

Drought indices predict changes in live fuel moisture content (LFMC) across French Mediterranean shrublands: a multisource and machine learning approach

Xiangzhuo Liu^{a,*}, Nicolas Martin-StPaul^a, Julien Ruffault^a, Gengke Lai^b, Kevyn Raynal^a, RodriguezSuquet Raquel^c, Huan Wang^d, Russell Parsons^e, Jean-Luc Dupuy^a, Francois Pimont^a

^a URFM, INRAE, 84000 Avignon, France

^b International Institute for Earth System Sciences, Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing University, Nanjing, China

^c CNES, F-3100 Toulouse, France

^d College of Urban and Environmental Sciences, Peking University, Beijing 100871, China

^e USDA Forest Service, Fire Sciences Lab, Missoula, MT, 59808, USA

ARTICLE INFO

Keywords:

Live fuel moisture content
XGBoost
Explainable machine learning
Drought indices
French Mediterranean

ABSTRACT

The French Mediterranean region is among the most wildfire-prone areas in Europe, and fire activity is projected to intensify further under climate change. Live fuel moisture content (LFMC), a key driver of fire behavior, is critical for fire danger monitoring. However, the scarcity of high-quality LFMC data in this region hampers understanding of wildfire drivers and limits robust fire danger assessment. To address this gap, we developed an LFMC prediction model for the French Mediterranean fire season using the extreme gradient boosting (XGB) algorithm, hereafter referred to as the LFMC-XGB model. The input variables of the LFMC-XGB model were selected from the remote sensing (both optical and microwave) and weather-related variables using the leave-one-site-out cross-validation method. The results showed that the best performance was achieved by combining the Drought Code and Duff Moisture Code calculated from SAFRAN weather reanalysis data, the root zone soil moisture from the Soil Moisture Active Passive product, and the non-tree percentage from the Moderate Resolution Imaging Spectroradiometer product. When evaluated against in-situ LFMC measurements, the LFMC-XGB model yielded an R^2 of 0.56 with an RMSE of 13.95 %. The model also performed well in simulating site temporal fluctuations ($R^2 = 0.58$ and $RMSE = 12.21$ %) and site means ($R^2 = 0.52$ and $RMSE = 6.77$ %). Shapley additive explanation (SHAP) analysis identified the Drought Code and Duff Moisture Code as the most influential predictors, exhibiting monotonically decreasing contributions with increasing LFMC. Non-tree cover and root zone soil moisture followed, showing stepwise response patterns. Using the LFMC-XGB model, we generated daily LFMC maps at 250 m resolution for 2016–2022. The analysis of LFMC maps highlighted the potential of the LFMC-XGB model to strengthen operational fire danger prediction and improve understanding of fire regime drivers in the French Mediterranean.

1. Introduction

The Mediterranean region is one of three widely recognized global wildfire hotspots, where fire activity threatens ecosystems, humans and infrastructure, and where fire frequency and intensity are projected to increase under ongoing global changes (Amraoui et al., 2013; Cunningham et al., 2024; Dupuy et al., 2020; Pimont et al., 2023; Pimont et al., 2022; Ruffault et al., 2020). The French Mediterranean region, in

particular, is expected to experience significant expansions of fire-prone areas and prolonged fire seasons (Pimont et al., 2022). Live fuel moisture content (LFMC), defined as the ratio of water mass and dry mass in fresh foliage (Quan et al., 2021b), plays an important role in wildfire behavior. It is a key variable for predicting fire size, especially when it falls below a critical threshold, at which point the likelihood of large fires increases (Dennison and Moritz, 2009; Luo et al., 2019; Nolan et al., 2016). It also drives fire spread rate (Pimont et al., 2019), a crucial

* Corresponding author at: ISPA, INRAE, Bordeaux, France.

E-mail addresses: xiangzhuo.liu@inrae.fr, xiangzhuo.liu@hotmail.com (X. Liu).

parameter for estimating wildfire behavior. Hence, the development of an accurate LFMC product for the French Mediterranean region is critical for improving wildfire prediction, prevention and management.

LFMC variability is primarily controlled by two physiological mechanisms: (i) plant responses to drought, which affect water mass condition, and (ii) vegetation phenology changes, which drive fluctuations in dry biomass (Jolly et al., 2014; Jolly and Johnson, 2018; Nolan et al., 2022). Both mechanisms are influenced by meteorological conditions, soil characteristics, and vegetation traits (Ruffault et al., 2023; Ruffault et al., 2018; Scarff et al., 2021). Meteorological and remote sensing data have been widely utilized to estimate LFMC. Estimation methods can be broadly categorized into empirical models, radiative transfer models, and artificial intelligence-based methods. Empirical models establish linear or non-linear relationships between in-situ LFMC measurements and drought indices derived from weather data or vegetation indices obtained from satellite or aerial sensors. Although these models have shown good performance across various regions (Arganaraz et al., 2016; Dennison et al., 2003; Peterson et al., 2008), including some Mediterranean countries (Castro et al., 2003; Costa-Saura et al., 2021; Dimitrakopoulos and Bemmerzouk, 2003; Jolly et al., 2025; Marino et al., 2020; Pellizzaro et al., 2007), they are typically calibrated for specific species or sites with similar environmental conditions, which limits their generalizability and operational applicability at larger spatial scales (Ruffault et al., 2018). Radiative transfer models simulate the propagation and interaction of radiation within the canopy using physical principles that explicitly link vegetation physiological and morphological and optical traits (specific leaf area, leaf properties) to canopy or leaf spectral responses (Houborg et al., 2007), thereby allowing the retrieval of LFMC independent from sensor specifications and site conditions (Yebra et al., 2013). Radiative transfer models have been successfully used at the country level (Yebra et al., 2018) and in gable-scale LFMC mapping (Quan et al., 2021b). However, the parameterization of radiative transfer models is complicated and needs prior pieces of knowledge to constrain the ill-posed problem (Yebra and Chuvieco, 2009). More importantly, the lack of radiative transfer models specifically developed for shrublands introduces greater uncertainty in LFMC retrievals for these ecosystems compared to other vegetation types (Lai et al., 2022). With the recent availability of more extensive in-situ LFMC measurements and explanatory data (Yebra et al., 2024; Yebra et al., 2019), artificial intelligence models have been increasingly employed to estimate LFMC because they can uncover complex non-linear relationships and integrate diverse data sources (Miller et al., 2022; Rao et al., 2020; Zhu et al., 2021). However, the way in which the variables driving LFMC influence model outcomes has not yet been explored in artificial intelligence models.

There are still some limitations in previous studies, although they have achieved good results in different regions. First, in most cases, these models were calibrated and evaluated over an entire vegetation season. Such an approach allows a clear discrimination between dry (summer) and wet (winter) conditions, but it may conceal the more operationally relevant challenge: during the core fire season, when LFMC values are generally low and their range is narrow, the key issue is to distinguish between dryness states beyond the critical threshold. Furthermore, the respective contributions of temporal and spatial performance to overall model accuracy remain largely underexplored in most previous LFMC studies (Costa-Saura et al., 2021; Cunill Camprubí et al., 2022; Miller et al., 2023). As a result, it makes it difficult to assess whether prediction errors arise primarily from inconsistencies across time or biases across space. Finally, a rigorous assessment of the model is needed. In a lot of cases, a random subset of the data is held out for evaluation (Benali et al., 2025; Tanase et al., 2022), which is unable to fully assess a model's effectiveness in spatial mapping (Meyer et al., 2019). The rigorous spatial cross-validation strategies, such as leave-one-site-out (Arcos et al., 2024; Costa-Saura et al., 2021) were less used.

In the French Mediterranean region, LFMC modelling and prediction

appear to be particularly challenging due to a large variability in soil and plant species (Ruffault et al., 2018). Fan et al. (2018) evaluated the ability of various remote sensing microwave and optical indices to capture the temporal dynamics of LFMC across different shrublands species in southeastern France and found that the optical indices outperformed microwave indices. In that study, LFMC data was obtained from the French "Reseau Hydrique" database (Martin-StPaul et al., 2018), which provides in situ LFMC measurements at several sites in southeastern France during the fire season (June to September). In addition, Ruffault et al. (2018) used 14 years (2000–2013) of LFMC measurements from the same database to evaluate the performance of various meteorological drought indices in predicting LFMC using a linear equation. The results indicate that combining all sites reduces spatial predictability, which complicates LFMC product development in this region. Currently, no LFMC product has been specifically developed for the French Mediterranean region. The available LFMC maps in this area include two products retrieved from Moderate Resolution Imaging Spectroradiometer (MODIS) observations: the global LFMC product based on radiative transfer models (Quan et al., 2021b) and the random forest model-based LFMC product (Cunill Camprubí et al., 2022) for the Mediterranean region (Spain, France, Italy, and Tunisia), respectively. The evaluation of the random-forest-based model LFMC product (Cunill Camprubí et al., 2022) demonstrated greater agreement with field measurements than the global LFMC product over the Mediterranean region. However, its overall performance remains suboptimal ($R^2 = 0.37$), and the performance in the French Mediterranean region is unknown. Unfortunately, neither of the products has been updated. Recent studies have shown that combining remote sensing and climatic variables can enhance LFMC estimation in the Spanish Mediterranean region (Arcos et al., 2024; Costa-Saura et al., 2021; Viñuales et al., 2024) and in Portugal (Benali et al., 2025), however, their effectiveness in the French Mediterranean region remains unknown.

Within this context, we hypothesized that the combination of available data sources from remote sensing and meteorological variables can predict LFMC at a large scale using advanced prediction and validation techniques based on machine learning. This study aims to i) identify the most influential variables affecting LFMC in the French Mediterranean over a large range of potential predictors; ii) develop an accurate LFMC product during the core of the fire season for the French Mediterranean region; iii) explore how these influential variables drive LFMC changes. To address these objectives, we collected a large range of potential predictors and used the eXtreme Gradient Boosting (XGB) model with leave-one-site-out evaluation strategy to select the most influential variables for developing the LFMC mapping model, named the LFMC-XGB model. Furthermore, to enhance interpretability of the LFMC-XGB model, we applied the TreeExplainer-based Shapley additive explanation (SHAP) method (Lundberg et al., 2020), which allows us to quantify the dependence of LFMC changes on the input variables. Finally, the potential of the LFMC product over the French Mediterranean is illustrated by performing comparisons with burned area.

2. Materials and methods

2.1. Study region and in-situ LFMC

This study focused on the French Mediterranean region, located in southeastern France (Fig. 1). The area features a typical Mediterranean climate, characterized by hot and dry summers. Vegetation is dominated by shrublands, alongside coniferous and sclerophyll evergreen forests. Shrubs constitute a major component of fine and highly flammable fuels and play a key role in fire ignition and spread in southeastern France (Pimont et al., 2019). As a result, this region is highly fire-prone and accounts for the majority of the burned area in France during the summer season (Castel-Clavera et al., 2025).

The in-situ LFMC measurements were obtained from the 'Reseau Hydrique' network, managed by the French National Forest Service.

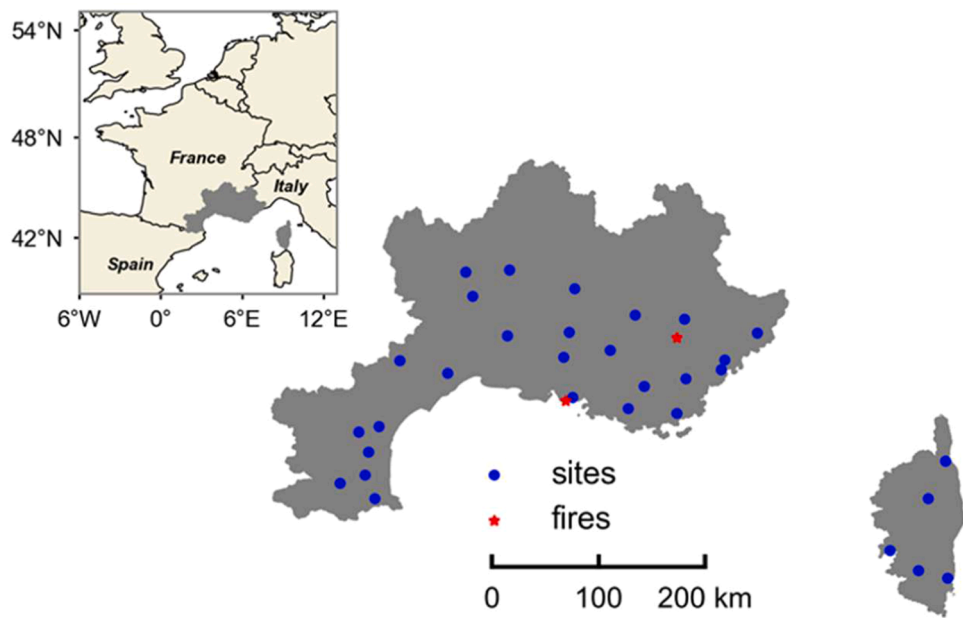


Fig. 1. Location of the study region and sites used to build the LFM retrieval model.

LFMC was measured nearly weekly since 1996 during the fire season (June to September) without rainfall events. At each site, two to three dominant shrubland species were sampled, with five samples per species and date, representing the surrounding 1000 m² landscape. To facilitate scientific use, [Martin-StPaul et al. \(2018\)](#) curated the raw observations and developed the robust “Reseau Hydrique” database, which is updated as new measurements and sites became available.

The study period spans from 2016 to 2022, corresponding to the overlapping period of the remote sensing and meteorological variables (see [Sections 2.2](#)). To ensure a robust evaluation, only sites with at least four years of observations within the study period were considered. At each site, LFM was calculated by averaging species-level measurements collected on each sampling date, yielding a site-level observed LFM. In total, 2000 samples from 31 sites, encompassing 19 species,

were kept. Most sites represent Mediterranean shrubland species, although four sites also include tree species. The spatial distribution of sites is shown in [Fig. 1](#), and detailed information is provided in [Table S1](#).

2.2. Remote sensing and weather- derived data

Several variables were extracted or derived from remote sensing and weather data to select the optimal variables to predict LFM in this study. The main characteristics of each variable are shown in [Table 1](#), with further details provided below.

2.2.1. Vegetation information data

Two types of vegetation data, dynamic and static, were evaluated for LFM estimation. The dynamic vegetation information included five

Table 1
The variables used to test the prediction of LFM.

Name	Category	Source	References
Normalized Difference Vegetation Index (NDVI)	Remote-sensing based	Sentinel-2	(Rouse et al., 1973)
Normalized Difference Water Index (NDWI)			(Gao, 1996)
Soil Adjusted Vegetation Index (SAVI)			(Huete, 1988)
Enhanced Vegetation Index (EVI)			(Huete et al., 2002)
Visible Atmospherically Resistant Index (VARI)			(Gitelson et al., 2002)
Cross-ratio (CR)		Sentinel-1	(Vreugdenhil et al., 2018)
Radar Vegetation Index (RVI)			(Nasirzadehdizaji et al., 2019)
Non-tree cover fraction (Notree)		MODIS	(DiMiceli et al., 2015)
CanopyHeight (CH)		Sentinel-1 & Sentinel-2	(Schwartz et al., 2023)
Aboveground biomass (AGB)			
Surface Clay (ClaySurf)		SoilGrids 2.0	(Poggio et al., 2021)
Surface Sand (SandSurf)			
Surface Silt (SiltSurf)			
Surface soil organic carbon (SOCSurf)			
Root zone Clay (ClayRoot)	Weather-derived	GMTED 2010	(Danielson and Gesch, 2011)
Root zone Sand (SandRoot)			
Root zone Silt (SiltRoot)			
Surface soil organic carbon (SOCRoot)		MODIS	(Wan et al., 2021b)
Digital elevation model (DEM)			
Slope		SMAP	(Reichle et al., 2022)
Land surface temperature (LST)			
Root zone soil moisture (smRoot)		SAFRAN	(Wagner, 1987)
Surface soil moisture (smSurf)			
Duff Moisture Code (DMC)		SAFRAN	(Yuan et al., 2019)
Drought Code (DC)			
Vapour pressure deficit (VPD)			

Vegetation indices (VIs) computed from the optical satellite Sentinel-2 reflectance data, two VIs computed from the imaging radar satellite Sentinel-1 backscatter coefficient data, and the non-tree (Notree) cover fraction data from MODIS Vegetation Continuous Fields product. The static vegetation variables were the canopy height (CH) and above-ground biomass (AGB) data, derived from Forest Multiple Source (FORMS) products (Schwartz et al., 2023). The five optical VIs included Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Soil Adjusted Vegetation Index (SAVI), Enhanced Vegetation Index (EVI), and Visible Atmospherically Resistant Index (VARI), which were widely used in estimating LFMC (Lai et al., 2022). Two microwave VIs were cross-ratio (CR) and Radar Vegetation Index (RVI), which were used in vegetation water content studies (Huang et al., 2016; Khabbaza et al., 2022; Vreugdenhil et al., 2018). Since in-situ measurements were taken in shrubland, the Notree variable was used as a proxy for shrubland fraction (Lai et al., 2022). CH and AGB were used to describe the vegetation structure and biomass.

Sentinel-2 data provided by the French Space Agency was created from Sentinel-2 Level-2A product with atmosphere correction and resampled at 250 m resolution. The image processing of this dataset takes into account the mask of zones that were not covered by forests and shrublands (provided by the French Forest Service) and the cloud mask, in order to consider only the reflectance of relevant vegetation. For consistency, Sentinel-1 Ground Range Detected data downloaded from the Google Earth Engine platform was also resampled to 250 m after preprocessing, including thermal noise removal, radiometric calibration, and terrain correction. Sentinel-1 data from ascending orbit (nighttime overpass) were used because they provided more observations that match the timing of the observed LFMC than data from the descending orbit. In Europe, the temporal resolution of Sentinel-1 and Sentinel-2 can reach 6 days and 5 days, respectively. This study merged the Sentinel-1 and Sentinel-2 data into weekly data for use. The MODIS Vegetation Continuous Fields data were obtained from the MOD44B version 6.1 dataset (DiMiceli et al., 2022), which provides annual 250 m resolution data on non-tree cover fraction. FORMS provides the best CH and AGB product for France currently and contains data exclusively from the year 2020. In FORMS products, the CH map was first estimated at 10 m spatial resolution by training a deep learning U-Net model from Global Ecosystem Dynamics Investigation height data, Sentinel-1 and Sentinel-2 data. Then the CH map was converted to the 30 m AGB through the allometric equations calibrated by the National Forest Inventory AGB plot data. FORMS products were resampled to a 250 m grid using the averaging method.

2.2.2. Soil moisture and land surface temperature data

In addition to including vegetation information from remote sensing products, we also included surface and root zone soil moisture and land surface temperature (LST) products, which have been used for LFMC estimation in some studies (Chuvieco et al., 2004; Sow et al., 2013; Vinodkumar et al., 2021). In this study, the soil moisture (SM) at the surface (0–5 cm, represented by smSurf) and root zone (0–100 cm, represented by smRoot) were obtained from the Soil Moisture Active Passive (SMAP) Level 4 product, which provides the 3-hourly, 9-km resolution SM data by assimilating SMAP L-band brightness temperature observations into the National Aeronautics and Space Administration Catchment land surface model (Reichle and Koster, 2005). Evaluation based on core validation site observations showed a significant improvement in SMAP L4 SM compared to model-based SM alone (Reichle et al., 2017). The daily average smSurf and smRoot were used in this study. The weekly LST data were merged from the 1 km MODIS MOD11A1 daily LST product (Wan et al., 2021a) after being filtered using the data quality flag. The smSurf, smRoot, and LST data were resampled to 250 m grid by using the bilinear interpolation.

2.2.3. Soil properties and terrain data

Soil properties and topography affect the water supply and energy

balance of the vegetation (Dymond et al., 2017), which indirectly affects the vegetation LFMC changes. Therefore, those parameters are also included. Corresponding to smSurf and smRoot, four soil property data, including, clay, sand, silt and soil organic carbon (soc) content, for the surface (0–5 cm) and root zone (0–100 cm) were obtained from the SoilGrids 2.0 product (Poggio et al., 2021). The terrain variables are the digital elevation model (DEM) from 7.5 arc-seconds resolution (around 231 m) Global Multi-resolution Terrain Elevation Data 2010 (Danielson and Gesch, 2011) and the corresponding slope. The SoilGrids 2.0 products are already at 250 m, and the terrain data were resampled to 250 m.

2.2.4. Drought indices

Following Ruffault et al. (2018), we selected the Duff Moisture Code (DMC) and Drought Code (DC), two components of the Canadian Fire Weather Index System (Wagner, 1987) as predictors of LFMC. These indices represent the drying of organic fuels at moderate (5–10 cm) and deep (10–20 cm) depths, respectively. DC reflects long-term drought conditions (over a 53-day timescale), while DMC represents short- to medium-term drought conditions (over a 15-day timescale). In addition, the vapour pressure deficit (VPD), describing the degree of air dryness, was also considered as it is one of the main drivers of plant transpiration (Grossiord et al., 2020; Martin-StPaul et al., 2023). The calculations of both DMC and DC use temperature and rainfall, but DMC additionally incorporates relative humidity. VPD was calculated by using the temperature, relative humidity and altitude. The DC and DMC were calculated following the procedures outlined in (Wang et al., 2017), while VPD was computed using the method described in (Yuan et al., 2019). The inputs of each variable calculation were extracted from the hourly SAFRAN product at 8 km × 8 km resolution (Vidal et al., 2009). We used the DC, DMC and VPD at noon, corresponding to the time of in-situ LFMC measurements, and then resampled them into 250 m spatial resolution.

2.2.5. Ancillary dataset

In addition, the mask data used during the preprocessing of Sentinel-2 imagery and burned area data from the MODIS MCD64A1 version 6.1 dataset (Giglio et al., 2021) were also used in the LFMC mapping and analysis. The mask data was derived from the fuel susceptibility map produced by the French Forest Service (Office National des Forêts, 2022). As with the Sentinel-2 data, regional LFMC maps were restricted to shrubland and forest ecosystems to ensure consistency between model training data and its spatial application. Burned area data were derived from the MODIS monthly global burned area product (500 m spatial resolution), which provides the day of burning for each pixel. Annual burned area was aggregated from June to September for each year, and two representative fire events (Fig. 1) were selected to analyze the LFMC product.

2.3. LFMC-XGB model development

To develop a robust model for predicting LFMC for the French Mediterranean region, we used the XGB algorithm (Chen and Guestrin, 2016). XGB algorithm is a tree-based machine learning method that has been shown to outperform traditional deep learning models when applied to tabular datasets (Borisov et al., 2024; Grinsztajn et al., 2022; Lundberg et al., 2020), and such methods are widely used for estimating vegetation characterization parameters (Chen et al., 2022; Geng et al., 2021; Jia et al., 2024; Li et al., 2019). Tree-based models are generally robust to the effects of multicollinearity (Chowdhury et al., 2022). The advantage of the XGB algorithm is that it reduces prediction error by integrating multiple decision trees to optimize the model iteratively and gradually. In each iteration, the XGB algorithm optimizes the objective function by gradient descent and introduces a regularization term to control the complexity of the model and prevent overfitting.

Before the development of the LFMC-XGB model, remote sensing and

weather-derived variables were extracted for the pixels whose center coordinates were closest to the observed LFMFC locations. All variables listed in Table 1 constituted the initial predictor set considered for LFMFC estimation. Variable elimination and selection were then applied to develop an efficient model that achieves high performance with a reduced number of input variables (Li et al., 2019). During variable elimination, all in-situ LFMFC at each site were first pooled to calculate their correlations with each variable in the initial predictor set, as well as the pairwise correlations among all variables. When the absolute value of the cross-correlation coefficient between a pair of variables exceeded 0.7, only the variable more strongly correlated with in situ LFMFC was retained, following common practice (Konings et al., 2024). For example, if NDVI and VARI exhibited a cross-correlation of $R > 0.7$, only the variable showing the stronger correlation with in situ LFMFC is retained. The variables retained after this step were defined as candidate predictors.

The recursive feature elimination method (Guyon et al., 2002) was then applied to the candidate predictors for variable selection. Specifically, the XGB model was initially fitted 500 times using the candidate predictors, and the average importance of each variable was calculated. Subsequently, the least important variable was iteratively removed from the model until only one variable remained. In each iteration, leave-one-site-out cross-validation (LOCV) (Costa-Saura et al., 2021) was used to evaluate the performance of each model based on the evaluation metrics in Section 3.2. The subset of variables yielding the best LFMFC prediction performance constituted the final predictor set, and the XGB model trained with these variables is hereafter referred to as the LFMFC-XGB model. The *GridSearchCV* method (Scikit-learn developers, 2023) was used to find optimized parameters of the model. For reference, the approach was also implemented using the random forest (RF) model, which has demonstrated strong performance in estimating LFMFC for Mediterranean species (Arcos et al., 2024; Ferrer Palomino et al., 2022). Statistical comparisons between the XGB and RF models were conducted at the site level using paired tests, with the Wilcoxon signed-rank test was applied to assess statistical significance. The XGB and RF models were implemented using the Python packages 'xgboost' (Chen and Guestrin, 2016) and 'scikit-learn' (Pedregosa et al., 2011), respectively.

2.4. Evaluation metrics

For the LFMFC evaluation, the site-level observed LFMFC were decomposed into the site-mean LFMFC ($\overline{LFMFC}$) and the site temporal fluctuations LFMFC ($LFMFC'$). While $\overline{LFMFC}$ only varies spatially, $LFMFC'$ and the observed LFMFC vary both spatially and temporally. Such a decomposition is useful for better understanding the influence of each predictor, as variables related to vegetation spatial distribution are expected to better explain spatial differences, whereas meteorological conditions and vegetation structure are likely more relevant for capturing temporal variation dynamics. The calculations of $\overline{LFMFC}$ and $LFMFC'$ over each site are as follows:

$$\overline{LFMFC}(s) = \left(\sum_{t=1}^n LFMFC(s, t) \right) / n$$

$$LFMFC'(s) = LFMFC(s, t) - \overline{LFMFC}(s)$$

where $LFMFC(s, t)$ is the observed LFMFC measurement at site s on date t , n is the number of the days for which observed LFMFC is available.

Three metrics were used to evaluate the model performance on observed LFMFC, $\overline{LFMFC}$ and $LFMFC'$. The coefficient of determination (R^2) was employed to indicate the model's goodness of fit. The root mean square error (RMSE) quantified the magnitude of prediction errors, while bias was used to assess systematic errors within the model.

$$R^2 = 1 - \frac{\sum_{i=1}^m (Y_{obs} - Y_{pre})^2}{\sum_{i=1}^m (Y_{obs} - \bar{Y})^2}$$

$$RMSE = \sqrt{\left(\sum_{i=1}^m (Y_{obs} - Y_{pre})^2 \right) / m}$$

$$Bias = \left(\sum_{i=1}^m (Y_{pre} - Y_{obs}) \right) / m$$

where Y_{obs} and Y_{pre} are the observed and predicted LFMFC, $\overline{LFMFC}$ or $LFMFC'$. m is the number of Y_{obs} , and $\bar{Y}$ is the average of Y_{obs} .

2.5. Shapley explanation method

To interpret the contribution of each variable in the LFMFC-XGB model, we applied the TreeExplainer framework (Lundberg et al., 2020) in the Shapley Additive exPlanation (SHAP) method (Štrumbelj and Kononenko, 2013). It was designed for the tree-based machine learning method. SHAP values quantify how much each variable contributes to a prediction by comparing the model output with and without the variable across many possible combinations of features. Concretely, they represent differences between the predicted LFMFC value and the expected value, i.e., the mean LFMFC across all observations. For a sample x with M features, the SHAP value can be expressed as follows:

$$f(x) = \varphi_o + \sum_{i=1}^M \varphi_i(x) = \sum_{i=1}^M [\varphi_{i,main}(x) + \varphi_{i,inter}(x)]$$

where $f(x)$ is the predicted LFMFC for sample x , φ_o is the expected LFMFC, and $\varphi_i(x)$ is the SHAP value for feature i . Each $\varphi_i(x)$ can be further decomposed into a main effect ($\varphi_{i,main}(x)$) and interaction effect ($\varphi_{i,inter}(x)$). SHAP values are in the same unit as LFMFC (%). Their magnitude reflects the importance of a variable for the prediction, with larger absolute values indicating a stronger influence (positive or negative). All analyses were performed with the Python 'shap' package (Lundberg and Lee, 2017).

2.6. LFMFC product analyses

For the LFMFC product derived from the LFMFC-XGB model, we further investigated the frequency of LFMFC values falling below a certain threshold across the study region each year. The threshold was adopted from (Pimont et al., 2019), who found that LFMFC begins to have an exponential effect on the fire rate of spread below 67 % in Southern France. Accordingly, we calculated the percentage of area in which the number of days with LFMFC below 67 % exceeded one week per year and compared these results with yearly burned area data. In addition, we further investigated the daily changes in LFMFC in the two weeks prior to two representative fire events in the study region to demonstrate the potential of LFMFC for fire danger prediction.

3. Results

3.1. Variable elimination and selection

The correlation map of all 26 variables listed in Table 1 is shown in Fig. S1. After variable elimination, 14 variables were retained. The relative importance of variables for predicting LFMFC with the model fitted with 14 variables is shown in Fig. 2. Almost half of the variable importance was attributed to weather-related variables, with DC ranking in first place (29.88 % $\pm$ 0.84 %). This is followed by DMC (13.84 % $\pm$ 0.87 %) and smRoot (10.53 % $\pm$ 0.76 %), which together represent substantial secondary contributions. Notree and smSurf had

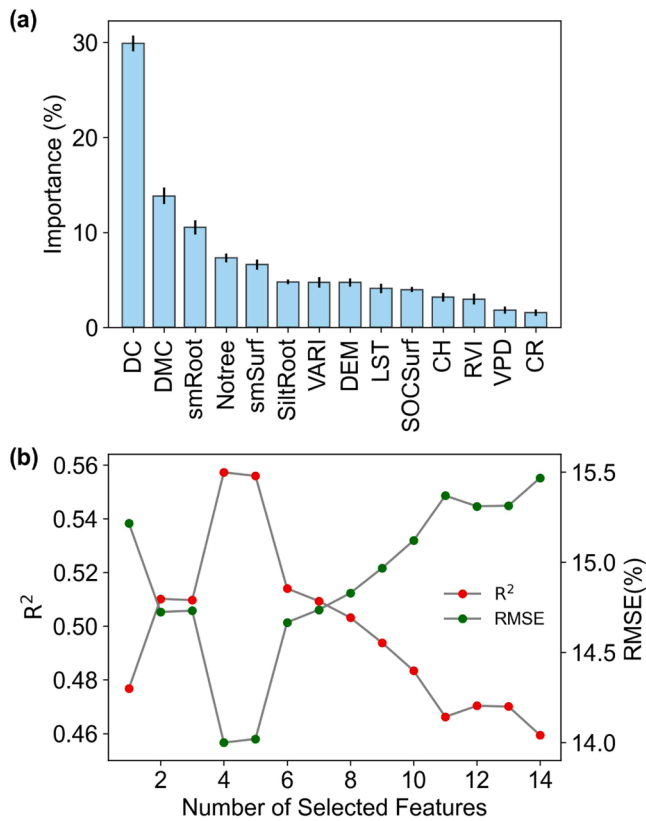


Fig. 2. (a) The average importance of the 14 candidate variables when all are used in the model, and (b) the model performance on observed LFMC with different selected variables.

similar importance, each contributing roughly 7 %. The remaining variables (siltRoot, VARI, DEM, LST, SOCSurf, CH, RVI, VPD, and CR) individually contribute smaller proportions, generally below 5 %, with CR having the smallest importance ($1.54 \% \pm 0.34 \%$).

Based on the rank in Fig. 2 (a), we removed the least important variable iteratively and calculated the model performance on predicting the original LFMC using the LOCV method. Fig. 2 (b) shows, increasing the number of variables initially improved the model performance up to a point beyond which model performance decreased. The best performance was obtained with four variables: DC, DMC, smRoot and Notree. This model was retained to estimate LFMC and hereafter referred as LFMC-XGB model.

Changes in model performance on site means ($\overline{LFMC}$) and temporal fluctuations ($LFMC'$) as a function of the selected features are shown in Fig. 3. The best performance in estimating $\overline{LFMC}$ was achieved when four selected variables were used (Fig. 3 (a)). Although the model performed best on $LFMC'$ with 13 variables (Fig. 3 (b)), its performance on observed LFMC and $\overline{LFMC}$ was much lower (Figs. 2 and 3 (a)). Notably, increasing the number of variables from three to four, by adding the Notree variable, significantly improved the model's performance on $\overline{LFMC}$, confirming the LFMC-XGB (with 4 variables) was the optimal model.

3.2. Spatiotemporal performance of LFMC-XGB model

Fig. 4(a)–(c) shows the scatter plot between predicted LFMC and observed LFMC values using the LFMC-XGB model. Fig. 4(a) shows the performance of LFMC-XGB model on observed LFMC data, with $R^2 = 0.56$, RMSE = 14.00 %, and bias = -0.18 %. The LFMC-XGB model predicted temporal fluctuations with an $R^2 = 0.58$ (Fig. 4(b)), while its performance in predicting site means was slightly lower, with an $R^2 =$

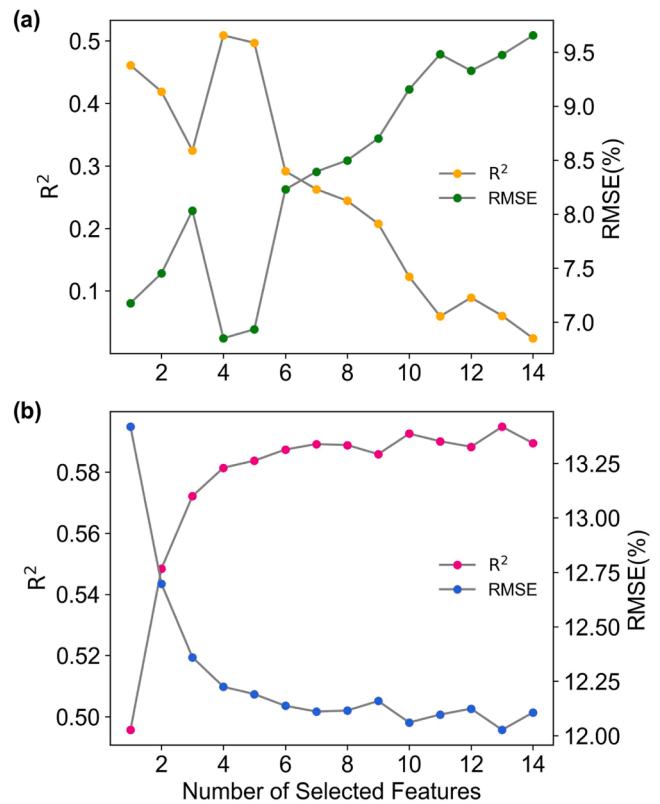


Fig. 3. The model performance on site-mean observed LFMC ($\overline{LFMC}$) (a) and the temporal fluctuations LFMC ($LFMC'$) with different selected variables (b).

0.51 (Fig. 4(c)). This indicates that the model captured temporal variation in LFMC slightly better than spatial variation. The detailed results for each site (Fig. 4(d)) can explain this. Of the 31 sites, 27 exhibited a temporal correlation (R) greater than 0.7 between predicted and observed LFMC values, while the coefficient of determination (R^2) exceeded 0.5 at 15 sites. Two sites exhibited very low R^2 values (< 0.3), while three other sites showed negative R^2 values, indicating poor model performance at these locations.

Two representative sites, one with the best quality and one with the worst, were selected to illustrate the detailed comparison between predicted and observed LFMC time series, as presented in Fig. 5. At site D06S3a, the LFMC-XGB model accurately captured both dynamics ($R = 0.90$) and magnitudes of LFMC, with predicted values closely matching measurements (bias = 1.45 %), resulting in a high R^2 equal to 0.83. In contrast, at site D2AS3, although the model captured the overall LFMC dynamics reasonably well ($R = 0.83$), it significantly overpredicted values in certain years, leading to a large bias (13.34 %) and a negative R^2 . Additionally, we found that the LFMC-XGB model struggled to capture large differences between two consecutive LFMC observations in those two sites.

The site-level comparison of the performance of the LFMC-XGB and LFMC-RF models in predicting observed LFMC and $LFMC'$ is summarized in Table 2. The LFMC-RF model also incorporated four input variables, namely DC, DMC, smRoot, and Notree. Across the 31 sites, LFMC-XGB model achieved significantly higher R^2 (0.44 ± 0.23) than LFMC-RF (0.36 ± 0.26 ; Wilcoxon signed-rank test, p value = 0.0008) on observed LFMC. Similarly, RMSE was significantly lower for LFMC-XGB (13.51 ± 4.23 %) model compared to LFMC-RF (14.25 ± 4.23 %; p value = 0.0016) model. In contrast, no significant difference was observed between the two models in terms of bias (p value = 0.247). Consistent results were obtained $LFMC'$. The LFMC-XGB model also showed significantly higher R^2 (0.58 ± 0.17 vs. 0.52 ± 0.18 ; p value = 0.0003) and lower RMSE (11.76 ± 3.88 % vs. 12.43 ± 3.59 %; p value =

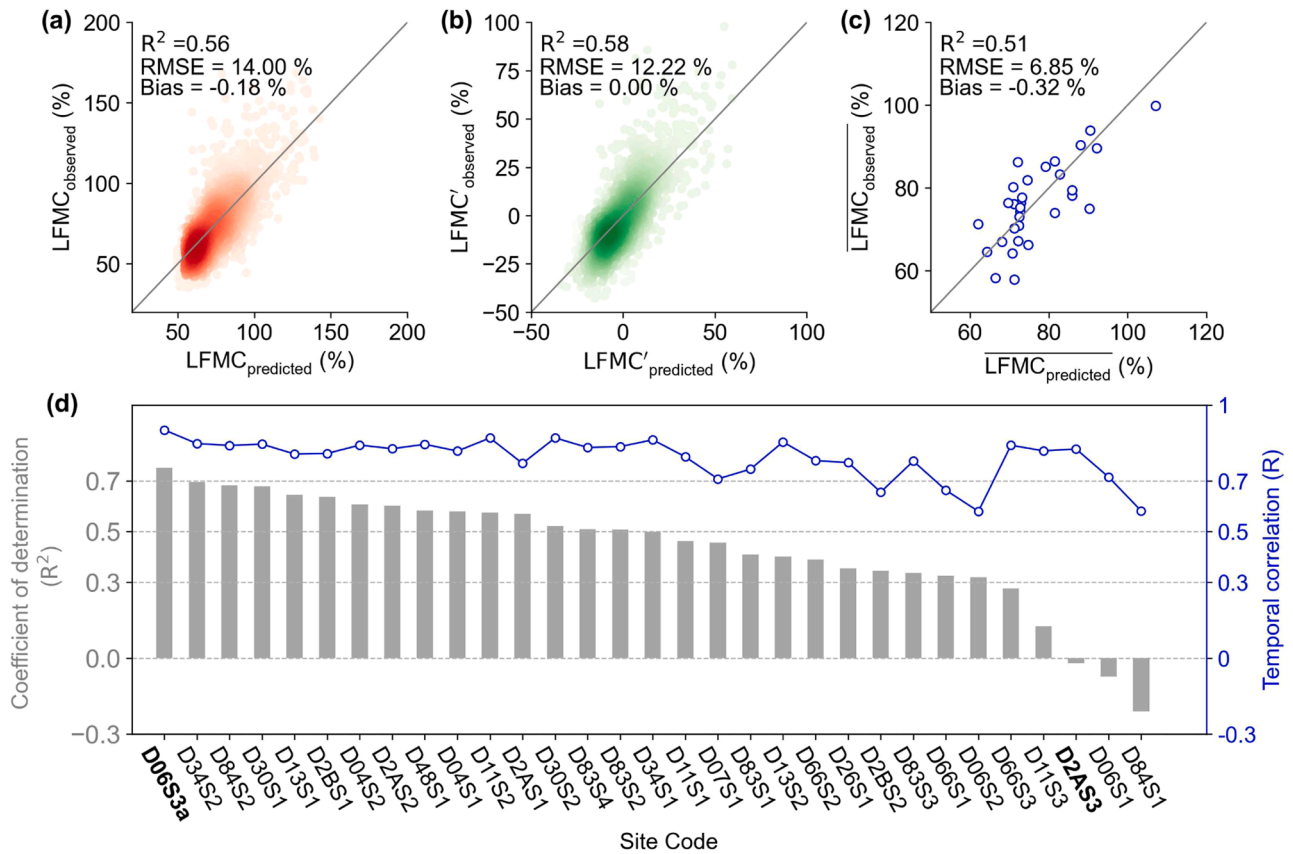


Fig. 4. Model performance based on selected variables for the (a) The observed LFMCM data, (b) temporal fluctuations of observed LFMCM data per site, (c) site means observed LFMCM data, and (d) the temporal correlation (R) and the coefficient of determination (R^2) between the predicted and observed LFMCM over each site. The changes in the color in (a) and (b) indicate the point density changes. The two sites with bold site names were used to illustrate the seasonal variations in LFMCM (time series in Fig. 5).

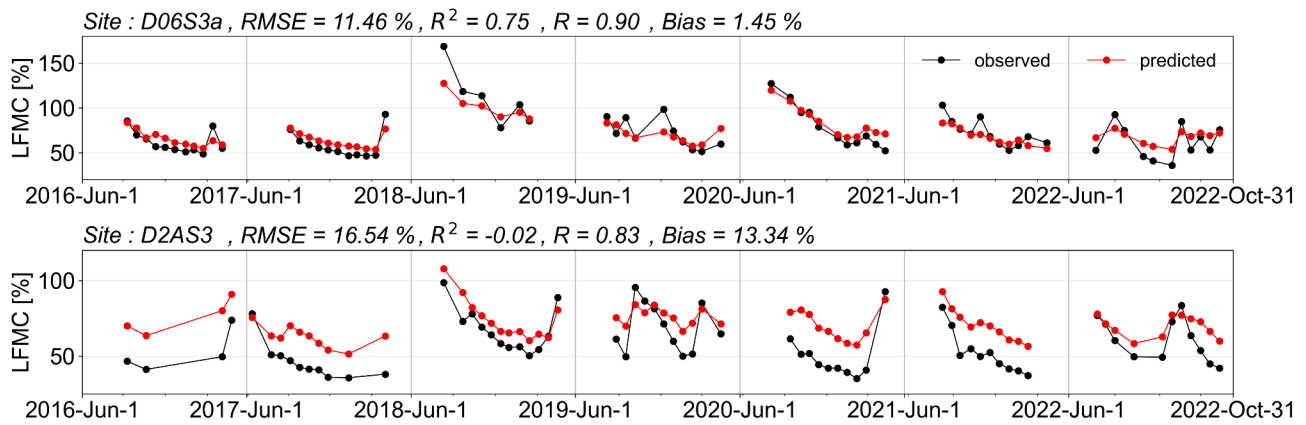


Fig. 5. Time series of predicted and observed LFMCM at sites D06S3a and D2AS3 in the period 2016-2022.

Table 2

Site-level comparison of model performance between LFMCM-XGB-X and LFMCM-RF model on observed LFMCM and $LFMCM'$.

LFMCM representation	Metric	LFMCM-XGB (Mean $\pm$ SD)	LFMCM-RF (Mean $\pm$ SD)	Wilcoxon_pvalue	Paired_t_pvalue
observed LFMCM	R^2	0.44 ± 0.23	0.36 ± 0.26	0.0008	0.0012
	RMSE (%)	13.51 ± 4.23	14.25 ± 4.23	0.0016	0.0025
	Bias (%)	-0.32 ± 6.96	-0.45 ± 7.40	0.2474	0.6751
$LFMCM'$	R^2	0.58 ± 0.17	0.52 ± 0.18	0.0003	0.0003
	RMSE (%)	11.76 ± 3.88	12.43 ± 3.59	0.0004	0.0004
	Bias (%)	0.00 ± 0.00	0.00 ± 0.00	0.6636	0.7167

Notes: Statistics are reported as mean $\pm$ standard deviation (SD) across sites ($n = 31$).

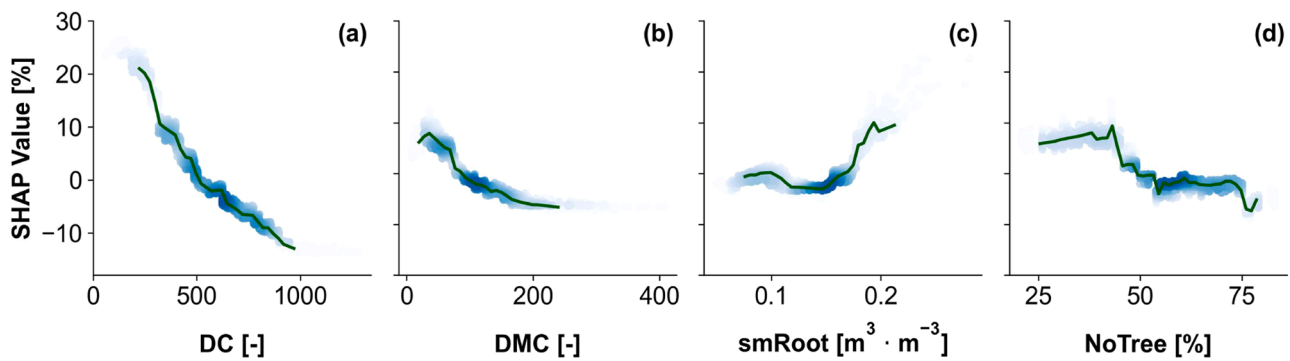


Fig. 6. Impacts of the drivers on LFM indicated by SHAP value. The color variations in panels (a)–(d) indicate changes in point density. The colored solid lines in panels (a)–(d) represent the SHAP value, computed as the median values within 50 bins of the x-axis variables. Binned medians are displayed only when the sample size within a given bin exceeds 20.

0.0004) than the LFM-RF model, while no significant difference in bias was observed (p value = 0.664).

3.3. Analysis of LFM drivers in LFM-XGB with SHAP value

To provide more detailed insights into the contribution of individual variable in LFM variability, the SHAP value of each variable was computed. Fig. 6 provides the SHAP value of the different variables across each sample. Overall, the SHAP value of DC had a wider range (–13.78 to 26.37 %) than those of the other three variables. Although DMC was the second most important variable, its SHAP value range was the narrowest (–7.29 to 12.19 %). The sample with the highest SHAP value of smRoot could reach 22.61 %, while there were many samples with SHAP values near zero.

Regarding the SHAP value trends, the SHAP values of DC and DMC (Fig. 6a & b) decreased as corresponding variable values increased,

indicating that these two drought indices had a negative effect on the LFM changes, i.e., an increasing drought decreased the LFM value. When the value of DC and DMC exceeded a certain threshold, their SHAP value leveled off. The SHAP value changes of Notree (Fig. 6c) and smRoot (Fig. 6d) showed more complex patterns than those for DC and DMC. When Notree was between 20–40 %, its SHAP values exhibited relatively limited variation, suggesting a weak and stable influence of this variable within this range. However, when Notree reached about 40–55 %, its SHAP values quickly decreased from 10 % to –3 %, showing that LFM changes were very sensitive to the Notree percentage change. After that, the SHAP values of Notree did not change much until the Notree value reached around 70 %, when it dropped again. Regarding the SHAP values of smRoot, they remained relatively flat when smRoot was around 0.11 to 0.15 $\text{m}^3 \cdot \text{m}^{-3}$, with values hovering around 0 %, which indicates that the change was very small. When smRoot exceeded 0.15 $\text{m}^3 \cdot \text{m}^{-3}$, its SHAP values increased sharply, rising rapidly from

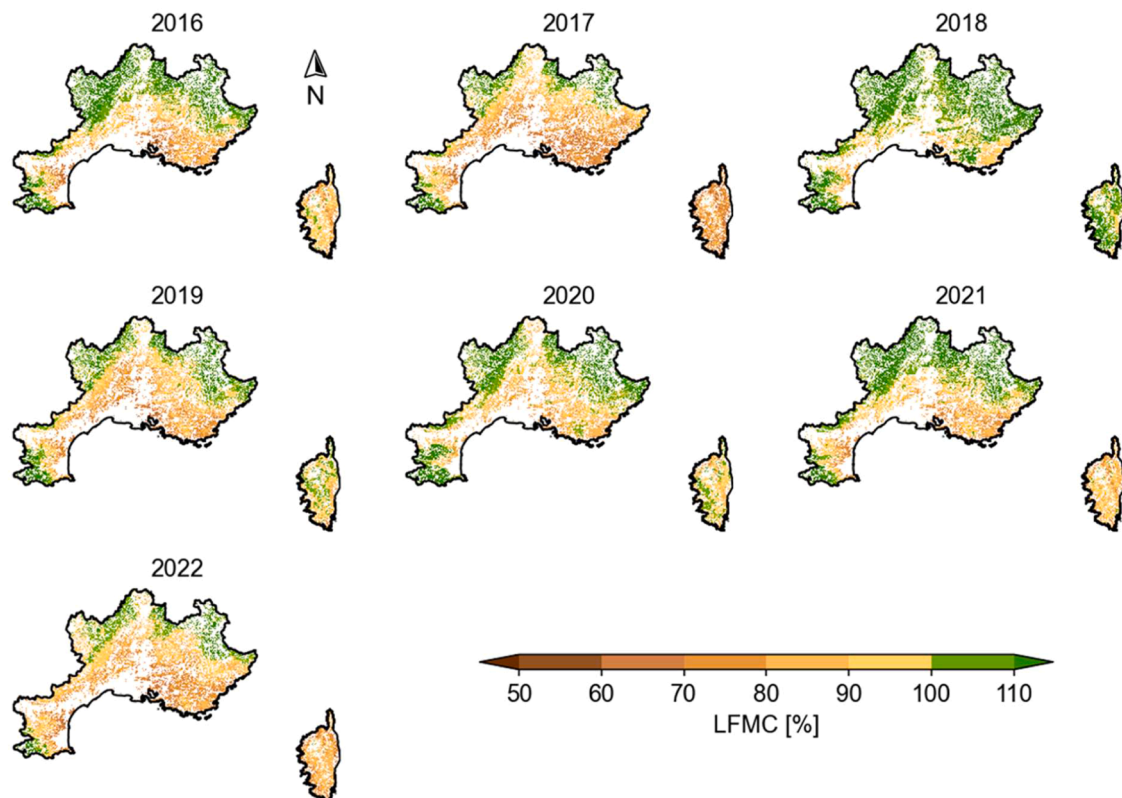


Fig. 7. Yearly averaged LFM from 2016 to 2022 (June to September) in the French Mediterranean region. White areas indicate no valid predicted LFM data.

nearly 0 % to approximately 20 % or higher, indicating higher LFCM values when smRoot is high.

Additionally, we computed the differences between the SHAP value and the SHAP main effect value for each variable, in order to evaluate the potential role of interactions (Fig. S2). The median value of the differences was close to 0 %, and the range was between -2 and 2 %. The DMC, smRoot and Notree had a wider range, and DC had the smallest range. The Mann-Whitney U test revealed a significant difference between the SHAP value and the SHAP main effect value for smRoot and Notree, indicating that the sum of interactions between smRoot or Notree and the other three variables also influenced changes in LFCM. However, for the DC and DMC, no difference was observed between the SHAP value and the SHAP main effect value, indicating that the overall contribution of interactions was negligible.

3.4. LFCM mapping and analysis

In light of the good performance of the LFCM-XGB model, we generated daily LFCM maps for French Mediterranean region from June to September between 2016 and 2022 at a 250 m resolution. Fig. 7 presents the yearly averaged LFCM maps at different years. Spatially, regardless of the year, we observed a clear latitudinal gradient, with the lower LFCM values located in the south of the study area, while higher LFCM values were found in the northern part. Temporally, we noticed that lower LFCM values were observed during the years 2017 and 2022.

Fig. 8 (a) shows the percentage of the study area where LFCM remained below 67 % for more than one week in each year, together with the corresponding annual burned area. With the exception of 2018, when only 7.97 % of the region experienced prolonged low LFCM conditions, more than 25 % of the study area was affected each year. Burned areas exceeding 100 km² occurred in all years except 2018–2020. Years characterized by a larger spatial extent of prolonged low LFCM generally corresponded to higher burned areas; however, this relationship was not systematic across all years, as illustrated by the mismatch observed during 2019–2020. These results suggest that prolonged periods of low LFCM reflect background drought conditions conducive to fire activity, while emphasizing that burned area is also strongly influenced by additional factors such as ignition occurrence, extreme fire weather, and fire suppression.

Fig. 8 (b) shows the time series of daily LFCM change in two representative fire events occurring in 2020 and 2022: one fire occurring during a year with extensive low LFCM conditions, and one fire occurring during a year with comparatively lower burned area despite similar drought conditions. In both cases, LFCM exhibited a clear decline within the burned area perimeters during the weeks preceding fire occurrence, approaching or falling below the 67 % threshold commonly associated with increased fire susceptibility. This indicates increasing fuel dryness prior to fire occurrence and highlights the relevance of LFCM as a background condition influencing fire activity.

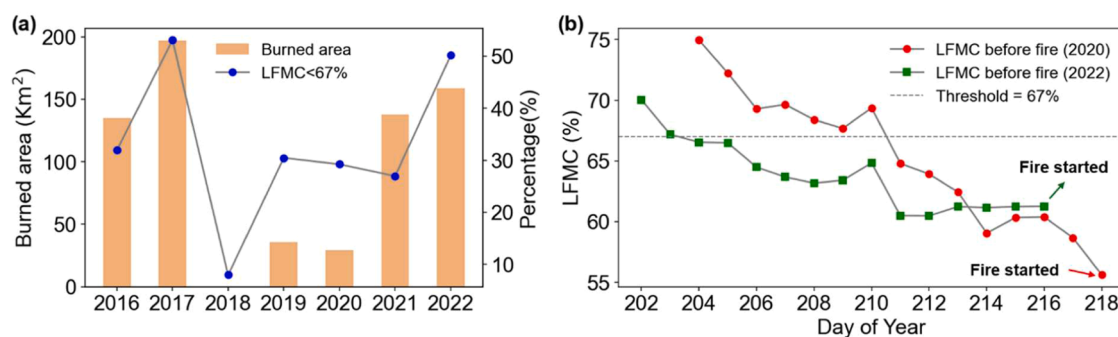


Fig. 8. (a) The summed burned area from June to September (left axis) and the percentage of the area with LFCM lower than 67% for more than one week (right axis), (b) LFCM time series before fire (14 days).

4. Discussion

The LFCM-XGB model was developed for mapping LFCM over the French Mediterranean under summer conditions in this study. The model showed a good agreement compared with the in-situ LFCM measurements. Moreover, the LFCM product indicated that the areas of low LFCM correspond well to regions with higher fire activity (Pimont et al., 2022), and years with intense fire activity, such as 2017 and 2022 in France, were associated with lower predicted LFCM values. The evaluation of the LFCM-XGB model provide confidence in (i) the insights gained into the main factors influencing LFCM and their responses, and (ii) the model's spatial performance, despite certain limitations. These points are elaborated below, followed by future perspectives.

4.1. Drivers of LFCM in the French Mediterranean

Three of the four parameters used in the LFCM-XGB model were related to the soil drought conditions. This strong influence can be attributed to the characteristics of the study area, where the Mediterranean summer features very high temperatures and low rainfall (Ruffault et al., 2020; Trambly et al., 2020). During this period, vegetation primarily relies on soil moisture to sustain plant hydration; consequently, soil moisture availability plays a key role in shaping LFCM dynamics. However, the influence of soil-water dynamics on LFCM is mediated by species-specific physiological and anatomical traits along the soil-plant-atmosphere continuum. Soil moisture exhibits distinct dynamics at different depths, and species with shallow root systems generally experience stronger LFCM fluctuations than species capable of accessing deeper water pools.

Our results indicate that DC, the long-term drought variable, had higher importance than the short- to medium-term drought parameter DMC, suggesting that cumulative soil drying exerts a dominant control on LFCM during summer. Although smRoot from SMAP reflects soil moisture conditions in deeper layers, its relatively lower contribution compared to DMC may be related to the generally shallow rooting systems of Mediterranean shrubs species, which limit access to deep soil water (Stocker et al., 2023; Tumber-Davila et al., 2022). In addition, stomatal regulation plays a critical role in modulating LFCM responses to atmospheric conditions. While a previous study has shown that VPD, an atmospheric dryness parameter, can explain LFCM variability (Griebel et al., 2023), VPD was not retained by the LFCM-XGB model in this study. This may reflect the focus on the summer period, when stomatal closure reduces plant sensitivity to atmospheric demand and limits the direct influence of VPD on LFCM (Martin-StPaul et al., 2023). Together, these results highlight that LFCM variability reflects not only soil-water dynamics at different depths, but also species-specific rooting strategies and stomatal control mechanisms.

Among all remote sensing-based variables introduced during the model-building phase, only the Notree variable, which serves as a proxy for shrubland cover, was selected when using a stringent cross-

validation method (LOCV). Results show that Notree mostly improved the spatial performance of the model (site means). The effect of Notree we observed in this study is consistent with the fact that LFMC tends to be lower in shrublands (high Notree values) than in forested areas (low Notree values). Although plausible, the effect of the Notree variable has not, to the best of our knowledge, been previously reported and should be further investigated in future studies.

Although several sites showed strong temporal correlations between in-situ LFMC measurements and other remote sensing products, such as those derived from Sentinel-2, our results also highlighted that these relationships did not hold when considering all sites together, thereby greatly limiting their potential use for capturing regional spatio-temporal patterns of LFMC. For example, regarding the vegetation indices derived from Sentinel-2, VARI exhibited the strongest correlation with LFMC, consistently with [Fan et al. \(2018\)](#), who found MODIS-derived VARI had the highest correlation with LFMC in the 'Réseau Hydrique' dataset. However, it was not selected in the LFMC-XGB model, as the LOCV method when selecting the variables tends to eliminate variables that exhibit poor correlations on some sites (**Table S2**). Microwave variables derived from Sentinel-1 also failed to contribute to the LFMC estimation in this study, as two microwave vegetation parameters from S1 exhibited a very weak relationship with in-situ LFMC, both at individual sites (**Table S3 and S4**) and when combining all sites (**Fig. S1**). This contrasted with findings from a study conducted in forest regions of the United States ([Rao et al., 2020](#)) and Germany ([Kranz et al., 2025](#)), but aligned with results reported for shrubland in the Madrid Region in Spain ([Tanase et al., 2022](#)). Because microwave observations are sensitive to both vegetation water content and structural characteristics ([Jackson and Schmugge, 1991](#); [Ulaby, 1986](#)), their contribution to LFMC estimation varies with vegetation type. In our study, we averaged the LFMC from different types of shrub species in each site, which makes the use of microwave variables especially challenging.

4.2. Uncertainty in the LFMC product

The uncertainties associated with the proposed LFMC product are mainly related to (i) the incorporation of the Notree variable, (ii) limitations inherent to drought indices, (iii) vegetation heterogeneity and species aggregation effects and (iv) constraints associated with machine learning-based approaches.

The Notree variable was used to represent the spatial proportion of shrubland within each grid cell, corresponding to the vegetation type sampled for LFMC. However, Notree has a larger extent than in situ plots and may encompass other non-shrubland vegetation types such as grassland or cropland. Information on whether such vegetation types in Notree grid cells was not explicitly available. Consequently, a lower actual proportion of shrubland within the Notree class may lead to greater uncertainty in LFMC estimation. Unfortunately, high-resolution vegetation cover products are usually only available for specific years and lack temporal continuity ([Zanaga et al., 2021](#)). To compute regional maps, the LFMC-XGB model was also applied to the forest region, where shrubland commonly forms the understory and acts as a primary source of fire spread. The Notree variable does not capture understory shrubland, and the absence of calibration sites within the understory increased the uncertainty of LFMC estimation data, which could potentially exhibit some bias. Additional LFMC data for understory vegetation are currently being collected by the French Forest Service to clarify this aspect. Notree also influences rain interception and the shading effect on VPD, which were not considered in the LFMC-XGB model and may introduce uncertainty into the LFMC product.

The second uncertainty of the LFMC product comes from the coarse-resolution drought indices. In this study, 8 km resolution meteorological data were downscaled to 250 m to match the Notree dataset and subsequently used to calculate DC and DMC for generating the LFMC product. This approach raises concerns regarding the effective spatial

resolution of the final LFMC product, which may not fully capture high-resolution spatial variability, especially in mountainous regions where meteorological conditions are highly heterogeneous but were represented at a coarse resolution. Although it is possible to generate a 250 m resolution LFMC product relying solely on remote sensing variables, model performance in this configuration is notably poorer (**Table S5**). Ongoing efforts by the French meteorological administration to develop fine-resolution (1.2 km) meteorological datasets are expected to improve LFMC mapping accuracy in future applications. Beyond input-data resolution, an additional source of uncertainty is related to the use of drought indices as the primary predictors of LFMC. While DC and DMC efficiently integrate atmospheric water demand and soil moisture proxies, they do not explicitly account for species-specific physiological adjustments, plant water-use strategies, leaf traits, or phenological variations in dry biomass. Consequently, the LFMC-XGB model inherits the intrinsic limitations of these drought indices, despite the improved predictive skill provided by machine learning and other explanatory variables. This limitation is particularly relevant in multispecific or functionally diverse ecosystems, where contrasting drought-response strategies (e.g., isohydric versus anisohydric behavior, shallow versus deep rooting systems) may lead to asynchronous LFMC dynamics that cannot be fully captured by weather-derived indices alone.

The third source of uncertainty arises from the use of site-level LFMC values averaged across different species. A total of 19 species were sampled in this study; however, most were represented by limited sample sizes and occurred at only one or two study sites, making the development of species-specific models infeasible. However, to assess the influence of vegetation diversity on model performance, we built specific LFMC-XGB models based on the functional type. All sampled species can be broadly classified into three plant functional types: seeders, resprouters, and trees. Seeder species, which are generally shallow-rooted and more tightly coupled to atmospheric drought, tend to exhibit LFMC dynamics that are more directly related to drought indices. In contrast, trees and resprouter species often access deeper water pools and exhibit stronger hydraulic buffering, resulting in weaker and more asynchronous LFMC responses to climatic drought. The performance of three functional-type LFMC-XGB models for seeder, resprouter, and tree species was compared with that of the site-averaged model at the site level (**Table S7 & S8**). Model performance tended to be higher in monospecific or functionally homogeneous sites, whereas multispecific sites combining contrasting drought-response strategies exhibited more variable, and in some cases lower, predictive skill. The poorest performance observed at site D84S1 illustrates this limitation, as this site combines seeder shrubs, resprouter shrubs, and trees representing distinct plant functional types with asynchronous LFMC dynamics. A similar effect is observed at site D11S3, which combines seeder and resprouter species with contrasting drought-response strategies. Such functional heterogeneity increases uncertainty when LFMC is aggregated at the site level. In contrast, the low predictive skill at D66S3, despite being dominated by two seeder species (*Cistus monspeliensis* and *Erica arborea*), suggests that species-level differences within the same functional type and site-specific factors may also contribute to LFMC prediction uncertainty. Overall, 51.6 % of the sites (16 out of 31) were better predicted using one of the functional-type models than using the site-averaged model. Although functional-type or species-specific models may improve the performance of the prediction and are ecologically relevant, species-level LFMC modeling remains challenging for regional-scale mapping applications due to the limited spatial coverage of LFMC observations and the lack of spatially explicit, temporally consistent species distribution data. We therefore adopted a site-averaged LFMC approach to produce a spatially continuous product suitable for regional analyses and operational applications.

This study represents, to our knowledge, the first application of the XGB model for LFMC estimation. Since machine learning methods are data-driven, the model's performance is inherently influenced by the dataset used. This study considered a broad set of variables to identify

those most suitable for LFMC estimation, but not all variables from different domains were incorporated. In addition, the LFMC-XGB model developed in this study exhibited an underestimation for high LFMC values, which was caused by the limited number of samples used in the model. However, it is not a major issue in practice, as the primary interest lies in the lower range (Benali et al., 2025; Marino et al., 2020). The results showed that LFMC-XGB demonstrated better performance than LFMC-RF in predicting LFMC in the French Mediterranean area.

4.3. Future application and improvement

A first and straightforward application of the LFMC product developed in this study would be its integration into regional operational and research fire danger models. When compared with the 'Vegetation Drought Sensitivity Index' (NSV2) index (Durand et al., 2023), which served as the operational proxy for LFMC in French fire agencies, LFMC-XGB outperformed the operational model (Table S6). Quan et al. (2021a) demonstrated that incorporating LFMC into the fire risk model further improves its accuracy in China. As previous fire danger assessments in the French Mediterranean area did not explicitly account for LFMC (Castel-Clavera et al., 2025; Castel-Clavera et al., 2022), evaluating the potential of the new LFMC product to enhance fire danger predictions would be a valuable next step. Another potential application would be to investigate how LFMC could change in the study area under different future climate scenarios (Rogelj et al., 2012; Ruffault et al., 2023; van Vuuren et al., 2011) using a simplified LFMC-XGB model, which only accounts for DC and DMC, excluding remote sensing variables that are challenging to project.

Regarding future enhancements, beyond the use of refined weather-derived data, the first aspect is related to soil moisture and its link to vegetation moisture. A potential future improvement is the installation of soil moisture sensors at LFMC sampling sites, which would allow recalibration of DC and DMC in line with Canadian Forest Service recommendations (Lawson and Armitage, 2008). In this study, due to the lack of in-situ soil moisture data, we used the default settings for the calculation, as most previous studies did (Ruffault et al., 2018). A fine calibration may improve the accuracy of DC and DMC, resulting in a higher-quality LFMC product. Moreover, a potential temporal lag may exist between soil and vegetation moisture and has not been considered in the present study. Jia et al. (2019) found that there is a two-month lag between SMAP Level 3 enhanced surface soil moisture and LFMC measurements in Southern California USA. Vinodkumar et al. (2021) developed a Fourier cosine series that accounts for the two-week delay between soil moisture from the Joint United Kingdom Land Environment Simulator land surface model and LFMC to predict LFMC over Australia. The time lag between soil moisture and LFMC in the study region should also be considered, as it can help characterize drought conditions at different time scales.

Other improvements might come from exploring a better use of thermal data in the model. Indeed, although LST was not used in the final model, it had a relatively high correlation with observed LFMC ($R = -0.28$) among the variables that were not selected during the variable selection process. This could be addressed through modeling evapotranspiration, which can be retrieved from thermal data (Penot and Merlin, 2023). In addition, as more and more sites are incorporated in the "Réseau Hydrique" database, it is possible to integrate species-based models into the regional mapping framework. Finally, insights from ecophysiological modelling could be incorporated (Jolly et al., 2025). For example, the SurEau-Ecos-FMC model has been shown to simulate LFMC effectively in the Mediterranean region at the species level (Ruffault et al., 2023). It would be interesting to test the model at larger spatial scales and explore integrating process-based model simulations and different datasets into machine-learning frameworks to further improve LFMC estimations.

5. Conclusion

This study integrated meteorological and remote sensing data with advanced machine learning methods to develop a new LFMC product for the French Mediterranean region. We found that combining DC, DMC, smRoot, and Notree in the LFMC-XGB model yielded the best performance for simulating LFMC ($R^2 = 0.56$, RMSE = 14.00 %, bias = -0.18 %). The model also performed well in predicting site-mean LFMC and the site-temporal fluctuations LFMC. Our results indicated that the LFMC changes in the French Mediterranean region are mainly driven by the soil drought conditions, especially the long-term drought variable. The inclusion of the Notree variable allowed the LFMC-XGB model to attain spatial performance, measured by site means, comparable to its performance on temporal fluctuations. Importantly, the LFMC products generated in this study demonstrated potential for operational fire danger assessment and for improving understanding of the drivers of burned areas and extreme fire seasons.

CRedit authorship contribution statement

Xiangzhuo Liu: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Nicolas Martin-StPaul:** Writing – review & editing, Resources, Formal analysis. **Julien Ruffault:** Writing – review & editing, Methodology, Formal analysis. **Gengke Lai:** Writing – review & editing, Methodology. **Kevyn Raynal:** Writing – review & editing, Data curation. **RodriguezSuquet Raquel:** Writing – review & editing, Data curation. **Huan Wang:** Writing – original draft. **Russell Parsons:** Writing – review & editing, Funding acquisition. **Jean-Luc Dupuy:** Writing – review & editing, Supervision, Funding acquisition. **Francois Pimont:** Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgments

The re-aggregated and filtered Sentinel-2 reflectance data were provided by CNES as part of the SCO ALEOFEU project. We want to thank for the support from US DoD Strategic Environmental Research and Development Program (SERDP), RC20-1025 Closing Gaps project, through Forest Service Agreement 20-IJ-11221637-178. Finally, special thanks for the French Forest Service (ONF) for collecting and sharing the Réseau Hydrique data, funded by Ministry of Agriculture. The authors would like to thank the reviewers for their constructive comments and suggestions on the original manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.agrformet.2026.111093](https://doi.org/10.1016/j.agrformet.2026.111093).

Data availability

The LFMC data for the study region for 2016 to 2022 can be downloaded from: <https://figshare.com/s/29d1c58295d2577103cc>

References

- Amraoui, M., et al., 2013. Fire activity over Mediterranean Europe based on information from Meteosat-8. *For. Ecol. Manag.* 294, 62–75.
- Arcos, M.A., Balaguer-Beser, A., Ruiz, L.A., 2024. Evaluating the performance of spectral indices and meteorological variables as indicators of live fuel moisture content in Mediterranean shrublands. *Ecol. Indic.* 169.

- Arganaraz, J.P., et al., 2016. Estimation of live fuel moisture content from MODIS images for fire danger assessment in Southern Gran Chaco. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 9 (12), 5339–5349.
- Benali, A., et al., 2025. A near-real-time operational live fuel moisture content (LFMC) product to support decision-making at the national level. *Fire* 8 (5).
- Borisov, V., et al., 2024. Deep neural networks and tabular data: a survey. *IEEE Trans. Neural Netw. Learn. Syst.* 35 (6), 7499–7519.
- Castel-Clavera, J., et al., 2025. A comparative analysis of fire-weather indices for enhanced fire activity prediction with probabilistic approaches. *Agric. For. Meteorol.* 361.
- Castel-Clavera, J., et al., 2022. Disentangling the factors of spatio-temporal patterns of wildfire activity in south-eastern France. *Int. J. Wildland Fire*.
- Castro, F., Tudela, A., Sebastià, M.T., 2003. Modeling moisture content in shrubs to predict fire risk in Catalonia (Spain). *Agric. For. Meteorol.* 116 (1–2), 49–59.
- Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 785–794.
- Chen, X., et al., 2022. A long-term reconstructed TROPOMI solar-induced fluorescence dataset using machine learning algorithms. *Sci. Data* 9 (1), 427.
- Chowdhury, S., Lin, Y., Liaw, B., Kerby, L., 2022. Evaluation of tree based regression over multiple linear regression for non-normally distributed data in battery performance. In: *2022 International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*, pp. 17–25.
- Chuvieco, E., et al., 2004. Combining NDVI and surface temperature for the estimation of live fuel moisture content in forest fire danger rating. *Remote Sens. Environ.* 92 (3), 322–331.
- Costa-Saura, J.M., Balaguer-Beser, Á., Ruiz, L.A., Pardo-Pascual, J.E., Soriano-Sancho, J. L., 2021. Empirical models for spatio-temporal live fuel moisture content estimation in mixed mediterranean vegetation areas using sentinel-2 indices and meteorological data. *Remote Sens.* 13 (18).
- Cunill Camprubi, A., González-Moreno, P., Resco de Dios, V., 2022. Live fuel moisture content mapping in the Mediterranean Basin using random forests and combining MODIS spectral and thermal data. *Remote Sens.* 14 (13).
- Cunningham, C.X., Williamson, G.J., Bowman, D.M.J.S., 2024. Increasing frequency and intensity of the most extreme wildfires on Earth. *Nat. Ecol. Evol.*
- Danielson, J.J., Gesch, D.B., 2011. Global multi-resolution terrain elevation data 2010 (GMTED2010). US Department of the Interior, US Geological Survey.
- Dennison, P.E., Moritz, M.A., 2009. Critical live fuel moisture in chaparral ecosystems: a threshold for fire activity and its relationship to antecedent precipitation. *Int. J. Wildland Fire* 18 (8), 1021.
- Dennison, P.E., et al., 2003. Modeling seasonal changes in live fuel moisture and equivalent water thickness using a cumulative water balance index. *Remote Sens. Environ.* 88 (4), 442–452.
- DiMiceli, C., et al., 2015. MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V006. NASA EOSDIS Land Processes DAAC.
- DiMiceli, C., Sohlberg, R., Townsend, J., 2022. MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V061. NASA EOSDIS Land Processes Distributed Active Archive Center.
- Dimitrakopoulos, A.P., Bemmerzouk, A.M., 2003. Predicting live herbaceous moisture content from a seasonal drought index. *Int. J. Biometeorol.* 47 (2), 73–79.
- Dupuy, J.-L., et al., 2020. Climate change impact on future wildfire danger and activity in southern Europe: a review. *Ann. For. Sci.* 77 (2).
- Durand, J.M., Piveteau, V., Cannard, P., Leuret, C. and Mortier, F., 2023. Politique de prévention et de lutte contre l'incendie de forêt dans un contexte d'extension et d'intensification du risque dû au changement climatique.
- Dymond, S.F., et al., 2017. Topographic, edaphic, and vegetative controls on plant-available water. *Ecohydrology* 10 (8).
- Fan, L., et al., 2018. Evaluation of microwave remote sensing for monitoring live fuel moisture content in the Mediterranean region. *Remote Sens. Environ.* 205, 210–223.
- Ferrer Palomino, A., Sanchez Espino, P., Borrego Reyes, C., Jimenez Rojas, J.A., Rodriguez, Y.S.F., 2022. Estimation of moisture in live fuels in the mediterranean: linear regressions and random forests. *J. Env. Manage.* 322, 116069.
- Gao, B.-c., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* 58 (3), 257–266.
- Geng, L., Che, T., Ma, M., Tan, J., Wang, H., 2021. Corn biomass estimation by integrating remote sensing and long-term observation data based on machine learning techniques. *Remote Sens.* 13 (12).
- Giglio, L., Justice, C., Boschetti, L., Roy, D., 2021. MODIS/Terra+Aqua Burned Area Monthly L3 Global 500m SIN Grid V061, Editor. N.L.P.D.A.A. Center.
- Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* 80 (1), 76–87.
- Griebel, A., et al., 2023. Specific leaf area and vapour pressure deficit control live fuel moisture content. *Funct. Ecol.* 37 (3), 719–731.
- Grinstajn, L., Oyallon, E., Varoquaux, G., 2022. Why do tree-based models still outperform deep learning on typical tabular data? *Adv. Neural Inf. Process. Syst.* 35, 507–520.
- Grossiord, C., et al., 2020. Plant responses to rising vapor pressure deficit. *New. Phytol.* 226 (6), 1550–1566.
- Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using support vector machines. *Mach. Learn.* 46 (1–3), 389–422.
- Houborg, R., Soegaard, H., Boegh, E., 2007. Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using Terra and Aqua MODIS reflectance data. *Remote Sens. Environ.* 106 (1), 39–58.
- Huang, Y., Walker, J.P., Gao, Y., Wu, X., Moneris, A., 2016. Estimation of vegetation water content from the radar vegetation index at L-band. *IEEE Trans. Geosci. Remote Sens.* 54 (2), 981–989.
- Huete, A., et al., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83 (1–2), 195–213.
- Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25 (3), 295–309.
- Jackson, T.J., Schmugge, T.J., 1991. Vegetation effects on the microwave emission of soils. *Remote Sens. Environ.* 36 (3), 203–212.
- Jia, S., Kim, S.H., Nghiem, S.V., Kafatos, M., 2019. Estimating live fuel moisture using SMAP L-band radiometer soil moisture for Southern California, USA. *Remote Sens.* 11 (13).
- Jia, Z., et al., 2024. Grassland biomass spatiotemporal patterns and response to climate change in eastern Inner Mongolia based on XGBoost model estimates. *Ecol. Indic.* 158.
- Jolly, W.M., Conrad, E.T., Brown, T.P., Hillman, S.C., 2025. Combining ecophysiology and combustion traits to predict conifer live fuel moisture content: a pyro-ecophysiological approach. *Fire Ecol.* 21 (1).
- Jolly, W.M., Hadlow, A.M., Hugué, K., 2014. De-coupling seasonal changes in water content and dry matter to predict live conifer foliar moisture content. *Int. J. Wildland Fire* 23 (4).
- Jolly, W.M., Johnson, D.M., 2018. Pyro-ecophysiology: shifting the paradigm of live wildland fuel research. *Fire* 1 (1).
- Khabbazan, S., et al., 2022. The influence of surface canopy water on the relationship between L-band backscatter and biophysical variables in agricultural monitoring. *Remote Sens. Environ.* 268.
- Konings, A.G., et al., 2024. Tree species explain only half of explained spatial variability in plant water sensitivity. *Glob. Change Biol.* 30 (7).
- Kranz, J., et al., 2025. Assessing predictors for fuel moisture content in Central European forests. *Agric. For. Meteorol.* 371.
- Lai, G., Quan, X., Yebra, M., He, B., 2022. Model-driven estimation of closed and open shrublands live fuel moisture content. *GIScience Remote Sens.* 59 (1), 1837–1856.
- Lawson, B.D. and Armitage, O., 2008. Weather guide for the Canadian forest fire danger rating system.
- Li, Y., Li, C., Li, M., Liu, Z., 2019. Influence of variable selection and forest type on forest aboveground biomass estimation using machine learning algorithms. *Forests* 10 (12).
- Lundberg, S.M., et al., 2020. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* 2 (1), 56–67.
- Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. In: *Advances in neural information processing systems*, p. 30.
- Luo, K., Quan, X., He, B., Yebra, M., 2019. Effects of live fuel moisture content on wildfire occurrence in fire-prone regions over Southwest China. *Forests* 10 (10).
- Marino, E., et al., 2020. Investigating live fuel moisture content estimation in fire-prone shrubland from remote sensing using empirical modelling and RTM simulations. *Remote Sens.* 12 (14).
- Martin-StPaul, N., et al., 2018. Live fuel moisture content (LFMC) time series for multiple sites and species in the French Mediterranean area since 1996. *Ann. For. Sci.* 75 (2).
- Martin-StPaul, N., et al., 2023. How much does VPD drive tree water stress and forest disturbances? *Authorea Preprints*.
- Meyer, H., Reudenbach, C., Wöllauer, S., Nauss, T., 2019. Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction. *Ecol. Model.* 411.
- Miller, L., Zhu, L., Yebra, M., Rüdiger, C., Webb, G.I., 2022. Multi-modal temporal CNNs for live fuel moisture content estimation. *Environ. Model. Softw.* 156.
- Miller, L., Zhu, L., Yebra, M., Rüdiger, C., Webb, G.I., 2023. Projecting live fuel moisture content via deep learning. *Int. J. Wildland Fire*.
- Nasirzadehdizaji, R., et al., 2019. Sensitivity analysis of multi-temporal sentinel-1 SAR parameters to crop height and canopy coverage. *Appl. Sci.* 9 (4).
- Nolan, R.H., Boer, M.M., Resco de Dios, V., Caccamo, G., Bradstock, R.A., 2016. Large-scale, dynamic transformations in fuel moisture drive wildfire activity across southeastern Australia. *Geophys. Res. Lett.* 43 (9), 4229–4238.
- Nolan, R.H., et al., 2022. Drought-related leaf functional traits control spatial and temporal dynamics of live fuel moisture content. *Agric. For. Meteorol.*, p. 319.
- Office National des Forêts, 2022. Les rendez-vous techniques de l'ONF. 1763-6442. Office national des forêts.
- Pedregosa, F., et al., 2011. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Pellizzaro, G., Cesaraccio, C., Duce, P., Ventura, A., Zara, P., 2007. Relationships between seasonal patterns of live fuel moisture and meteorological drought indices for Mediterranean shrubland species. *Int. J. Wildland Fire* 16 (2).
- Penot, V., Merlin, O., 2023. Estimating the water deficit index of a Mediterranean holm Oak forest from landsat optical/thermal data: A phenomenological correction for trees casting shadow effects. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 16, 5867–5880.
- Peterson, S.H., Roberts, D.A., Dennison, P.E., 2008. Mapping live fuel moisture with MODIS data: A multiple regression approach. *Remote Sens. Environ.* 112 (12), 4272–4284.
- Pimont, F., et al., 2023. Projections des effets du changement climatique sur l'activité des feux de forêt au 21ème siècle: Rapport final Contribution technique à la mission d'expertise conjointe sur l'extension des zones à risque d'incendie de forêt et de végétation à échéance du milieu et fin du siècle dans le contexte du changement climatique. INRAE.
- Pimont, F., Ruffault, J., Martin-StPaul, N.K., Dupuy, J.L., 2019. Why is the effect of live fuel moisture content on fire rate of spread underestimated in field experiments in shrublands? *Int. J. Wildland Fire* 28 (2).
- Pimont, F., et al., 2022. Future expansion, seasonal lengthening and intensification of fire activity under climate change in southeastern France. *Int. J. Wildland Fire*.

- Poggio, L., et al., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *Soil* 7 (1), 217–240.
- Quan, X., Xie, Q., He, B., Luo, K., Liu, X., 2021a. Integrating remotely sensed fuel variables into wildfire danger assessment for China. *Int. J. Wildland Fire*.
- Quan, X., et al., 2021b. Global fuel moisture content mapping from MODIS. *Int. J. Appl. Earth Obs. Geoinf.* 101.
- Rao, K., Williams, A.P., Flefil, J.F., Konings, A.G., 2020. SAR-enhanced mapping of live fuel moisture content. *Remote Sens. Environ.* 245.
- Reichle, R.H. et al., 2022. Soil Moisture Active Passive (SMAP) mission level 4 surface and root zone soil moisture (L4_SM) product specification document.
- Reichle, R.H., et al., 2017. Assessment of the SMAP level-4 surface and root-zone soil moisture product using In situ measurements. *J. Hydrometeorol.* 18 (10), 2621–2645.
- Reichle, R.H., Koster, R.D., 2005. Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. *Geophys. Res. Lett.* 32 (2).
- Rogelj, J., Meinshausen, M., Knutti, R., 2012. Global warming under old and new scenarios using IPCC climate sensitivity range estimates. *Nat. Clim. Change* 2 (4), 248–253.
- Rouse Jr, J.W., Haas, R.H., Schell, J. and Deering, D., 1973. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation.
- Ruffault, J., et al., 2020. Increased likelihood of heat-induced large wildfires in the. *Mediterr. Basin. Sci. Rep.* 10 (1), 13790.
- Ruffault, J., et al., 2023. Plant hydraulic modelling of leaf and canopy fuel moisture content reveals increasing vulnerability of a Mediterranean forest to wildfires under extreme drought. *New. Phytol.* 237 (4), 1256–1269.
- Ruffault, J., Martin-StPaul, N., Pimont, F., Dupuy, J.-L., 2018. How well do meteorological drought indices predict live fuel moisture content (LFMC)? An assessment for wildfire research and operations in Mediterranean ecosystems. *Agric. For. Meteorol.* 262, 391–401.
- Scarff, F.R., et al., 2021. Effects of plant hydraulic traits on the flammability of live fine canopy fuels. *Funct. Ecol.* 35 (4), 835–846.
- Schwartz, M., et al., 2023. FORMS: forest multiple source height, wood volume, and biomass maps in France at 10 to 30 m resolution based on Sentinel-1, Sentinel-2, and GEDI data with a deep learning approach. *Earth Syst. Sci. Data Discuss* 2023, 1–28.
- Scikit-learn developers, 2023. GridSearchCV: Exhaustive Search Over Specified Parameter Values for an Estimator.
- Sow, M., Mbow, C., Hély, C., Fensholt, R., Sambou, B., 2013. Estimation of herbaceous fuel moisture content using vegetation indices and land surface temperature from MODIS data. *Remote Sens.* 5 (6), 2617–2638.
- Stocker, B.D., et al., 2023. Global patterns of water storage in the rooting zones of vegetation. *Nat. Geosci.* 16 (3), 250–256.
- Štrumbelj, E., Kononenko, I., 2013. Explaining prediction models and individual predictions with feature contributions. *Knowl. Inf. Syst.* 41 (3), 647–665.
- Tanase, M.A., et al., 2022. Characterizing live fuel moisture content from active and passive sensors in a Mediterranean environment. *Forests* 13 (11).
- Tramblay, Y., et al., 2020. Challenges for drought assessment in the Mediterranean region under future climate scenarios. *Earth-Sci. Rev.* 210.
- Tumber-Davila, S.J., Schenk, H.J., Du, E., Jackson, R.B., 2022. Plant sizes and shapes above and belowground and their interactions with climate. *New. Phytol.* 235 (3), 1032–1056.
- Ulaby, F.T., 1986. Microwave remote sensing active and passive volume 3.
- van Vuuren, D.P., et al., 2011. The representative concentration pathways: an overview. *Clim. Change* 109 (1–2), 5–31.
- Vidal, J.P., Martin, E., Franchistéguy, L., Baillon, M., Soubeyroux, J.M., 2009. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *Int. J. Climatol.* 30 (11), 1627–1644.
- Vinodkumar, V., Dharssi, I., Yebra, M., Fox-Hughes, P., 2021. Continental-scale prediction of live fuel moisture content using soil moisture information. *Agric. For. Meteorol.*, p. 307.
- Vinuales, A., et al., 2024. Real-time assessment of live forest fuel moisture content and flammability by using space-time universal kriging. *Ecol. Model.* 498.
- Vreugdenhil, M., et al., 2018. Sensitivity of sentinel-1 backscatter to vegetation dynamics: an Austrian case study. *Remote Sens.* 10 (9).
- Wagner, C.V., 1987. Development and structure of the Canadian forest fire weather index system.
- Wan, Z., Hook, S., Hulley, G., 2021a. MODIS/Terra land surface temperature/emissivity daily L3 Global 1km SIN Grid V061 [Data set], Editor. N.E.L.P.D.A.A. Center.
- Wan, Z., Hook, S., Hulley, G., 2021b. MODIS/Terra land surface temperature/emissivity daily L3 Global 1km SIN Grid V061 [Data set]. NASA EOSDIS Land Processes Distributed Active Archive Center.
- Wang, X., et al., 2017. cffdrs: an R package for the Canadian forest fire danger rating system. *Ecol. Process.* 6 (1).
- Yebra, M., Chuvieco, E., 2009. Linking ecological information and radiative transfer models to estimate fuel moisture content in the Mediterranean region of Spain: solving the ill-posed inverse problem. *Remote Sens. Environ.* 113 (11), 2403–2411.
- Yebra, M., et al., 2013. A global review of remote sensing of live fuel moisture content for fire danger assessment: moving towards operational products. *Remote Sens. Environ.* 136, 455–468.
- Yebra, M., et al., 2018. A fuel moisture content and flammability monitoring methodology for continental Australia based on optical remote sensing. *Remote Sens. Environ.* 212, 260–272.
- Yebra, M., et al., 2024. Globe-LFMC 2.0, an enhanced and updated dataset for live fuel moisture content research. *Sci. Data* 11 (1).
- Yebra, M., et al., 2019. Globe-LFMC, a global plant water status database for vegetation ecophysiology and wildfire applications. *Sci. Data* 6 (1).
- Yuan, W., et al., 2019. Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Sci. Adv.* 5 (8), eaax1396.
- Zanaga, D. et al., 2021. ESA WorldCover 10m 2020 v100.
- Zhu, L., et al., 2021. Live fuel moisture content estimation from MODIS: A deep learning approach. *ISPRS J. Photogramm. Remote Sens.* 179, 81–91.