022, H22) and Wang et al. (2022, W22), two studies which used machine learning to predict a next-day fire mask and monthly PM_{2.5}, respectively. H22 uses gridded fuel, weather, topography, and human data to produce gridded next-day predictions of fire location. Similar to our work, H22 finds that the previous day's fire location (persistence information) is key to giving a reliable forecast of next-day fire spread. Unlike our work, they have topography-related variables as another key feature, which lends credence to the idea that spatially resolved land surface features contain more useful information than area-averaged land surface features. In W22, the authors predict monthly PM_{2.5} from GFED (Kaiser et al., 2012) using fuel, weather, topography, and human variables. They achieve higher correlation with observations spatially than our work, and we attribute that to our daily data being noisier than their monthly data. In their feature importance ranking, weather is the most important predictor. We may conclude that our work is comparable to the machine learning models present in H22 and W22, and that most differences are attributable to methodological differences and different output variables. Between our work and the work of H22 and W22, a consensus seems to be emerging about the high importance of weather and persistence information for forecasting fire and emissions progression using machine learning.

5.2. Uncertainties and Limitations

Since the random forest method is applied as a post-processing step, uncertainties in the input data propagate through to the final model output. These uncertainties can be associated with measurements, such as gap-filling (Li et al., 2022) or missing (Schroeder et al., 2014) smoke- or cloud-obscured fire detections or the best estimates of resources applied to a fire event (Text S1 in Supporting Information S1). Additionally, other models were used to generate the training data. For instance, the operational HRRR model was used here, which has a longer record but did not include smoke radiative feedbacks until the end of 2020 (Ahmadov et al., 2017; Dowell et al., 2022a, 2022b). Additionally, model biases such as those for relative humidity in MERRA-2 which are used to calculate FWI and components (Field, 2020a, 2020b), and the outsize impact of temperature on some of the fuel moisture estimates (McCandless et al., 2020) could contribute to uncertainty in these variables and how they behave in the random forest. Finally, population density, PWS, fuel loading, and terrain slope and elevation were assumed to be static over the 2019–2021 period, which does not consider how either past burns or fuel treatments have altered the landscape. Future work to improve this modeling technique could consider adding georeferenced firefighting information, such as the location of firelines, as such information has proven useful in fire spread modeling (Turney et al., 2023).

Many variables used in this study are not available outside of the 2019–2021 period (Table 1), so the specific trained models evaluated may not be directly applicable to beating persistence for other fire seasons. However, our sensitivity analysis showed that the random forest can utilize wildfire drivers that were identified as far back as the 1970s (Countryman, 1972). We also show that subsets of the predictors can be used to beat persistence. This means that the model is likely learning real relationships between fire weather, fuels, topography, and firefighting, and that this technique is robust to the data sets chosen to represent different sides of the fire behavior triangle. Training the model on subsets of the data could allow future studies to investigate years beyond 2019–2021, because such models would not be limited by data sets which only exist for this limited period. Likely this would lead to improved model performance due to a training data set which samples the input space more fully. Finally, the configuration here requires a priori knowledge of fire location to extract predictors in the relevant location and generate future forecasts. This work was also limited by the fact that we had no multi-day predictions of fire location over which to extract predictors. This may explain why some features, such as fuel loading, moisture, population, and terrain did not show predictive power, since they may be spatially heterogeneous, and their variability is not captured through persisting polygon-averaged values throughout the forecast window. Therefore, we recommend for future study coupling the random forest technique with fire spread modeling (Bakhshaii & Johnson, 2019).

6. Conclusions

We described the development of data-driven FRE forecasting products that can be used as inputs for predicting wildfire emissions. We test the skill of several models at making forecasts of daily wildfire FRE using data extracted over the area affected by the wildfire. Scaled fire indices (SFI), and random forest (RF)-based model predictions were evaluated with respect to the persistence forecast. Overall, the RF showed the largest increase in skill in the 1-day (2-day) forecasts, with an $R^2 = 0.48$ ($R^2 = 0.25$) compared to 0.02 (-0.62) for persistence, $SS_{RMSE} = 0.27$ ($SS_{RMSE} = 0.32$), and $SS_{MAD} = 0.28$ ($SS_{MAD} = 0.33$) when all variables were used to train the model. The random forest was also the most skilled in forecasting whether day-to-day changes in fire behavior would occur, with $F1_{decrease} = 0.26$ ($F1_{decrease} = 0.44$) and $F1_{increase} = 0.48$ ($F1_{increase} = 0.57$). We find that the random forest reduces error compared to persistence across all states, for a majority of the days in Summer 2020, and for all severity categories. Error in both the random forest model and in the persistence model tends to increase as FRE increases. Additionally, the random forest model tends toward a persistence forecast at extreme FWI and for low-FRE fires.

The sensitivity analysis confirmed that most forecast skill came from including reanalysis-derived surface weather variables and persistence FRE as features. Other variables such as vertical atmospheric stability, fuel moisture, human influence, fuel loading, and PWS, seem to be most useful for fine tuning the predictions. Comparable skill was achieved by training several random forests on subsets of these variables based on data available to certain agencies. Therefore, there is significant value in using the random forest over a persistence or scaled fire index approach, and we recommend random forests be applied to scale emissions in future work.

We showed that machine learning plus knowledge of the human and natural drivers of wildfire can dramatically improve upon the current persistence assumption, with the biggest gains made via predicting increases or decreases in FRE. Our work also shows that the daily change in weather indices (HWP, HDW) can be very useful in predicting changes in FRE. FRE is proportional to emissions, so this method may be used to predict daily changes in FRP-based emissions inventories (Darmenov et al., 2015; Kaiser et al., 2012; Li et al., 2022). Many regional models (HRRR-Smoke, Ahmadov et al., 2017) use these FRP-based inventories, and so are poised to take advantage of the methodology presented here. While the methodology outlined here is most suitable for forecasting FRE and computing emissions offline, coupled chemical transport model-machine learning frameworks have been shown to work (e.g., Kelp et al., 2022) and are a promising future application of this work. Thus, data-driven derived approaches for scaling wildfire emissions should be considered in the next generation of air quality and atmospheric composition forecasting.

Data Availability Statement

VIIRS active fire detection data were acquired from NASA FIRMS (<https://firms.modaps.eosdis.nasa.gov/download/>). MTBS fire perimeter data were acquired from MTBS using the interactive viewer (Eidenshink et al., 2007b). RAVE FRE data comes from the Reprocessed RAVE Product acquired from SDSU GSCE (<https://sites.google.com/view/rave-emission/products?authuser=0>). GridMET data were acquired using wget with a script generated at the Gridmet site (Abatzoglou, 2013b). HRRR pressure level data were acquired using Amazon Web Services (Dowell et al., 2022b). GFWED calculations of the Canadian Fire Weather Indices were acquired from NASA NCCS via FTP (Field, 2015). Blended soil moisture observations were acquired from NOAA CLASS via data order (Liu et al., 2016b; requires login). Machine learning-derived living and dead fuel moisture estimates were acquired from NCAR/UCAR GDEX (Kosovic et al., 2019). Evaporative Stress Index maps were acquired via FTP from NOAA NIDIS (<https://gis1.servirglobal.net/data/esi/4WK/>). The PWS map was acquired via Krishna Rao's Github page (https://github.com/kkraoj/wildfire_from_lfmc). Fuel loading maps available at Thapa (2024). Terrain slope and elevation maps are available through LANDFIRE (https://www.landfire.gov/version_download.php). SIT-209 reports are acquired through the Wildland Fire Application Information Portal (Jamieson, 2005b). Note, Microsoft Access is required to use these files. Population density maps are acquired from NASA GPW v4 via (CIESIN-Columbia University, 2018b; requires login). The daily, polygon-averaged data (FRE, fuel, weather, topography, and fire-fighting) used to train and validate the random forest models and scaled weather index methods will be made available by time of publication via Zenodo (Thapa, 2024).

Acknowledgments

This work has been supported by the following Grants: NASA 80NSSC18K0629, NASA 80NSSC20K1650, NASA 80NSSC18K0685, NASA 80NSSC21K1456, NOAA NA18OAR4310107, NSF 2013461, NSF 2238338. This work was supported by funding from the Anthony and Jeanne Pritzker Family Foundation and the Eugene V. Cota-Robles Fellowship. L.H. T. would like to thank the Randerson Lab at UC Irvine for advice relating to heatwaves and other meteorological wildfire drivers. R.A. and E.J. thank NOAA's JPSS PGR program for funding and the rest of the HRRR-Smoke team and collaborators for helping with the model development. JB would like to gratefully acknowledge SubGrant 1559841 to the University of California, Los Angeles, from the University of Colorado Boulder under NASA Prime Grant agreement 80NSSC20K1580. The views expressed in this manuscript are those of the authors and do not reflect the views of NOAA or the Department of Commerce.

References

- Abatzoglou, J. T. (2013a). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, 33(1), 121–131. <https://doi.org/10.1002/joc.3413>
- Abatzoglou, J. T. (2013b). GridMET [Dataset]. *UC Merced Climatology Lab*. <https://www.climatologylab.org/wget-gridmet.html>
- Ahmadov, R., Grell, G., James, E., Csiszar, I., Tsidulko, M., Pierce, B., et al. (2017). Using VIIRS fire radiative power data to simulate biomass burning emissions, plume rise and smoke transport in a real-time air quality modeling system. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017-July, 2806–2808. <https://doi.org/10.1109/IGARSS.2017.8127581>
- Anderson, M. C., Hain, C., Wardlow, B., Pimstein, A., Mecikalski, J. R., & Kustas, W. P. (2011). Evaluation of Drought indices based on thermal Remote sensing of evapotranspiration over the continental United States. *Journal of Climate*, 24(8), 2025–2044. <https://doi.org/10.1175/2010JCLI3812.1>
- Arienti, M. C., Cumming, S. G., & Boutin, S. (2006). Empirical models of forest fire initial attack success probabilities: The effects of fuels, anthropogenic linear features, fire weather, and management. *Canadian Journal of Forest Research*, 36(12), 3155–3166. <https://doi.org/10.1139/x06-188>
- Bakhshaii, A., & Johnson, E. A. (2019). A review of a new generation of wildfire-atmosphere modeling. *Canadian Journal of Forest Research*, 49(6), 565–574. <https://doi.org/10.1139/cjfr-2018-0138>
- Berman, M. T., Ye, X., Thapa, L. H., Peterson, D. A., Hyer, E. J., Soja, A. J., et al. (2023). Quantifying burned area of wildfires in the Western United States from polar-orbiting and geostationary satellite active-fire detections. *International Journal of Wildland Fire*, 32(5), 665–678. <https://doi.org/10.1071/WF22022>
- Bradshaw, L. S., Deeming, J. E., Burgan, R. E., & compilers Cohen, J. D. (1984). *The 1978 national fire-danger rating system: Technical documentation. General technical report INT-169* (Vol. 44). U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 169. <https://doi.org/10.2737/INT-GTR-169>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J., & Wara, M. (2021). The changing risk and burden of wildfire in the United States. *Proceedings of the National Academy of Sciences*, 118(2), e2011048118. <https://doi.org/10.1073/pnas.2011048118>
- Carter, T. S., Heald, C. L., Jimenez, J. L., Campuzano-Jost, P., Kondo, Y., Moteki, N., et al. (2020). How emissions uncertainty influences the distribution and radiative impacts of smoke from fires in North America. *Atmospheric Chemistry and Physics*, 20(4), 2073–2097. <https://doi.org/10.5194/acp-20-2073-2020>
- Center for International Earth Science Information Network-CIESIN-Columbia University. (2018a). Documentation for the gridded population of the world, version 4 (GPWv4), revision 11 data sets. <https://doi.org/10.7927/H45Q4T5F>
- Center for International Earth Science Information Network—CIESIN—Columbia University. (2018b). Gridded Population of the World, version 4 (GPWv4): Population density, revision 11 [Dataset]. *NASA Socioeconomic Data and Applications Center (SEDAC)*. <https://doi.org/10.7927/H49C6VHW>
- Countryman, C. M. (1972). The fire environment concept. Vol. 15.
- Darmenov, A. S., Da Silva, A. M., & Koster, R. D. (2015). Technical report series on global modeling and data assimilation, volume 38 the Quick Fire Emissions Dataset (QFED): Documentation of versions 2.1, 2.2 and 2.4. <http://www.sti.nasa.gov>
- Di Giuseppe, F., Rémy, S., Pappenberger, F., & Wetterhall, F. (2017). Improving forecasts of biomass burning emissions with the fire weather index. *Journal of Applied Meteorology and Climatology*, 56(10), 2789–2799. <https://doi.org/10.1175/JAMC-D-16-0405.1>
- Di Giuseppe, F., Rémy, S., Pappenberger, F., & Wetterhall, F. (2018). Using the Fire Weather Index (FWI) to improve the estimation of fire emissions from Fire Radiative Power (FRP) observations. *Atmospheric Chemistry and Physics*, 18(8), 5359–5370. <https://doi.org/10.5194/acp-18-5359-2018>
- Dowell, D. C., Alexander, C. R., James, E. P., Weygandt, S. S., Benjamin, S. G., Manikin, G. S., et al. (2022a). The High-Resolution Rapid Refresh (HRRR): An hourly updating convection-allowing forecast model. Part I: Motivation and system description. *Weather and Forecasting*, 37(8), 1371–1395. <https://doi.org/10.1175/WAF-D-21-0151.1>
- Dowell, D. C., Alexander, C. R., James, E. P., Weygandt, S. S., Benjamin, S. G., Manikin, G. S., et al. (2022b). High Resolution Rapid Refresh (HRRR) model [Dataset]. *Amazon Web Services*. <https://rapidrefresh.noaa.gov/hrrr/>
- Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.-L., Quayle, B., & Howard, S. (2007a). A Project for monitoring Trends in burn severity. *Fire Ecology*, 3(1), 3–21. <https://doi.org/10.4996/fireecology.0301003>
- Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.-L., Quayle, B., & Howard, S. (2007b). A interactive viewer-monitoring Trends in burn severity: Burned area boundaries [Dataset]. *USDA Forest Service/US Geological Survey*. (Accessed on April 2023). Retrieved from <https://www.mtbs.gov/viewer/index.html>
- Field, R. D. (2015). Global Fire WEather Dabase (GFWED) [Dataset]. *NASA NCCS*. <https://portal.nccs.nasa.gov/datashare/GlobalFWI/v2.0/fwiCalcs.MERRA2/Default/IMERG.FINAL.v6>
- Field, R. D. (2020a). Evaluation of global fire weather Database reanalysis and short-term forecast products. *Natural Hazards and Earth System Sciences*, 20(4), 1123–1147. <https://doi.org/10.5194/NHESS-20-1123-2020>
- Field, R. D. (2020b). Using satellite estimates of precipitation for fire danger rating. In V. Levizzani, C. Kidd, D. B. Kirschbaum, C. D. Kummerow, K. Nakamura, & F. J. Turk (Eds.), *Satellite precipitation measurement* (Vol. 2, pp. 1131–1154). Springer International Publishing. https://doi.org/10.1007/978-3-030-35798-6_33
- Field, R. D., Spessa, A. C., Aziz, N. A., Camia, A., Cantin, A., Carr, R., et al. (2015). Development of a global fire weather Database. *Natural Hazards and Earth System Sciences*, 15(6), 1407–1423. <https://doi.org/10.5194/nheiss-15-1407-2015>
- García-Llamas, P., Suárez-Seoane, S., Taboada, A., Fernández-Manso, A., Quintano, C., Fernández-García, V., et al. (2019). Environmental drivers of fire severity in extreme fire events that affect Mediterranean pine forest ecosystems. *Forest Ecology and Management*, 433, 24–32. <https://doi.org/10.1016/j.foreco.2018.10.051>
- Gillies, S., van der Wel, C., Van den Bossche, J., Taves, M. W., Arnott, J., Ward, B. C., et al. (2007). Shapely: (Version 2.0.6) [Computer software]. <https://doi.org/10.5281/zenodo.5597138>
- Graff, C. A., Coffield, S. R., Chen, Y., Fofoula-Georgiou, E., Randerson, J. T., & Smyth, P. (2020). Forecasting daily wildfire activity using Poisson regression. *IEEE Transactions on Geoscience and Remote Sensing*, 58(7), 4837–4851. <https://doi.org/10.1109/TGRS.2020.2968029>
- Gray, M. E., Zachmann, L. J., & Dickson, B. G. (2018). A weekly, continually updated dataset of the probability of large wildfires across Western US forests and woodlands. *Earth System Science Data*, 10(3), 1715–1727. <https://doi.org/10.5194/essd-10-1715-2018>
- Hernandez, C., Keribin, C., Drobinski, P., & Turquety, S. (2015). Statistical modelling of wildfire size and intensity: A step toward meteorological forecasting of summer extreme fire risk. *Annales Geophysicae*, 33(12), 1495–1506. <https://doi.org/10.5194/angeo-33-1495-2015>

- Hu, X., & Ntairo, L. (2009). Integrated simulation and optimization for wildfire containment. *ACM Transactions on Modeling and Computer Simulation*, 19(4), 1–29. <https://doi.org/10.1145/1596519.1596524>
- Huot, F., Hu, R. L., Goyal, N., Sankar, T., Ihme, M., & Chen, Y.-F. (2022). Next day wildfire spread: A machine learning dataset to predict wildfire spreading from remote-sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–13. <https://doi.org/10.1109/TGRS.2022.3192974>
- Ichoku, C., Martins, J. V., Kaufman, Y. J., Wooster, M. J., Freeborn, P. H., Hao, W. M., et al. (2008). Laboratory investigation of fire radiative energy and smoke aerosol emissions. *Journal of Geophysical Research*, 113(D14), D14S09. <https://doi.org/10.1029/2007JD009659>
- Jamieson, G. (2005a). NIMS and the incident Command system. In *International Oil Spill conference* (pp. 291–294).
- Jamieson, G. (2005b). The SIT-209 Database [Dataset]. *Wildland Fire Application Information Portal*. (Accessed on 2022). Retrieved from <https://www.wildfire.gov/application/sit209>
- Jang, E., Kang, Y., Im, J., Lee, D.-W., Yoon, J., & Kim, S.-K. (2019). Detection and monitoring of forest fires using Himawari-8 geostationary satellite data in South Korea. *Remote Sensing*, 11(3), 271. <https://doi.org/10.3390/rs11030271>
- Jolly, W. M., Freeborn, P. H., Page, W. G., & Butler, B. W. (2019). Severe fire danger index: A forecastable metric to inform firefighter and community wildfire risk management. *Fire*, 2(3), 47. <https://doi.org/10.3390/fire2030047>
- Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., et al. (2012). Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, 9(1), 527–554. <https://doi.org/10.5194/bg-9-527-2012>
- Kelp, M. M., Jacob, D. J., Lin, H., & Sulprizio, M. P. (2022). An online-learning neural network chemical solver for stable long-term global simulations of atmospheric chemistry. *Journal of Advances in Modeling Earth Systems*, 14(6), e2021MS002926. <https://doi.org/10.1029/2021MS002926>
- Kosovic, B., Massie, S., McCandless, T., Petzke, B., Jimenez Munoz, P., DeCastro, A., & Siems-Anderson, A. (2019). Fuel moisture content (live and dead) over the conterminous United States [Dataset]. *Version 1.0. UCAR/NCAR—GDEX*. <https://doi.org/10.5065/qt42-zd40>
- Krueger, E. S., Levi, M. R., Achieng, K. O., Bolten, J. D., Carlson, J. D., Coops, N. C., et al. (2022). Using soil moisture information to better understand and predict wildfire danger: A review of recent developments and outstanding questions. *International Journal of Wildland Fire*, 32(2), 111–132. <https://doi.org/10.1071/WF22056>
- Li, F., Zhang, X., Kondragunta, S., Lu, X., Csizsar, I., & Schmidt, C. C. (2022). Hourly biomass burning emissions product from blended geostationary and polar-orbiting satellites for air quality forecasting applications. *Remote Sensing of Environment*, 281, 113237. <https://doi.org/10.1016/j.rse.2022.113237>
- Liu, J., Zhan, X., Hain, C., Yin, J., Fang, L., Li, Z., & Zhao, L. (2016a). NOAA Soil Moisture Operational Product System (SMOPS) and its validations. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 3477–3480. <https://doi.org/10.1109/IGARSS.2016.7729899>
- Liu, J., Zhan, X., Hain, C., Yin, J., Fang, L., Li, Z., & Zhao, L. (2016b). Soil moisture operational product system (SMOPS) [Dataset]. *NOAA CLASS*. https://www.aev.class.noaa.gov/saa/products/search?sub_id=0&datatype_family=SMOPS&submit.x=30&submit.y=10
- McCandless, T. C., Kosovic, B., & Petzke, W. (2020). Enhancing wildfire spread modelling by building a gridded fuel moisture content product with machine learning. *Machine Learning: Science and Technology*, 1(3), 035010. <https://doi.org/10.1088/2632-2153/aba480>
- McClure, C. D., Pavlovic, N. R., Huang, S., Chaveste, M., & Wang, N. (2023). Consistent, high-accuracy mapping of daily and sub-daily wildfire growth with satellite observations. *International Journal of Wildland Fire*, 32(5), 694–708. <https://doi.org/10.1071/WF22048>
- Munoz-Alpizar, R., Pavlovic, R., Moran, M., Chen, J., Gravel, S., Henderson, S., et al. (2017). Multi-year (2013–2016) PM2.5 wildfire pollution exposure over North America as determined from operational air quality forecasts. *Atmosphere*, 8(12), 179. <https://doi.org/10.3390/atmos8090179>
- Nicodemus, K. K., & Malley, J. D. (2009). Predictor correlation impacts machine learning algorithms: Implications for genomic studies. *Bioinformatics*, 25(15), 1884–1890. <https://doi.org/10.1093/bioinformatics/btp331>
- Partanen, T. M., & Sofiev, M. (2022). Forecasting the regional fire radiative power for regularly ignited vegetation fires. *Natural Hazards and Earth System Sciences*, 22(4), 1335–1346. <https://doi.org/10.5194/nhess-22-1335-2022>
- Peterson, D., Hyer, E., & Wang, J. (2013). A short-term predictor of satellite-observed fire activity in the North American boreal forest: Toward improving the prediction of smoke emissions. *Atmospheric Environment*, 71, 304–310. <https://doi.org/10.1016/j.atmosenv.2013.01.052>
- Peterson, D. A., Thapa, L. H., Saide, P. E., Soja, A. J., Gargulinski, E. M., Hyer, E. J., et al. (2022). Measurements from inside a Thunderstorm driven by wildfire: The 2019 FIREX-AQ field experiment. *Bulletin of the American Meteorological Society*, 103(9), E2140–E2167. <https://doi.org/10.1175/BAMS-D-21-0049.1>
- Pinto, M. M., DaCamara, C. C., Hurdac, A., Trigo, R. M., & Trigo, I. F. (2020). Enhancing the fire weather index with atmospheric instability information. *Environmental Research Letters*, 15(9), 0940b7. <https://doi.org/10.1088/1748-9326/ab9e22>
- Powers, D. (2008). Evaluation: From precision, recall and F-factor to ROC, informedness, markedness and correlation. *Mach. Learn. Technol.*, 2.
- Preisler, H. K., & Westerling, A. L. (2007). Statistical model for forecasting monthly large wildfire events in Western United States. *Journal of Applied Meteorology and Climatology*, 46(7), 1020–1030. <https://doi.org/10.1175/JAM2513.1>
- Rao, K., Williams, A. P., Diffenbaugh, N. S., Yebra, M., & Konings, A. G. (2022). Plant-water sensitivity regulates wildfire vulnerability. *Nature Ecology & Evolution*, 6(3), 332–339. <https://doi.org/10.1038/s41559-021-01654-2>
- Rodrigues, M., Alcasena, F., Gelabert, P., & Vega-García, C. (2020). Geospatial modeling of containment probability for escaped wildfires in a Mediterranean region. *Risk Analysis*, 40(9), 1762–1779. <https://doi.org/10.1111/risa.13524>
- Rothermel, R. C. (1972). A mathematical model for predicting fire spread in wildland fuels. *INT*, 115, 40.
- Rothermel, R. C. (1991). *Predicting behavior and size of crown fires in the northern rocky mountains*. U.S. Department of Agriculture, Forest Service, Intermountain Research Station.
- Saide, P. E., Peterson, D. A., da Silva, A., Anderson, B., Ziemba, L. D., Diskin, G., et al. (2015). Revealing important nocturnal and day-to-day variations in fire smoke emissions through a multiplatform inversion. *Geophysical Research Letters*, 42(9), 3609–3618. <https://doi.org/10.1002/2015GL063737>
- Sazib, N., Bolten, J. D., & Mladenova, I. E. (2022). Leveraging NASA soil moisture active passive for assessing fire susceptibility and potential impacts over Australia and California. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 779–787. <https://doi.org/10.1109/JSTARS.2021.3136756>
- Schroeder, W., Oliva, P., Giglio, L., & Csizsar, I. A. (2014). The New VIIRS 375m active fire detection data product: Algorithm description and initial assessment. *Remote Sensing of Environment*, 143, 85–96. <https://doi.org/10.1016/j.rse.2013.12.008>
- Seiler, W., & Crutzen, P. J. (1980). Estimates of gross and net fluxes of carbon between the biosphere and the atmosphere from biomass burning. *Climatic Change*, 2(3), 207–247. <https://doi.org/10.1007/BF00137988>
- Soja, A. J., Cofer, W. R., Shugart, H. H., Sukhinin, A. I., Stackhouse, P. W., Jr., McRae, D. J., & Conard, S. G. (2004). Estimating fire emissions and disparities in boreal Siberia (1998–2002). *Journal of Geophysical Research*, 109(D14), D14S06. <https://doi.org/10.1029/2004JD004570>

- Srock, A. F., Charney, J. J., Potter, B. E., & Goodrick, S. L. (2018). The hot-dry-windy index: A new fire weather index. *Atmosphere*, 9(7), 279. <https://doi.org/10.3390/atmos9070279>
- Thapa, L. H. (2024). Forecasting daily fire radiative energy using data driven methods and machine learning [Dataset, Software]. *Zenodo*. <https://doi.org/10.5281/zenodo.10231992>
- Thapa, L. H., Ye, X., Hair, J. W., Fenn, M. A., Shingler, T., Kondragunta, S., et al. (2022). Heat flux assumptions contribute to overestimation of wildfire smoke injection into the free troposphere. *Communications Earth & Environment*, 3(1), 236. <https://doi.org/10.1038/s43247-022-00563-x>
- Turney, F. A., Saide, P. E., Jimenez Munoz, P. A., Muñoz-Esparza, D., Hyer, E. J., Peterson, D. A., et al. (2023). Sensitivity of burned area and fire radiative power predictions to containment efforts, fuel density, and fuel moisture using WRF-fire. *Journal of Geophysical Research: Atmospheres*, 128(18), e2023JD038873. <https://doi.org/10.1029/2023JD038873>
- Viegas, D. X. (2004). Slope and wind effects on fire propagation. *International Journal of Wildland Fire*, 13(2), 143. <https://doi.org/10.1071/WF03046>
- Wagner, C. E. V. (1987). Development and structure of the Canadian forest fire weather index system. *Minister of Supply and Services Canada*.
- Wang, S. S.-C., Qian, Y., Leung, L. R., & Zhang, Y. (2022). Interpreting Machine Learning prediction of fire emissions and comparison with FireMIP process-based models. *Atmospheric Chemistry and Physics*, 22(5), 3445–3468. <https://doi.org/10.5194/acp-22-3445-2022>
- Westerling, A. L., Hidalgo, H. G., Cayan, D. R., & Swetnam, T. W. (2006). Warming and earlier spring increase Western U.S. forest wildfire activity. *Science*, 313(5789), 940–943. <https://doi.org/10.1126/science.1128834>
- Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., & Soja, A. J. (2011). The Fire INventory from NCAR (FINN): A high resolution global model to estimate the emissions from open burning. *Geoscientific Model Development*, 4(3), 625–641. <https://doi.org/10.5194/gmd-4-625-2011>
- Willmott, C., & Matsuura, K. (2005). Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in assessing average model performance. *Climate Research*, 30, 79–82. <https://doi.org/10.3354/cr030079>
- Wooster, M. J., Roberts, G., Perry, G. L. W., & Kaufman, Y. J. (2005). Retrieval of biomass combustion rates and totals from fire radiative power observations: FRP derivation and calibration relationships between biomass consumption and fire radiative energy release. *Journal of Geophysical Research*, 110(D24), D24311. <https://doi.org/10.1029/2005JD006318>
- Ye, X., Arab, P., Ahmadov, R., James, E., Grell, G. A., Pierce, B., et al. (2021). Evaluation and intercomparison of wildfire smoke forecasts from multiple modeling systems for the 2019 Williams Flats fire. *Atmospheric Chemistry and Physics*, 21(18), 14427–14469. <https://doi.org/10.5194/acp-21-14427-2021>
- Young, J. D., Thode, A. E., Huang, C.-H., Ager, A. A., & Fulé, P. Z. (2019). Strategic application of wildland fire suppression in the Southwestern United States. *Journal of Environmental Management*, 245, 504–518. <https://doi.org/10.1016/j.jenvman.2019.01.003>
- Zhuang, Y., Fu, R., Santer, B. D., Dickinson, R. E., & Hall, A. (2021). Quantifying contributions of natural variability and anthropogenic forcings on increased fire weather risk over the Western United States. *Proceedings of the National Academy of Sciences*, 118(45), e2111875118. <https://doi.org/10.1073/pnas.2111875118>

References From the Supporting Information

- Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., et al. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, 28(15), 2031–2064. <https://doi.org/10.1002/joc.1688>
- Dowdy, A. J., & Pepler, A. (2018). Pyroconvection risk in Australia: Climatological changes in atmospheric stability and surface fire weather conditions. *Geophysical Research Letters*, 45(4), 2005–2013. <https://doi.org/10.1002/2017GL076654>
- Jenkins, M. A. (2002). An examination of the sensitivity of numerically simulated wildfires to low-level atmospheric stability and moisture, and the consequences for the Haines Index. *International Journal of Wildland Fire*, 11(4), 213–232. <https://doi.org/10.1071/wf02006>
- LANDFIRE. (2016). Slope degrees. Retrieved from https://www.engineeringtoolbox.com/slope-degrees-gradient-grade-d_1562.html
- Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., et al. (2004). The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research*, 109(D7), D07S90. <https://doi.org/10.1029/2003JD003823>
- Ottmar, R., Sandberg, D., Riccardi, C., & Prichard, S. (2007). An overview of the fuel characteristic classification system—Quantifying, classifying, and creating fuelbeds for resource planning This article is one of a selection of papers published in the special forum on the fuel characteristic classification system. *Canadian Journal of Forest Research—Revue Canadienne De Recherche Forestiere—CAN J FOREST RES*, 37(12), 2383–2393. <https://doi.org/10.1139/X07-077>
- Perkins, S. E., Alexander, L. V., & Nairn, J. R. (2012). Increasing frequency, intensity and duration of observed global heatwaves and warm spells. *Geophysical Research Letters*, 39(20), L20714. <https://doi.org/10.1029/2012GL053361>
- Potter, B. (2018). The Haines index – It's time to revise it or replace it. *International Journal of Wildland Fire*, 27(7), 437. <https://doi.org/10.1071/WF18015>
- Potter, B. E. (2005). The role of released moisture in the atmospheric dynamics associated with wildland fires. *International Journal of Wildland Fire*, 14(1), 77–84. <https://doi.org/10.1071/WF04045>
- Schmidt, C. C., Ssec, U.-M., Hoffman, J., Ssec, U.-M., Prins, E., Ssec, U.-M., et al. (2012). GOES-R advanced baseline imager (ABI) algorithm theoretical basis document for fire/hot spot characterization.
- Tory, K. J., & Kepert, J. D. (2021). Pyrocumulonimbus firepower threshold: Assessing the atmospheric potential for pyroCb. *Weather and Forecasting*, 36(2), 439–456. <https://doi.org/10.1175/WAF-D-20-0027.1>
- Yin, J., Zhan, X., & Liu, J. (2020). NOAA Satellite Soil Moisture Operational Product System (SMOPS) version 3.0 generates higher accuracy blended satellite soil moisture. *Remote Sensing*, 12(17), 2861. <https://doi.org/10.3390/rs12172861>
- Warneke, C., Schwarz, J. P., Dibb, J., Kalashnikova, O., Frost, G., Al-Saad, J., et al. (2023). Fire Influence on Regional to global Environments and Air Quality (FIREX-AQ). *Journal of Geophysical Research: Atmospheres*, 128(2), e2022JD037758. <https://doi.org/10.1029/2022JD037758>