Enhancing wildfire spread modelling by building a gridded fuel moisture content product with machine learning

IOP Publishing Mach. Learn.: Sci. Technol. 1 (2020) 035010 https://doi.org/10.1088/2632-2153/aba480

MACHINE LEARNING
Science and Technology

OPEN ACCESS

RECEIVED
20 March 2020

REVISED
3 July 2020

ACCEPTED FOR PUBLICATION
9 July 2020

PUBLISHED
18 August 2020

PAPER

Enhancing wildfire spread modelling by building a gridded fuel moisture content product with machine learning

Tyler C McCandless[1], Branko Kosovic[2] and William Petzke[2]

[1] E Source, Boulder, CO 80301, United States of America
[2] National Center for Atmospheric Research, Research Applications Laboratory, Boulder, CO, United States of America

E-mail: tyler_mccandless@esource.com

Keywords: random forest, artificial neural network, gradient boosted regression, machine learning interpretability, wildfire modeling, satellite data

Abstract

Wildland fire decision support systems require accurate predictions of wildland fire spread. Fuel moisture content (FMC) is one of the important parameters controlling the rate of spread of wildland fire. However, dead FMC measurements are provided by a relatively sparse network of remote automatic weather stations (RAWS), while live FMC is relatively infrequently measured manually. We developed a high resolution, gridded, real-time FMC data sets that did not previously exist for assimilation into operational wildland fire prediction systems based on ML. We used surface observations of live and dead FMC to train machine learning models to estimate FMC based on satellite observations. Moderate Resolution Imaging Spectrometer Terra and Aqua reflectances are used to predict the live and dead FMC measured by the Wildland Fire Assessment System and RAWS). We evaluate multiple machine learning methods including multiple linear regression, random forests (RFs), gradient boosted regression and artificial neural networks. The models are trained to learn the relationships between the satellite reflectances, surface weather and soil moisture observations and FMC. After training on data corresponding to the temporally and spatially nearest grid points to the irregularly spaced surface FMC observations, the machine learning models could be applied to all grid cells for a gridded product over the Conterminous United States (CONUS). The results show generally that the rule-based approaches have the lowest errors likely due to the sharp decision boundaries among the predictors, and the RF approach that utilizes bagging to avoid over-fitting has the lowest error on the test dataset. The errors are typically between 25%−33% the typical variability of the FMC data, which indicate the skill of the RF in estimating the FMC based on satellite data and surface characteristics. The FMC gridded product based on the RF runs operationally daily over CONUS and can be assimilated into WRF-Fire for more accurate wildland fire spread predictions.

1. Introduction

Wildland fire modeling is a growing area of research and development in recent years due to the multitude of historical fires in many countries including Brazil, Greece, Australia and United States. Atmospheric conditions, fuel type and load, as well as fuel moisture content are all important factors affecting the rate of spread from wildland fires. In this study, we focus on improving estimates of fuel moisture content for integration into a wildland fire spread model, specifically the Weather Research and Forecast (WRF)-Fire model [1].

The WRF-Fire model attempts to achieve an optimal balance between physical processes that control the wildfire fire spread and the interaction with the atmosphere with the need for real-time simulations, which is an advancement to the WRF model [2]. WRF-Fire [1] is a coupled wildland fire spread that advances standard numerical weather prediction (NWP) inputs by assimilating fuel data and includes parameterizations of both biomass burning processes and rate of fire spread. Previous work has shown the

sensitivity of the WRF-Fire model to the fuel moisture content. Fuel moisture content is typically two separate measurements, the dead and live fuel moisture contents (LFMCs), which are combined to derive the fuel moisture content used in the Rothermel fire spread model [3] used within WRF-Fire with the level set approach to tracking the fire front [4].

Currently, in the United States the National Fuel Moisture Database [5], which is available via the Wildland Fire Assessment System (WFAS), provides information about FMC based on interpolation of sparse manual sampling of LFMC and surface observations from the Remote Automated Weather Stations (RAWS). These interpolations can produce very large errors due to the fact that each RAWS and manual sample may be located in an area of different geography, fuel type or climatology, which is especially true in mountainous areas of the western United States.

One way to improve the resolution of estimates is by utilizing information provided by remote sensing, or satellite-based observations. In this study we use remote sensing data along with surface data and estimates from NWP models to build a predictive relationship to the fuel moisture content in order to better estimate the fuel moisture content in areas that may be far from the nearest observation site. Other studies have utilized machine learning on remote sensing data to improved gridded representation of surface or subsurface data. In [6], a random forest (RF) and gradient boosted are used to improve a version of the SoilGrids system at 250 m resolution. A genetic programming models was used in [7] to derive spatiotemporal patterns of phosphorus concentrations in a coastal bay with MODIS data. In [8], support vector machines were used to generate soil moisture estimates using remote sensing data for the Lower Colorado River Basin in the Western United States. A machine learning approach to predict drought conditions at high resolution in ungauged areas based on both climate forecasts and remote sesnsing data was used in [9]. Artificial neural networks (ANNs) and multiple linear regression were used in [10] to predict moisture content indices for a temperature humid forecast in northesatern Iran based on MODIS data.

This paper is organized as follows: section 2 describes the datasets, including the fuel moisture content, satellite data, WRF-Hydro/National Water Model (NWM) and derived predictors. Section 3 describes the system methodology including training, testing and validation and an explanation of the machine learning models. Section 4 presents the results of the machine learning models compared to multiple linear regression. The final section 5 discusses the results, interpretability and describes the operational fuel moisture content system, including where the data and maps are publicaly available.

2. Data

There are three different datasets that are merged together in this study. The predictands are the LFMC and the dead fuel moisture content (DFMC) which come from surface observations. The satellite data is from the polar orbiting MODIS Aqua and Terra satellites. Surface characteristics and observations come from the WRF-Hydro/NWM [11, 12]. These datasets are described in more detail in sub-sections 2.1-2.3. Additional derived variables are described in sub-section 2.4.

2.1. Fuel moisture content

In situ observations of DFMC are provided by RAWS locations, which are available at over 2000 sites over the conterminous United States (CONUS) [13]. The DFMC is provided for non-living fuels of different class sizes based on how long it typically takes that type of material to dry; however, in this study we evaluate 10-h DFMC. The DFMC observations are provided hourly by automated surface observations stations. The LFMC observations are measured for living fuels on a species by species basis such as pine, oak, aspen, etc. These observations are collected infrequently, every few days and in general only during a fire season and can be accessed through daily by the US Wildland WFAS. The LFMC is determined by weighing the materials before and after daying them in an oven [14]. We use the LFMC as a universal measure of all living fuels dryness. It is important to note that LFMC observations can be much greater than 100% when there is more water weight to a fuel than the plant material.

2.2. Satellite data

The satellite data used in this study are from the Moderate Resolution Imaging Spectroradiometer (MODIS), which is an instrument on board the NASA Aqua and Terra satellites. Terra passes from the north to the south during the morning hours (typically around 10:30AM local time in the USA) while Aqua passes from south to north over the equater in the afternoon (typically around 1:30PM local time in the USA). Our analysis reads in the data at 1-km resolution over the CONUS for surface reflectance (MOD09, MYD09), MODIS cloud mask (MOD35, MYD35), and land surface temperature (MOD11, MYD11). For training, the surface reflectance bands (1–7) are included if data is available for a given grid cell and if there are clouds present, then the previous day's reflectances are persistented in a real-time forecasting environment. More

Table 1. List of predictors by data source.

MODIS	WRF-Hydro	Derived
Band 1	Soil Saturation	X-Slope
Band 2	Accumulated Evapotranspiration	Y-Slope
Band 3	Land Use Category	Northwest
Band 4	Elevation	Southwest
Band 5		East
Band 6		
Band 7		
Land Surface Temperature		

details about the surface reflectance products can be found in [15, 16]. The surface reflectance values assume no atmospheric scattering or absorption and atmospheric corrections are applied to the MODIS radiances to produce surface reflectance fields in seven bands: Band 1 (620–670 nm), 2 (841–876 nm), 3 (459–479), 4 (545–565 nm), 5 (1230–1250 nm), 6 (1628–1652 nm), and 7 (2105–2155 nm). The bands provide insight into the representation of the surface with the GOES channels representing differences in soil or vegetaton, green or brown vegetation, leaf or canopy differences, snow or land properties. Specifically, bands 1, 4 and 3 represent red, green and blue channels that can help identify both the type of surface (land vs water) and the dryness (green or wet vs brown or dry). In addition to the surface reflectances we also utilize the surface temperature data product to be used as a predictor in the machine learning models [17]. The land surface temperature can add predictive information on seasonality, the rate of drying, and climate or agricultural zones.

2.3. WRF-Hydro/National Water Model (NWM)

In order to capture the impact of soil saturation and plant dynamics on the FMC estimation, we utilize inputs into NWM Implementation of WRF-Hydro [18]. The NWM became operational in August 2016 to address the needs and applications of the water management community. In this study we utilize three predictor variables from the NWM: accumulated total evapotranspiration, soil moisture content and land use category. Evapotranspiration measures the amount of water transferred from the land to the atmosphere via evaporation and transpiration from plants. The volumetric soil moisture content (i.e. soil saturation) is averaged over all layers to the depth of 2 m and quantifies the amount of water in the soil, which provides additional information on the ability of a plant to maintain a fuel moisture content in relatively dry atmospheric conditions. The elevation used by WRF and the land use category, which is from the United States Geological Survey (USGS) Landuse product as input into WRF, are used as predictors [19].

2.4. Derived predictors

The final predictors used as inputs to the machine learning model were derived to capture the affect of surface characteristics. In addition to using the elevation as a predictor, the slope over a one-kilometer grid cell in the meriodional (y) and zonal (x) directions are computed and used as predictors. The slope can help characterize the effect of solar radiation and the rate of drying for living and non-living plant material. The dataset also included predictors for regions that were one-hot encoded to capture the predictability of the region in the US that the data was in: Northwest, Southwest, or East as shown in figure 1. The final list of predictors is detailed in table 1.

After joining the data together, we performed a correlation heatmap analysis to determine the collinearity among the predictors and with the predictands (LFMC, figure 2; DFMC, figure 3). Here, we can see that generally the reflectance bands are positively correlated (blue) with each other, as expected, while most other predictors have correlations close to zero. The LFMC has mostly low correlations with each of the predictors, which is also true for the DFMC with each of the predictors. The strongest negative correlation is with the elevation predictor that makes physical sense as we would expect generally drier conditions at higher altitudes. It is also evident that the reflectance bands have a slightly negative correlation with the DFMC but near zero to slightly positive correlation with the LFMC. Note that the land use category is a categorical variable, which makes interpretability of the pearson correlation coefficient more difficult than a continuous variable.

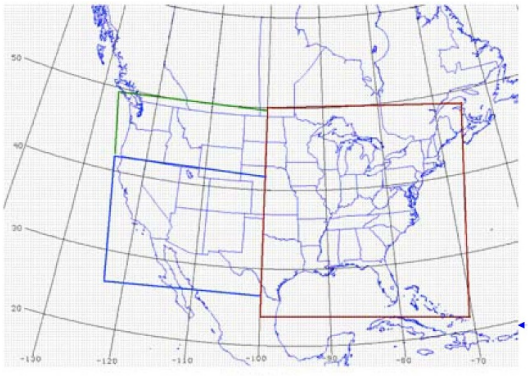

Figure 1. Regions of the USA used as predictors (Northwest, Southwest and East).

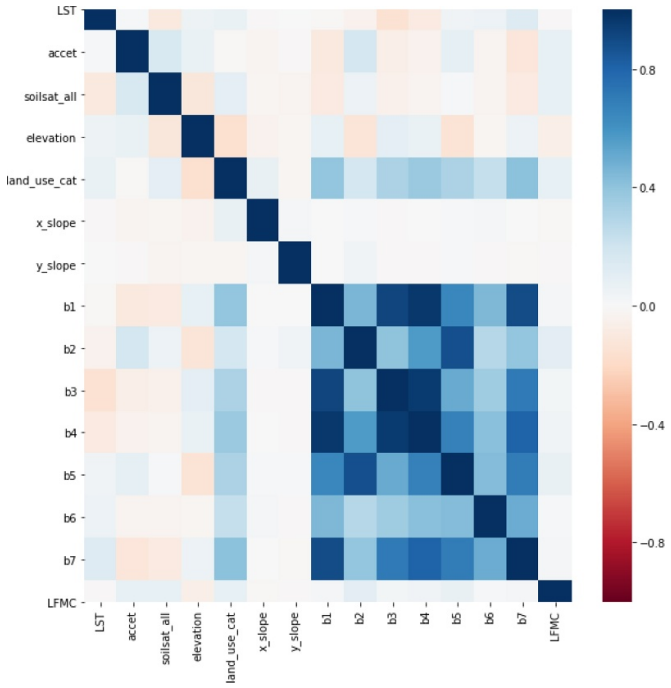

Figure 2. Correlation heatmap for the LFMC prediction.

3. Methodology

In order to assemble datasets for training machine learning models to predict fuel moisture content, a large volume of disparate data was combined into a flattened table format that machine learning code libraries, such as scikit-learn, can ingest.

3.1. Dataset generation

The data included approximately 2 TB of MODIS reflectance data from the Terra and Aqua satellites and 2 TB of WRF-Hydro model data, both covering a 1 km resolution CONUS grid. The number of WFAS observation locations was approximately 500 and RAWS was approximately 1500 as shown in figure 4. An interpolation procedure was performed where the observed fuel moisture values from RAWS or WFAS were matched with their nearest MODIS grid cell in time and space and then inverse distance weighted interpolation was performed for all surrounding grid cells within a radius of 2.5 km. This interpolation procedure generates a larger training dataset since the set of matched observation samples and MODIS reflectance pairs is small for any given day. The optimal radius was determined using cross validation on a subset of the training data over Colorado, United States, where there were sufficient observation sites. Next, the WRF-Hydro data was interpolated to the same grid as the MODIS data using a simple nearest neighbor interpolation, which was required due to coordinate system differences. The flattened WRF-Hydro variables

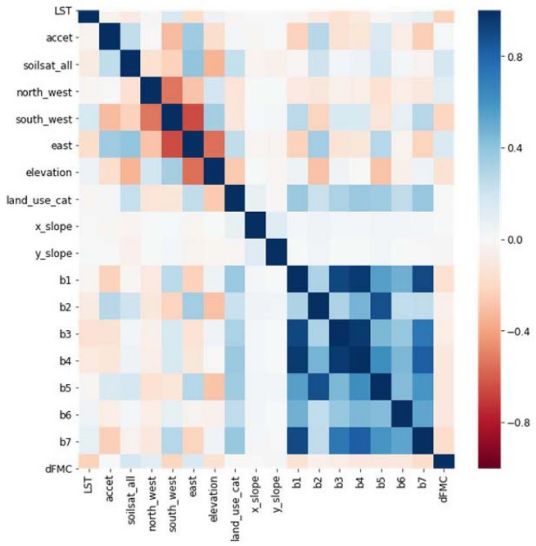

Figure 3. Correlation heatmap for the DFMC prediction.

Figure 4. On the left, Live Fuel Moisture Content (LFMC) observation sites from the WFAS network, and on the right the Dead Fuel Moisture Content (DFMC) observation network from RAWS.

were joined to the existing training set consisting of target fuel moisture values, MODIS reflectance values, and band ratio indices. Finally, ancillary data such as elevation and land use category, for the 1 km CONUS grid were joined with the other data rows at training time.

3.1.1. Hardware and software.

Python 3 was used to perform these calculations and generate the resulting training data sets. Each of the grid files for the WRF-Hydro and MODIS satellites contains on the order of 18×10^6 grid cells which placed burden on the CPU, memory, and disk of the system the data processing was being performed on. For this reason, the compute capabilities of the US National Center for Atmospheric Research (NCAR) Casper cluster supercomputer were utilized in the data processing and storage steps. Whenever possible precomputed interpolations were used however, the interpolation of nearby grid cells to RAWS or WFAS stations is dynamic due to the fact that available stations change over time likely due to campaign field station use and power or communication issues. The use of Python 3 meant that computation ran more slowly than with a language like C++; however, it also allowed rapid software development and testing. To generate the CONUS dataset, the Python multiprocessing library was used to parallelize code and utilize available cluster node cores on Casper. This was combined with the used of the Python Numba just-in-time compiler library [20] which helped speed up the computation needed to perform interpolation, data joining, and spatial operations. The results were not yet satisfactory since large chunks of the input data was being stored in memory at once burdening the system and heavy utilization of cluster CPU cores involved manual code restructuring to utilize the multiprocessing library.

For these reasons the Dask parallelization library [21] was used to replace many Numpy and Pandas functions with analog parallel functions. Dask implements blocked algorithms that break up large arrays or data frames into many smaller structures that are scheduled on available CPU cores to perform operations on a subset of the data. Blocked algorithms compute a large result like 'take the sum of these trillion numbers' with many small computations like 'break up the trillion numbers into one million chunks of size one million, sum each chunk, then sum all of the intermediate sums' [22]. Although Dask is not a complete

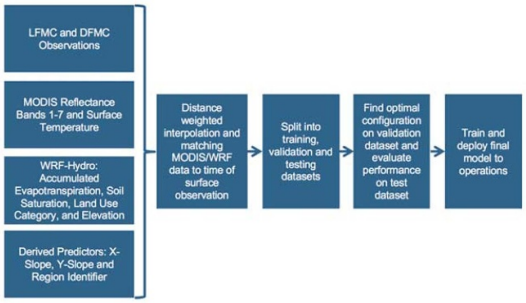

Figure 5. Training system design with the first four boxes on the left indicating the different data sources to merge into one dataset that is then subset to the nearest grid cells to the observation locations. This is then split into training, testing and validation datasets.

parallel replacement for all Numpy and Pandas operations, utilizing available functions in select code hotspots allowed harnessing more cluster processor cores, reduced wall clock time significantly, and allowed files larger than available RAM to be loaded and processed. Dask algorithms were the most efficient in parts of the computation similar to the parallel sum example previously mentioned where many independent computations can be completed at once followed by a simple reduction operation at the end of the computation. Joining a large number of flattened WRF-Hydro variable rows with existing MODIS and fuel moisture rows on location and time fields was still the main code bottleneck even with the use of Dask DataFrame and its corresponding parallel operations.

3.2. Training, testing and validation design

The data sources had different time intervals, grid resolutions and time of data retrievals; therefore, the first step in the dataset creation was to integrate all data sources as illustrated in figure 5. The first four boxes on the left indicate the different data sources to merge into one dataset that is then subset to the nearest grid cells to the observation locations. This is then split into training, testing and validation datasets. Ultimately, the FMC prediction models were to be used for a verification of the WRF-Fire coupled atmosphere wildfire spread model during fires in 2016. Thus, a specific subset of available days in the data between May and November 2016 were subset as the independent test dataset to quantify operational performance. These 25 d occurred in June, July, August, September and October. The remaining days were subset into training and validation. In order to avoid over-training to observation location specific biases, we split the data randomly by site into 80% training and 20% validation where each cluster of grid cells within 2.5-km of an observation location was considered one site.

3.3. Machine learning models

We test multiple machine learning methods in order to determine the optimal method to use in the real-time FMC prediction system.

3.3.1. Multiple linear regression.

Multiple Linear Regression (MLR) is a generalized form of linear regression with more than one predictor variable [23]. MLR is represented by equation (1) where $y'$is the predictand, which in this case is the DFMC or LFMC, b_0 is the intercept, b_k is the coefficient for each predictor x_k, and K is the number of predictors.

$$y' = b_0 + b_1x_1 + b_2x_2 + \ldots + b_Kx_K \tag{1}$$

Each of the regression coefficients are calculated by minimizing the sum of squared residuals through simultaneously solving K + 1 equations for each of the regression coefficients and b_0 in equation (1). MLR assumes that the predictand is a linear combination of each of the regression coefficients and predictor variables, and will not capture non-linear relationships between a predictor and the predictand unless the predictor data is transformed. MLR also assumes homoscedasticity, which is not entirely correct in predicting FMC since there is a minimum value (0%) and typically higher variance with higher FMC values. Here, we use MLR as a baseline technique since it models the linear relationships between the predictors and the predictand. It is important to note that the land use category is a categorical variable, which is not readily interpretable for MLR. We did not transform the categorical variable so that the predictors were consistent across our methods tested.

3.3.2. Random forest.
The RF algorithm has grown in popularity in recent years due to its interpretability and its general ability to avoid over-fitting. RFs are a combination of a specified number of decision trees [24]. Decision trees are a rule-based technique where the entropy of the target and the entropy of each branch based on predictor values are calculated. The difference between the entropy of the target and the entropy of the branch is the information gain. The decision tree starts with creating a rule to branch the tree based on the highest information gain and proceeds until the final branch has an entropy equal to 0 or reaching a rule for stopping such as max depth, which is the final leaf node. The final prediction is the mean of the instances in the final leaf node for an instance that follows the rules of the branches down to the final leaf. While the RF performs well for interpolating, it does not extrapolate due to the mean value for the training data being used as the prediction in the final leaves of the trees. The RF specifically handicaps each tree in the forest with a subset of the available predictors and the available training data; however, by taking an ensemble average of the prediction from each tree in the forest the model tends to avoid over-fitting and having a lower error on average than any tree in the forest. The optimal configuration of the RF for the DFMC prediction, as determined by the lowest error on the test dataset, had 250 trees, 25 minimum samples per split and 25 minimum samples per leaf. For the LFMC prediction, the RF was similar except 1000 trees were used.

3.3.3. Gradient boosted regression.
Gradient boosted regression (GBR) is similar to the RF in that it is fundamentally based on decision trees. The GBR tree works by building regression trees sequentially, rather than independently like the RF. Each sequential tree is trained to predict the residual of the previous tree until the mean residual no longer decreases or the maximum number of trees is reached. The RF and the GBR techniques are both implemented from Python's Scikit-Learn library [25]. The GBR configuration used for the DFMC prediction had 500 trees, a maximum depth of 12, 100 minimum samples per split and a learning rate of 0.3. The LFMC had the same configuration except a maximum tree depth of 8. The differences between RF and GBR are further discussed in sections 4 and 5.

3.3.4. Artificial neural network.
The ANN used here is a multi-layer perceptron trained with backpropagation [26, 27]. The multi-layer perceptron was modelled after how the neurons in the human brain work, with input data being fed into neurons that adjust parameters to output a learned value. Activation functions are used to determine the output of a neuron where more activated neurons have more relevance for the model's prediction. The training process iteratively adjusts the weights and biases until the error is minimized or the number of training iterations, or epochs, is reached. The model used here in the Python TensorFlow Keras model [28] built sequentially with two hidden layers of 10 neurons each and 200 training epochs for the DFMC prediction and 500 training epochs for the LFMC prediction. The configuration of the ANN approach was determined based on a sensitivity study performing a manual grid search of candidate configurations. The number of layers and neurons was chosen as the configuration that had the lowest error on the sensitivity study test set while minimizing difference between the sensitivity training and testing error.

4. Results

The results of the machine learning methods are quantified on the training and testing datasets where the testing dataset includes all locations for 25 d in 2016. The results for the DFMC predictions are shown in table 2. The RF method produces the lowest mean absolute error (MAE) on the test datasets for both Aqua and Terra satellites while maintaining a small difference between the training and testing errors. The GBR technique and the ANN both perform considerably better than the MLR method, which indicates that the relationships between the predictors is non-linear. This also provides evidence that rule based techniques such as the RF and GBR perform best here because there may be distinct decision boundaries in the predictors, such as the predictors with fuel type and climatology information. Note that the mean value of the DFMC across all locations was 9.4% and the standard deviation was 4.5%, which means the RF errors are approximately 33% of the value of the standard deviation.

The results for the LFMC prediction show a similar order of skill for the machine learning methods, with the RF producing the lowest MAEs on the test data. The results for the LFMC predictions are shown in table 3. The GBR technique and the ANN both perform considerably better than the MLR method, which indicates that the relationships between the predictors is non-linear. This also provides evidence that rule based techniques such as the RF and GBR perform best here where there may be distinct decision boundaries in the predictors, such as the predictors with fuel type and climatology information. In order to put these errors into perspective, the mean value of the LFMC across all locations was 94.8% and the standard

Table 2. Model evaluation of the machine learning techniques for predicting DFMC. The RF has the lowest mean absolute error on the test dataset while maintaining minimal difference between training and testing.

	Aqua		Terra	
Method	Train	Test	Train	Test
MLR	2.37%	2.36%	3.28%	3.27%
ANN	1.95%	1.94%	2.65%	2.64%
GBR	1.40%	1.73%	1.97%	2.33%
RF	1.30%	1.56%	1.82%	2.14%

Table 3. Model Evaluation of the machine learning techniques for predicting LFMC. Similar to the DFMC prediction, the RF has the lowest mean absolute error on the test dataset while maintaining minimal difference between training and testing.

	Aqua		Terra	
Method	Train	Test	Train	Test
MLR	30.97%	30.37%	30.30%	30.39%
ANN	28.38%	28.58%	27.15%	27.76%
GBR	19.59%	23.87%	18.28%	23.56%
RF	18.97%	21.92%	19.28%	22.06%

Table 4. Predictor importance, as a percentage, for the RF method predicting DFMC on the Terra satellite dataset.

DFMC Predictor (CONUS, Terra)	Importance
Land Surface Temperature	40%
Elevation	15%
Soil Saturation All Layers	7%
Accumulated Evapotranspiration	7%
Band 7	5%
East Region	4%
Band 2	3%
East West Slope	3%
North South Slope	3%
Band 1	2%
Band 3	2%
Band 5	2%
Band 6	2%
Band 4	1%
Land Use Category	1%
Northwest Region	1%
Southwest Region	1%

deviation was 86.2%, which means the RF errors are approximately 25% of the value of the standard deviation. Recall that the LFMC values can be greater than 100% when there is more weight to the water within the fuels than the weight of dry fuels.

In order to better understand the decisions behind branches in the tree, we performed a predictor importance analysis to determine the predictors that provided the most value to the model. These are shown for the DFMC predictions in table 4 and the LFMC predictions in table 5 for the Terra satellite dataset. Although results are not shown for Aqua, the results were very similar except Band 6 has very little to no importance due to a known issue with the Band 6 reflectance from Aqua [29].

The predictor importance analysis for the DFMC prediction shows how the model first captures the importance of the seasonal climatology and to a lesser degree fuel types (i.e. different types of vegetation) with elevation and land surface temperature. The soil saturation and accumulated evapotranspiration have the next highest predictor importance, which are valuable predictors to quantify how moist the soil and fuels are with higher moisture fluxes when there is greater moisture content. The remaining predictors all provide additional predictive information in ranges of 1%–5% importance. For the LFMC prediction, the land surface temperature is less important with a much higher importance of soil saturation and accumulatione evapotranspiration. This makes physical sense as living fuels will be more susceptible to recent precipitation and more readily bring moisture through their roots into the plants.

In addition to correlation analysis and predictor importance, we evaluate predictor partial dependence plots to better interpret the RF model predictions. Partial dependence plots are useful for understanding how

Table 5. Predictor importance, as a percentage, for the RF method predicting LFMC on the Terra satellite dataset. Note that regions were not used as predictors for the LFMC models when they were included for the DFMC models.

LFMC Predictor (CONUS, Terra)	Importance
Elevation	22%
Soil Saturation All Layers	17%
Accumulated Evapotranspiration	11%
East-West Slope	7%
Band 2	6%
Band 6	6%
Land Surface Temperature	6%
North-South Slope	5%
Land Use Category	5%
Band 7	4%
Band 1	3%
Band 4	3%
Band 5	3%
Band 3	1%

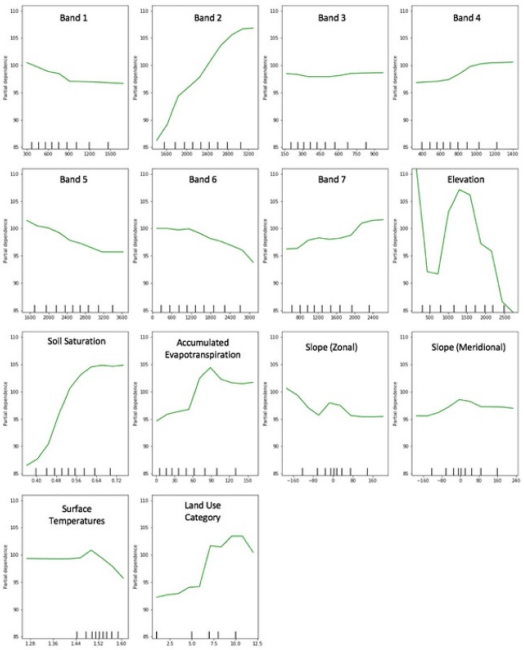

Figure 6. Partial dependence plots for the LFMC prediction on the Aqua dataset. The y-axis represents the LFMC prediction and the x-axis represents the change in predictor variable value.

each predictor impacts the model's predictions. Partial dependence plots are created after the model has been fit to the training data and then applied to make predictions while holding the predictor of interest constant. After all predictions are made on the dataset, the predictor of interest is increased and predictions are made for the same dataset. We iterate this process over the range of predictor values in increments of 1/10 of the range of the predictor data. This analysis allows interpretability of when the predictions change based on how the value of the predictor variable of interest changes. Figure 6 shows the partial dependence plots for the LFMC prediction while figure 7 shows the partial dependence plots for the DFMC prediction for the models trained on MODIS Aqua. The analysis for the LFMC predictions shows a few important patterns. First, the elevation, which is the most important predictor has a non-linear shape with a minimum near 500 m elevation, an increase until 1500 m elevation and then a steady decrease as the elevation increases beyond 1500 m. This is likely capturing specific patterns of the types of living fuels present in these elevation ranges. The soil saturation, which was the second most important prediction, has a shape as expected where the LFMC prediction increases as the soil saturation increases. Similar is true for the accumulated evapotranspiration although it reaches a peak before slowly decreasing after 90 mm day^{-1}.

The DFMC predictions in figure 7 show how three of the four most important predictors (land surface temperature, elevation and soil saturation) have patterns in the partial dependence. Other variables may

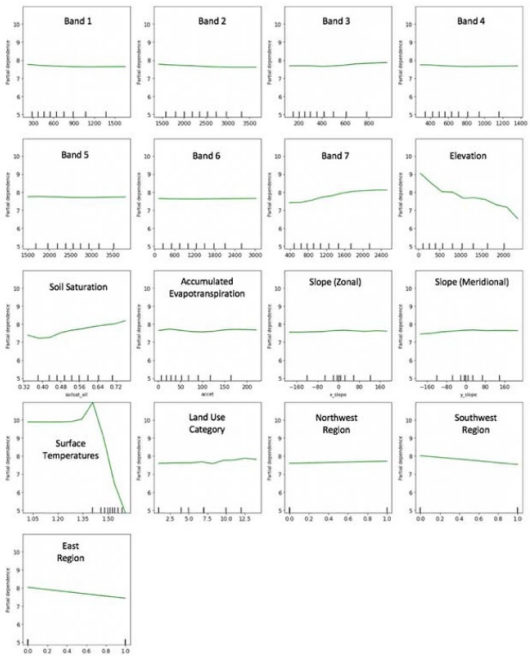

Figure 7. Partial dependence plots for the DFMC prediction on the Aqua dataset. The y-axis represents the DFMC prediction and the x-axis represents the change in predictor variable value.

improve the overall prediction accuracy, but they may only minimally change the overall prediction or interact with multiple other variables in a way that does not show the specific partial dependence of that variable on the DFMC prediction. The land surface temperature, which has by far the most predictor importance, shows a relatively flat pattern until at higher temperatures the DFMC predictor importance drops significantly. This makes physical sense as we would expect the hottest days in the summer to generally be the most dry.

5. Discussion and real-time system

In this study we evaluated a range of machine learning approaches to capture the non-linear relationships between predictors from satellite data, weather observations surface characteristics and the predictands of live andDFMC. We compared these methods to a MLR approach that does not capture non-linear relationships and found that all machine learning techniques improve upon MLR. Additionally, the MLR model uses the same available predictors as the machine learning approaches, which the predictors are a variety of continuous, binary and categorical variables that are better suited for machine learning approaches with decision boundaries such as the RF or GBR. Not surprisingly, the model that ultimately performed best was the RF technique, although both the RF and GBR had lower errors than the ANN.

One reason that the RF and GBR techniques tend to perform well is that they are rule-based techniques that are best used in preditive situations with sharp decision boundaries. The physical problem we are trying to model is the fuel moisture content of living and dead fuels, which has fundamental categories such as fuel types, soil moisture types and different weather or climatological regions. The RF and GBR models can differentiate these regions and then further refine the fuel moisture content predictions based on the predictive skill from the satellite reflectances and other surface characteristics. The ANN, however, is unable to as easily separate the decision boundaries and learn as much of the predictive ability of the available data as the RF and GBR.

The benefit of the RF method over the GBR method is that it tends to avoid over-fitting the noise in the dataset by bagging. Bagging attempts to approve the stability and accuracy of regression by reducing the variance. Due to the fact that there are many interpolations and assuming continuity within a 1-km grid cell, it is logical that there would be a considerable amount of noise in the dataset and that a bagging technique such as the RF would perform best.

The RF method was implemented in a real-time system in the early fall 2019. The system produces daily dead and LFMC maps and datasets at 1-km resolution, meeting an important need of the wildfire research and firefighting community. The real-time system runs at 00:10Z starting with a python runner script that is triggered by the crontab that builds up a list of the most recent MODIS and WRF-Hydro files needed for the

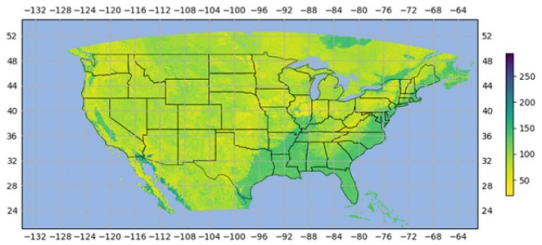

Figure 8. LFMC estimates from the RF for November 12th, 2019. Darker greens indicate higher moisture content and brighter yellow indicate lower moisture content.

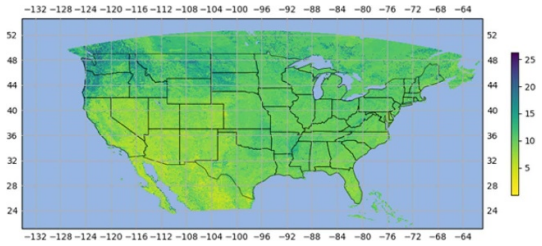

Figure 9. DFMC estimates from the RF for November 12th, 2019. Darker greens indicate higher moisture content and brighter yellow indicate lower moisture content.

predictions and then calls the trained RF models to make a prediction. These prediction datafiles in NetCDF format are availabe for download daily through the NCAR's Research Application Laboratory's (RAL) web site at https://ral.ucar.edu/projects/improved-wildland-fire-spread-prediction. The daily FMC maps are also uploaded to the same site in .png format. Historical datafiles and maps can be downloaded form NCAR/RAL's FTP site at ftp://ftp.rap.ucar.edu/pub/nasafmc. In addition to the publicaly available datasets, The predictions of LFMC for November 12th and DFMC also for November 12th are shown in figures 8 and 9. It is interesting to point out the irrigated central valley of California that shows higher moisture content in the LFMC predictions in figure 8, while dead fuels are not purposely irrigated and there is a much less prominent central valley of California in figure 9.

Acknowledgments

The authors would like to thank the many colleagues at NCAR who collaborated on building a machine learning based real-time system for a gridded fuel moisture content system, including Pedro Jimenez Munoz who downloaded and processed the MODIS data. The authors wish to thank collaborators from the University of Colorado, Steven Massie, and Brad Schmidt from the CoE Advanced Technology Aerial Firefighting. The Project was funded by NASA-AIST Award No. 16-0079. This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.5065/qt42-zd40. .

ORCID iD

Tyler C McCandless https://orcid.org/0000-0002-2068-4193

References

[1] Coen J L, Cameron M, Michalakes J, Patton E G, Riggan P J and Yedinak K M 2013 WRF-fire: coupled weather–wildland fire modeling with the weather research and forecasting model *J. Appl. Meteorol. Climatol.* **52** 16–38

[2] Skamarock W C 2004 Evaluating mesoscale NWP models using kinetic energy spectra *Mon. Weather Rev.* **132** 3019–32

[3] Rothermel R C 1972 A mathematical model for predicting fire spread in wildland fuels (available at: http://www.treesearch.fs.fed.us/pubs/download/32533.pdf) (Accessed February 10, 2017)
[4] Mandel J and Kulkarni V 2006 Construction of a level function for fireline data assimilation *UCDHSC/CCM Report No. 234* University of Colorado at Denver, Center for Computational Mathematics(available at: http://www-math.ucdenver.edu/ccm/reports/rep234.pdf)
[5] National Fuel Moisture Database (NFMD) (available at: http://www.wfas.net/index.php/national-fuel-moisture-database-moisture-drought-103) (Accessed June 1, 2018)
[6] Hengl T *et al* 2017 SoilGrids250m: global gridded soil information based on machine learning *Plos One* **12** e0169748
[7] Chang N-B, Xuan Z and Yang Y J 2013 Exploring spatiotemporal patterns of phosphorus concentrations in a coastal bay with MODIS images and machine learning models *Remote Sens. Environ.* **134** 100–10
[8] Ahmad S, Kalra A and Stephen H 2010 Estimating soil moisture using remote sensing data: A machine learning approach *Adv. Water Resour.* **33** 69–80
[9] Rhee J and Lm J 2017 Meteorological drought forecasting for ungauged areas based on machine learning: using long-range climate forecast and remote sensing data *Agric. Forest Meteorol.* **237-238** 105–22
[10] Adab H, Kanniah K D and Beringer J 2016 Estimating and up-scaling fuel moisture and leaf dry matter content of a temperate humid forest using multi resolution remote sensing data *Remote Sens.* **8** 25
[11] Cohen S, Praskievicz S and Maidment D R 2018 Featured collection introduction: national water model *J. Am. Water Resour. Assoc.* **54** 767–9
[12] Lin P, Yang Z-L, Gochis D J, Yu W, Maidment D R, Somos-Valenzuela M A and David C H 2018 Implementation of a vector-based river network routing scheme in the community WRF-Hydro modeling framework for flood discharge simulation *Environ. Model. Softw.* **107** 1–11
[13] Zachariassen J, Zeller K F, Nikolov N and McClelland T 2003 A review of the Forest Service Remote Automated Weather Station (RAWS) network *Technical Report* RMRS-GTR-119. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station
[14] Burgan R E, Andrews P L, Bradshaw L S, Chase C H, Hartford R A and Latham D J 1997 WFAS:wildlandfire assessment system *Fire Manag. Notes* **57** 14–17
[15] Vermote E 2015 MOD09A1 MODIS Surface Reflectance 8-Day L3 Global 500m SIN Grid V006 NASA EOSDIS Land Processes DAAC. (https://modis.gsfc.nasa.gov/data/dataprod/mod09.php)
[16] Vermote E F and Kotchenova S 2008 Atmospheric correction for the monitoring of land surfaces *J. Geophys. Res.* **113** 103–10
[17] Wang Y, Lyapustin A I, Privette J L, Cook R B, SanthanaVannan S K, Vermote E F and Schaaf C L 2010 Assessment of biases in MODIS surface reflectance due to Lambertian approximation *Remote Sens. Environ.* **114** 2791–801
[18] Gochis D J *et al* 2018 The WRF-Hydro modeling system technical description (Version 5.0) *Technical Note* NCAR pp 1–107 (available at: https://ral.ucar.edu/projects/wrf_hydro/technical-description-user-guide)
[19] U.S. Geological Survey 2011 Gap Analysis Program *Report* No. 20160513 GAP/LANDFIRE National Terrestrial Ecosystems (available at: https://doi.org/10.5066/F7ZS2TM0)
[20] Siu Kwan L, Pitrou A and Seibert S 2015. Numba: a LLVM-based Python JIT compiler, *Proc. of the Second Workshop on the LLVM Compiler Infrastructure in HPC* (November 15-15, 2015, Austin, Texas USA) pp 1–6
[21] Dask Development Team 2016 Dask: library for dynamic task scheduling (available at: https://dask.org)
[22] Rocklin M 2015 Dask: parallel computation with blocked algorithms and task scheduling *Proc. 14th Python in Sci. Conf.*
[23] Wilks D S 2011 *Statistical Methods in the Atmospheric Sciences: Third Edition* (Amsterdam: Elsevier)
[24] Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32
[25] Pedregosa F *et al* 2011 Machine learning in Python *J. Mach. Learn. Res.* **12** 2825–30
[26] Reed D R and Marks R J 1998 Neural smithing: supervised learning in feedforward artificial neural networks, ed D R Reed and R J Marks (Cambridge, MA: MIT Press) p 358
[27] Rosenblatt F 1958 The perceptron: a probabilistic model for information storage and organization in the brain *Psych. Rev.* **65** 386–408
[28] Chollet F *et al* 2015 Keras (available at: http://keras.io)
[29] Wang L, Qu J J, Xiong X, Hao X, Xie Y and Che N 2006 A new method for retrieving Band 6 of aqua MODIS *IEEE Geosci. Remote Sens. Lett.* **3** 267–70