Máster en Geomática, Teledetección y Modelos Espaciales Aplicados a la Gestión Forestal

TRABAJO FIN DE MÁSTER

Live Fuel Moisture Content Mapping in the Mediterranean Basin Using Random Forests and Combining MODIS Spectral and Thermal Data

Autor:

Àngel Ramon Cunill Camprubí

Director:

Pablo González Moreno

Córdoba, septiembre de 2022

Autorización para la Presentación a Evaluación del Trabajo Final de Máster

El Dr. D. Pablo González Moreno, profesor del Departamento de Ingeniería Forestal de la Escuela Técnica Superior de Ingeniería Agronómica y de Montes de la Universidad de Córdoba, director del Trabajo Fin de Master, "*Live Fuel Moisture Content Mapping in the Mediterranean Basin Using Random Forests and Combining MODIS Spectral and Thermal Data*" del que es autor el alumno D. Àngel Ramon Cunill Camprubí, con DNI ------------, y que presenta para optar al título del Máster en Geomática, Teledetección y Modelos Espaciales aplicados a la gestión forestal.

INFORMA:
Que el referido trabajo tiene carácter investigador y acredita su idoneidad y otorga el V.º B.º a su contenido para ir a Tribunal de Trabajo Fin de Máster.

Y para que conste, firma el presente escrito en Córdoba, a 1 de septiembre de 2022.

Director: Autor:

Pablo González Moreno Àngel R. Cunill Camprubí

Declaración de Autoría y Originalidad

Yo, Àngel Ramon Cunill Camprubí, con DNI --------------, y estudiante del Máster en Geomática, Teledetección y Modelos Espaciales aplicados a la gestión forestal de la Universidad de Córdoba, declaro ser el autor del presente Trabajo Final de Máster, con título "*Live Fuel Moisture Content Mapping in the Mediterranean Basin Using Random Forests and Combining MODIS Spectral and Thermal Data*", y que este es original en el sentido de que no se han utilizado fuentes sin citarlas debidamente.

También informo de que dicho trabajo ha generado una publicación en la revista internacional *Remote Sensing* con el mismo título, dando valor añadido a su contenido y validando su originalidad. La cita al artículo publicado es la siguiente:

Cunill Camprubí À, 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):3162.
https://doi.org/10.3390/rs14133162

Y para que conste, firmo el presente escrito en Córdoba, a 1 de septiembre de 2022.

Àngel R. Cunill Camprubí

Agradecimientos

Me gustaría mostrar mi más sincero agradecimiento a las personas e instituciones que han hecho posible la realización de este trabajo.

En primer lugar, a mi tutor el Dr. D. Pablo González Moreno, por su labor de guía y por transmitirme su experiencia y conocimientos a través de comentarios precisos y correcciones certeras que han ayudado a la mejora sustancial de esta memoria y a mi aprendizaje en general.

Quedo especialmente agradecido con el Dr. D. Víctor Resco de Dios, de la Universitat de Lleida, mi actual 'jefe', por ofrecerme la propuesta y permitirme desarrollar el trabajo dentro del entorno laboral al que estoy vinculado. Destaco su apoyo, sus aportaciones científicas, sus comentarios y sus correcciones durante todo el proceso de elaboración del trabajo, que han contribuido indudablemente a aumentar la calidad del mismo.

Agradezco también el apoyo de la Fundació AGROTECNIO, que es donde desarrollo mi actividad profesional y, por lo tanto, donde he podido desarrollar mi TFM.

Debo dar las gracias a la labor formativa que han realizado todos los docentes del Máster GEOFOREST y, en especial, a su director el Dr. D. Rafael María Navarro Cerrillo y a los miembros del IDEP por comprender mi situación tras tener un hijo mientras realizaba el trabajo y facilitarme los tramites para poder defender mi TFM.

Por último, a mi pareja Marta, por su paciencia y por apoyarme en el día a día de manera incondicional, y a todos aquellos que de alguna forma han ayudado a que este trabajo sea hoy una realidad.

Table of Contents

List of Figures

List of Tables

List of Abbreviatures

B*	Band (* = number of band)
CAL	Calibration, model validation inside the training period
Cat-LFMC	Catalan LFMC database
Cb	Penalty coefficient of the Lin's CCC
CCC	Lin's concordance correlation coefficient
DOY	Day of year
EXT	Extrapolation, model validation outside the training period
FFS	Forward Feature Selection
GEE	Google Earth Engine
Globe-LFMC	Global LFMC database
IGBP	International Geosphere-Biosphere Programme
LFMC	Live fuel moisture content
LFMC_{RF}	LFMC predictions with the RF model
LLOCV	Leave-location-out cross-validation
LST	Land surface temperature
MAE	Mean absolute error
MBE	Mean bias error (difference between predicted and observed)
ML	Machine learning
MODIS	Moderate Resolution Imaging Spectroradiometer
MP	General model performance assessment
NDVI_{CV}	Coefficient of variation of the Normalized Difference Vegetation Index
NIR	Near infrared
NR*	Nadir reflectance (* = number of band)
RF	Random forests
RMSE	Root mean square error
RTM	Radiative transfer model
SI	Spectral indices
SWIR	Short-wave infrared
ubRMSE	Unbiased root mean square error
VEcv	Variance explained by cross-validation
VIIRS	Visible Infrared Imaging Radiometer Suite
VPD	Vapor pressure deficit

Resumen

Los índices de vegetación derivados de la teledetección se han utilizado ampliamente para estimar el contenido de humedad del combustible vivo (LFMC por sus siglas en inglés), un factor importante del riesgo de incendios forestales, debido a la amplia disponibilidad de datos. Sin embargo, marcadas diferencias en la estructura de la vegetación afectan la relación entre LFMC medido en campo y la reflectancia captada por los sensores de los satélites, lo que limita la extrapolación espacial de estos índices. Para superar esta limitación, exploré el potencial de Random Forests (RF), una técnica de aprendizaje automático basada en la agregación de múltiples árboles de decisión, para estimar LFMC a escala subcontinental en la cuenca Mediterránea. Probé distintos modelos de RF usando una combinación de bandas espectrales de MODIS, índices de vegetación, la temperatura superficial terrestre y el día del año como predictores. Utilicé las bases de datos del Globe-LFMC y del programa catalán de seguimiento de LFMC como muestras de verdad-terreno (10.374 muestras). El proceso de modelado consistió en una selección de predictores y una validación cruzada espacial para evitar el sobreajuste espacial. El modelo final de $LFMC_{RF}$ se calibró y evaluó con muestras recolectadas entre 2000 y 2014, y se probó de forma independiente con muestras de 2015 a 2019, reportando valores generales de raíz del error cuadrático medio (RMSE por sus siglas en inglés) de 19,9% y 16,4%, respectivamente. Los resultados de $LFMC_{RF}$ fueron comparables a los enfoques actuales basados en modelos de transferencia radiativa (RMSE ~74–78%), introduciendo una alternativa confiable para aplicaciones a gran escala. Este estudio llena un importante vacío de investigación al crear un enfoque homogéneo para estimar LFMC en toda el área occidental de la cuenca Mediterránea. El modelo final fue usado para generar una base de datos pública con mapas de LFMC semanales extendidos a toda la cuenca Mediterránea propensa a incendios forestales.

Palabras clave: contenido de humedad del combustible vivo, incendio forestal, MODIS, índice espectral, temperatura superficial terrestre, Random Forest

Abstract

Remotely sensed vegetation indices have been widely used to estimate live fuel moisture content (LFMC), an important driver of wildfire risk, due to broad data availability. However, marked differences in vegetation structure affect the relationship between field-measured LFMC and reflectance, which limits spatial extrapolation of these indices. To overcome this limitation, I explored the potential of Random Forests (RF), a machine learning technique based on the ensemble of multiple decision trees, to estimate LFMC at the subcontinental scale in the Mediterranean basin wildland. I built RF models using a combination of MODIS spectral bands, vegetation indices, surface temperature, and the day of year as predictors. I used the Globe-LFMC and the Catalan LFMC monitoring program databases as ground-truth samples (10,374 samples). The modelling process consisted in a feature selection and two step spatial cross-validation in order to avoid spatial overfitting. The final $\mathrm{LFMC_{RF}}$ model was calibrated and evaluated with samples collected between 2000 and 2014, and independently tested with samples from 2015 to 2019 reporting an overall root mean square errors (RMSE) of 19.9% and 16.4%, respectively. The results from $\mathrm{LFMC_{RF}}$ were comparable to current approaches based on radiative transfer models (RMSE ~74–78%), introducing a reliable alternative for large-scale applications. This study fills an important research gap by creating a homogeneous approach to estimate LFMC across the Western Mediterranean basin. I used the final model to generate a public database with weekly LFMC maps extended to the fire-prone Mediterranean basin.

Keywords: live fuel moisture content, wildfire, MODIS, spectral indices, land surface temperature, Random Forests.

1. Introduction

Wildfires are a recurring phenomenon in many terrestrial biomes, where they play a fundamental role as drivers of ecosystem dynamics (Bowman et al., 2020). Interactions between climate, vegetation and fire occurrence have led to distinct fire regimes (Bowman et al., 2020). Human actions have modified intensively these fire regimes in many areas worldwide, including the Mediterranean basin, leading to undesirable impacts in ecosystems with carry-over effects for human safety, health and the economy (Duane et al., 2021; Karavani et al., 2018).

Live fuel moisture content (LFMC), the mass of water in the foliage and small twigs relative to its total dry mass, is a key factor affecting fire potential and determining wildfire danger and activity (Bradstock, 2010; Resco de Dios, 2020). Fuel moisture is directly related to the amount of energy needed to evaporate water before ignition (Jolly & Johnson, 2018; Resco de Dios, 2020). Consequently, high moisture values reduce, or even inhibit, ignitability and subsequent fire spread (Nelson, 2001).

Different studies conducted in a wide range of ecosystems have observed a significant correlation between burned area and LFMC (Dennison & Moritz, 2009; Luo et al., 2019; Nolan et al., 2016). More specifically, these studies report that large fires only occur once fuel moisture crosses critical dryness levels. In Mediterranean regions, longer summer drought periods along with increases in temperature have been projected under climate change (IPCC, 2022). Such climatic changes could significantly decline LFMC and consequently enhance the length of the fire season and the rate of high intensity fires (Dupuy et al., 2020). This situation could be exacerbated with intensifying fuel load accumulation and fuel connectivity as a result of rural exodus and widespread lack of land management. As a consequence, the probability and the frequency of extreme fire events is expected to increase (Dupuy et al., 2020). Accurate and comprehensive spatial and temporal estimations of LFMC are thus needed to assess wildfire danger (Chuvieco et al., 2020) and to develop early warning systems for the evolution of critical conditions (Boer et al., 2017).

Regional-scale assessment of LFMC is commonly obtained through expensive and time-consuming field inventories (Gabriel et al., 2021; Martin-StPaul et al., 2018) or through meteorological drought indices (e.g., Van Wagner, 1987). The latter allow spatially continuous measurements, but their validity for the Mediterranean area has been questioned in various studies (Caccamo et al., 2012; Ruffault et al., 2018; Soler Martin et al., 2017), as they do not take into account plant-specific differences and the influence of site conditions (e.g., soil water dynamics), often leading to poor predictions (Jolly & Johnson, 2018).

Remote sensing of LFMC using satellite information provides a valuable alternative to overcome the limitations of drought indices. Current approaches are mainly grouped into either physically-based simulation (Jurdao et al., 2013; Yebra et al., 2008, 2018) or empirical methods (Argañaraz et al., 2016; Caccamo et al., 2012; Chuvieco et al., 2004; Peterson et al., 2008). Generally, these methods measure how water absorption and leaf properties affect reflectance in the optical spectrum (Yebra et al., 2013). Physical approaches, such as radiative transfer models (RTM), are expected to be more robust than empirical methods (Chuvieco et al., 2020). This is because they are based on the physical associations between leaf-canopy properties and spectral reflectance, which are independent of sensor and site conditions (Yebra et al., 2008; Yebra & Chuvieco, 2009). However, they are also more complex to parameterize and require additional ecological information and prior knowledge over large geographical gradients to prevent unrealistic spectra simulations (Yebra & Chuvieco, 2009). In contrast, empirical approaches, which are commonly based on spectral indices (SI), are simpler and have shown similar or even better accuracies than physical models when applied locally (Marino et al., 2020; Yebra et al., 2008) or across specific vegetation types (Yebra et al., 2018).

Combinations of SI have been successfully employed to estimate LFMC (Argañaraz et al., 2016; Caccamo et al., 2012; Marino et al., 2020; Nolan et al., 2016). In addition, some authors found stronger predictive power by including land surface temperature (LST) along with optical data to the empirical relationships (Chuvieco et al., 2004; García et al., 2008; McCandless et al., 2020; Sow et al., 2013). The connection between LFMC and LST lies on the interaction between the plant energy balance mechanisms and its response to water stress (Yebra et al., 2013). Other recent studies implement microwave remote sensing to retrieve LFMC (Fan et al., 2018; Rao et al., 2020; Wang et al., 2019), but their use still has some limitations, such as data availability.

The application of empirical approaches at continental or global scales is precisely constrained by the availability of data for calibration during model development

(Argañaraz et al., 2016; Jurdao et al., 2013). The biophysical and structural differences among species impact the functional relationships between LFMC and remotely sensed reflectance (Ceccato et al., 2001; Zhu et al., 2021). Consequently, a large number of diverse sampling observations is required to reduce the effect of site dependence. Furthermore, the use of many predictive variables potentially related to LFMC may significantly improve the empirical estimations of the model (Yebra et al., 2018), but also increases its complexity.

Machine learning (ML) algorithms, such as Random Forests (RF), are a solid alternative to physically based RTM methods or the classical regression models on which the empirical approaches are commonly based. ML algorithms are highly efficient with high dimensional data and solve the problem of model complexity by applying different functional forms in the relation between predictors and LFMC, without make explicit a priori assumptions (Kuhn & Johnson, 2013). However, using ML to estimate LFMC from remote sensing is still very recent (Adab et al., 2016; McCandless et al., 2020; Rao et al., 2020; Zhu et al., 2021) and has not been used in the Mediterranean basin.

Despite the importance of wildfires in the Mediterranean basin, we are currently lacking a specific method to reliably estimate LFMC at the subcontinental scale. For example, the European Forest Fires Information System (EFFIS) is using the Australian operational system (Yebra et al., 2018) to estimate LFMC in the European extent, but this method has not been broadly assessed yet. Other studies have addressed LFMC modelling at local (Costa-Saura et al., 2021; Marino et al., 2020) or regional (Jurdao et al., 2013; Yebra & Chuvieco, 2009) scales and they are usually focused on specific vegetation types (e.g., grasslands or shrublands). Thus, we are still lacking a product that provides complete LFMC estimates for the Mediterranean basin. The only exception is the global LFMC product developed by Quan et al. (2021), which is based on an RTM, and it is not yet known whether LFMC estimates could be improved through ML approaches.

The present study aims to fill this knowledge gap by developing an RF algorithm to predict LFMC within the Western Mediterranean basin using the information of the widely used Moderate Resolution Imaging Spectroradiometer (MODIS), and comparing the results with the only other method available for this area, the physically-based estimations of Quan et al. (2021). I also aim to generalize the model over a wide range of fuel types with a unique formulation by combining a forward feature selection with a spatial cross-validation and ML techniques. Finally, the ultimate goal is to develop a database of LFMC for the Mediterranean basin using available data that improves beyond currently existing products.

2. Materials and Methods

2.1. Data

2.1.1. LFMC Field Measurements

I used all the LFMC data publicly accessible to date in the Mediterranean basin. Most of these data have been compiled in the Globe-LFMC database (Yebra et al., 2019, last accessed June 2021). The Globe-LFMC is a global compilation of 161,717 LFMC destructive field measurements of leaves and small twigs (<6 mm) from 1977 to 2018 at 1,383 sampling sites with different species and characteristics in 11 fire-prone countries (Yebra et al., 2019). I also found a more recent LFMC time series from Catalonia (Cat-LFMC) (Gabriel et al., 2021). This is a collection of 21 years (1998-2019) of biweekly field-sampled data compiled by the Catalan Forest Fire Prevention Service across nine sampling areas within this Spanish region, and focused on five species representatives of Mediterranean shrublands (Gabriel et al., 2021). Cat-LFMC was added to the Globe-LFMC to extend the total number of sites and the time interval within the Mediterranean area. Both datasets have already been technically validated by correcting inconsistencies and anomalies in LFMC, as described in the relevant publications (Gabriel et al., 2021; Yebra et al., 2019). All records are properly georeferenced and inform about the species collected, the sampling protocol, land cover type, and further eco-physiological and environmental properties not used in this study.

2.1.2. MODIS Data

The MODIS MCD43A4 Collection 6 product (Schaaf & Wang, 2015) was selected as a source of aboveground spectral information, as it has shown good performance in previous studies (Argañaraz et al., 2016; Marino et al., 2020; Zhu et al., 2021). MCD43A4 provides daily maps at 500 m spatial resolution from a 16-day composite of Nadir Bidirectional Distribution Function (NBDF)-Adjusted Reflectance for each of the 7 MODIS bands (channels 1-7, Table 1). Using a composite product may reduce the

Table 1. List of potential predictors of LFMC.

Variable	Description	Wavelength (nm)	Source
NR1	Nadir Reflectance Band 1 Red	620 - 670	MCD43A4
NR2	Nadir Reflectance Band 2 Near infrared (NIR1)	841 - 876	MCD43A4
NR3	Nadir Reflectance Band 3 Blue	459 - 479	MCD43A4
NR4	Nadir Reflectance Band 4 Green	545 - 564	MCD43A4
NR5	Nadir Reflectance Band 5 Near infrared (NIR2)	1230 - 1250	MCD43A4
NR6	Nadir Reflectance Band 6 Shortwave infrared (SWIR1)	1628 - 1652	MCD43A4
NR7	Nadir Reflectance Band 7 Shortwave infrared (SWIR2)	2105 - 2155	MCD43A4
SI	Vegetation spectral indices: NDVI, EVI, SAVI, VARI, VIgreen, Gratio, NDII6, NDII7, NDWI, GVMI, MSI, NDTI, STI		see Table A2
LST	Land surface temperature		MOD11A2
DOY_COS, DOY_SIN	Cosine and Sine of the Day of Year		

probability of cloud cover and shadows. The ‘Good quality’ flag from the simplified band specific quality layers (BRDF_Albedo_Band_Quality) associated with MCD43A4 was used to keep the full quality pixels of the composite.

The Terra MODIS Land Surface Temperature (LST) MOD11A2 Collection 6 product was included as a predictor of LFMC due to the impact of water availability in plant evapotranspiration and, consequently, on canopy temperature (Yebra et al., 2013). MOD11A2 is an 8-day pixel average from the MOD11A1, a daily product of LST measurements from the Terra satellite (Wan, 2014). I used the daytime composite values, instead of single day measurements, because MOD11A2 had fewer data gaps (8% vs 35%), and their effect on LFMC predictions, in terms of model RMSE, was the same (~20%, see Appendix A1). Daytime images cover the same period as MCD43A4, and they coincide better with the typical sample collection time, but at a 1000 m spatial resolution. They were resampled to the 500 m spatial resolution of MCD43A4 using a bilinear interpolation.

Additionally, the annual MODIS Land Cover Type (MCD12Q1) Collection 6 product (Sulla-Menashe et al., 2019) with the International Geosphere-Biosphere Programme (IGBP) classification scheme was used to distinguish between vegetation types in the analyses. This product replaced the land cover field included in the Globe-LFMC database, which is based on the ESA Climate Change Initiative Land Cover for the year 2015. This is because MCD12Q1 accommodates to the spatial resolution of the reflectance data and the temporal resolution of the field samples. It also was used for map production (e.g., masking water bodies and non-vegetation covers).

All MODIS images were downloaded from the NASA Land Processes Distributed Active Archive Center (LP DAAC) in the U.S. Geological Survey (USGS) Earth Resources Observation and Science Center (EROS) (https://lpdaac.usgs.gov/; accessed on June 2021).

2.1.3. Landsat Data

The Landsat Collection 1 surface reflectance data included in Google Earth Engine (GEE) (Gorelick et al., 2017) was used to assess the MODIS subpixel spatial heterogeneity corresponding to each sampling site in the LFMC dataset. The revisit time of these satellites is 16 days, and the resolution is 30 m for the reflective bands. Similarly to Quan et al. (2021), I employed Landsat 5 TM from Feb 2000 to Oct 2011 for high quality pixels, Landsat 7 ETM+ from Feb 2000 to Oct 2011 when Landsat 5 TM had poor quality pixels and also from Nov 2011 to Apr 2013, and Landsat 8 OLI from May 2013 until 2019. The use of Landsat 5 TM instead of Landsat 7 ETM+ was due to data gaps produced in the latter by failure in a sensor component (Quan et al., 2021). Snow, cloud, and shadow pixels were removed using the Landsat internal quality band.

2.1.4. Radiative Transfer Model (RTM) Database

The global RTM-based product developed by Quan et al. (2021) was used to compare the results of the ML-based approach proposed here. I chose this product because it is the only currently available database that has produced LFMC maps over the whole Mediterranean basin. It consists of a weekly collection of maps (2001-2019) generated by a physically-based remote sensing model.

2.2. Methods

The following sections describe all the steps I used to estimate LFMC (Figure 1). The first section explains how I prepared the data for analyses. The second section briefly introduces the modelling approach. The last sections describe the variable selection process, the calibration and validation methods, and the software used in all steps.

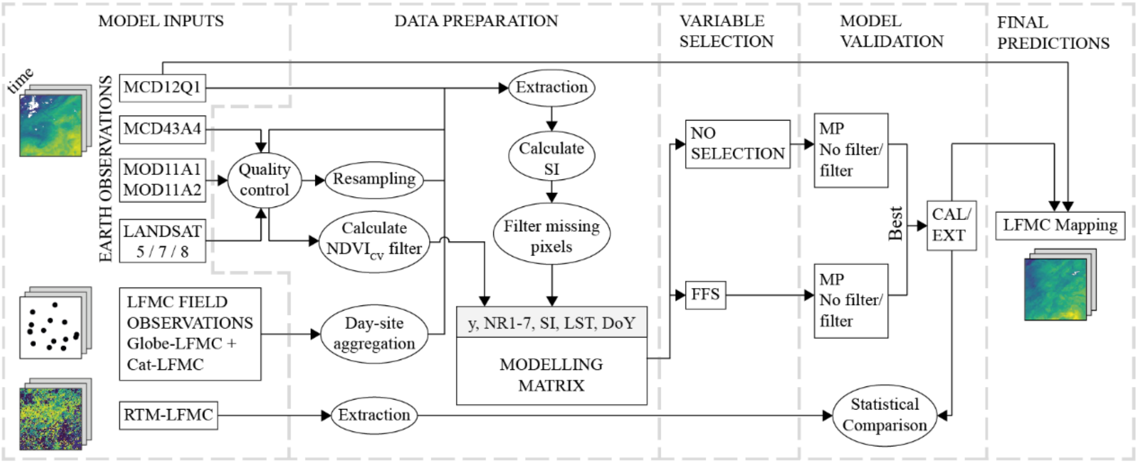

Figure 1. Overview of the pre-processing, modelling and analysis steps.

2.2.1. Data Preparation

First, I cropped the Globe-LFMC dataset to the Mediterranean region (Figure 2) and used it, along with the Cat-LFMC, only for the dates with available MODIS data. Then, LFMC samples collected within the same day and site, but corresponding to different species or vegetation layers (e.g., understory and canopy), were aggregated by arithmetic means to obtain a single value per site. Nolan et al. (2016) observed that average LFMC per site has a stronger correlation with spectral data than any individual vegetation layer alone. However, some studies have observed that spectral information may more closely reflect signals from the upper part of the canopy, particularly for closed forests (Yebra et al., 2013). I am interested in developing an indicator of LFMC representative of the entire canopy (upper canopy but also of the understory) because the understory often burns during a fire, which explains why I used the average LFMC value.

For each resultant LFMC sample, pixel values from remote sensing data were obtained by a simple pixel extraction (that is, the nearest grid cell centroid) matching their sampling date. I performed some preliminary tests observing that the simple pixel

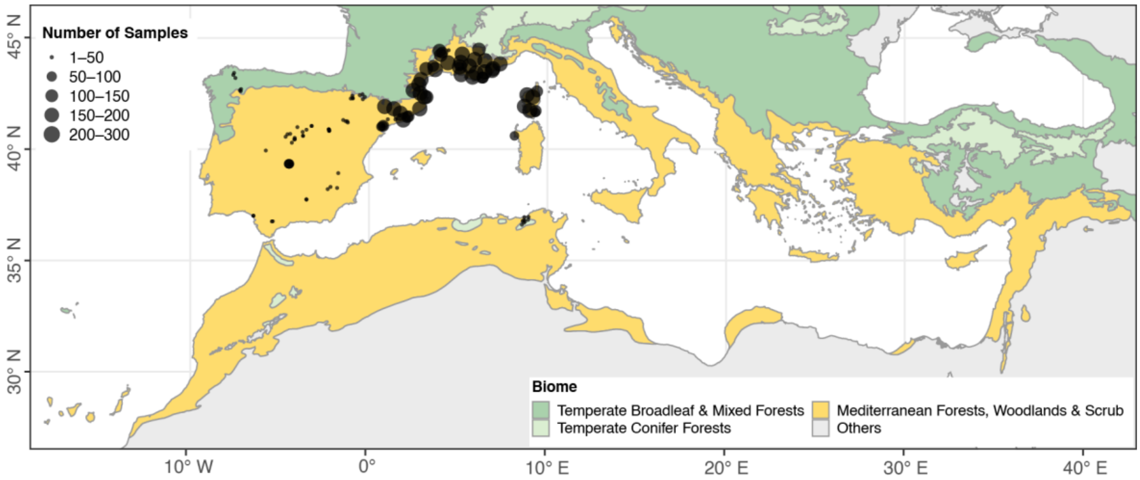

Figure 2. Distribution of sampling sites and extension of the mapping area for the database and future map productions. The background layer represents terrestrial biomes (Dinerstein et al., 2017). Gray areas were discarded from the predictions.

extraction method showed no significant differences (p-value = 0.9; see Appendix A2, Table A1) relative to conducting a focal mean (e.g., from a 3×3 window). Afterwards, various vegetation SI (Table A2) potentially related to LFMC were calculated by combining information from different MODIS spectral bands and used as predictors of LFMC in addition to atmospherically corrected reflectance. SI tend to reduce directional anisotropic and soil background effects and highlight specific properties of the vegetation canopy (Yebra et al., 2013). I also used LST from the MODIS LST 8-day average composite, as previously discussed (see Appendix A1). Finally, I added the day of year (DOY) of the ground LFMC samples as auxiliary variables to take into account the seasonal trends in LFMC (Chuvieco et al., 2004; García et al., 2008). To do so, DOY was normalized to [0, 1] and reconverted to $[-\pi, \pi]$, such that DOY 1 and DOY 366 corresponded to $-\pi$ and π, respectively. With the resulting values, I calculated the sine (DOY_SIN) and cosine (DOY_COS) to maintain the information on the periodicity as performed in Zhu et al. (2021). Consequently, DOY_SIN varied from -1 to 1 between the wettest and driest season, while DOY_COS varied from winter (coldest; -1) and summer (hottest; 1).

After defining the potential predictors described above (Table 1), I removed LFMC samples with missing data from any variable, and I discarded values outside the threshold 20-250%, which is considered the biological range of LFMC (Martin-StPaul et al., 2018). I then averaged multiple observations in the same day and MODIS-grid cell and randomly assigned the values to one of their locations to have a single daily LFMC value for a given pixel value. The resulting dataset contained a total of 10,374 LFMC field measurements between 2000 and 2019 from 118 sites located in Spain, France, Italy, and Tunisia (Figure 2; Figures B1, and B2). These sites are mostly concentrated in the ecoregions 'Northeast Spain and Southern France Mediterranean forests' and 'Italian sclerophyllous and semi-deciduous forests' (~80%). Ecoregions with Mediterranean woodlands and coniferous, broadleaf, and mixed forest formations are also represented to a minor degree. In conjunction, mean annual temperature ranges from 6 to 20 °C and mean annual rainfall ranges from 250 to 1100 mm (Dinerstein et al., 2017). Site altitudes ranges from 11 to 1660 m.

For model validation, Quan et al. (2021) RTM data were only acquired for the sample records that coincided with the available dates of such products. I also assigned land cover information from the MCD12Q1 layers to each ground sample by matching the year of sampling with the year of the layer. Misclassified sites (e.g., croplands, permanent wetlands, and urban covers) were discarded or manually corrected based on the species collected, location, and the land cover type field included in the Globe-LFMC database. To simplify the analyses, the IGBP land cover classes present in the study were re-classified into four vegetation (or fuel) types accounting for different structural characteristics (Table A3). These new land cover classes were defined as grasslands, shrublands, savannas (tree cover 10-60%), and forests (tree cover >60%).

Additionally, the NDVI coefficient of variation ($NDVI_{CV}$) derived from Landsat data were used to assess the homogeneity of vegetation 'greenness' surrounding each site coordinates, as performed in Quan et al. (2021). The authors suggest using these metrics to filter highly heterogeneous areas within a specific satellite footprint since they may not be suitable for predictive attributions (Yebra et al., 2019). Lower values correspond to more homogeneous sites. $NDVI_{CV}$ was calculated with the Landsat surface reflectance values from a 500×500 m^2 buffer that matched the MODIS cell where site coordinates were located. To do so, I adapted the GEE script publicly shared by Yebra et al. (2019), such that the $NDVI_{CV}$ value was the monthly average that corresponded to the sampling date. Monthly average maximizes the quality (unmasked pixels) and the stability of the $NDVI_{CV}$ statistic. Only values with more than 80% good quality pixels (without snow, clouds, or shadows) were retained.

2.2.2. Machine Learning Approach

Random Forests (RF) was the ML algorithm chosen to empirically estimate LFMC at the Mediterranean basin because of its simplicity, its ability to deal with a large number of covariates, and because it is not necessary to have prior knowledge of the functional form of the relationships between these covariates and the response. Furthermore, the presence of outliers does not have a great influence on its performance (Kuhn & Johnson, 2013).

RF is a non-parametric data-driven statistical method proposed by Breiman (2001), which is based on Classification and Regression Trees (CART, also called decision trees) and bagging. Several decision trees are constructed in different bootstrap samples of the data, on which every data split (node) is forced to consider an arbitrary subset of available predictors. All individual-tree responses are then aggregated to obtain the final output predictions. The hyperparameters needed for model calibration and used in the subsequent analyses are explained in the supplementary methods (Appendix A5). Full details on CART, bagging, and RF can be found in Kuhn and Johnson (2013).

2.2.3. Variable Selection: Forward Feature Selection

Variable selection was needed because many of the variables (or features) used as candidates to estimate LFMC were highly correlated with each other, as expected (Figure B3). This is because the SI were formed by close combinations of different spectral bands. On the other hand, predictor variables that are highly autocorrelated in space can be misinterpreted by the Random Forests algorithm, leading to poor predictions outside the locations of the training data (Meyer et al., 2018).

Here, I used the Forward Feature Selection (FFS) method proposed by Meyer et al. (2018) to eliminate uninformative predictors and reduce the spatial over-fitting. First, the algorithm trains models using all possible combinations of two predictor variables and keeps those with the lowest prediction error based on a spatial cross-validation that discards entire sampling sites, as described later. Then, FFS iteratively increases the number of variables and evaluates the new model until none of the remaining variables improves the performance of the current best model. Additionally, I introduced a modification of the original method that consisted of calculating the average error over 25 different data splits. This avoided the dependence of cross-validation data splitting and aimed at stabilizing the error estimation (Krstajic et al., 2014).

FFS is complex and computationally intensive to execute parallel with RF parameter selection (Krstajic et al., 2014), and this step was performed before model calibration using a fixed set of hyperparameters (Table A4).

2.2.4. Model Selection and Performance Evaluation

In order to select the final model, I first assessed the general performance of different forms of the RF (depending on the selected predictors and whether or not the $NDVI_{CV}$ filter for heterogeneous pixels was applied) independently from a specific model calibration. I then adjusted the best performing model and evaluated its predictions.

Initial model performance assessment (MP) consisted of a bias-reduced predictive performance evaluation done using a nested 5-fold leave-location-out cross-validation (LLOCV) (Krstajic et al., 2014). Nested cross-validation divides the data two times, first to develop the model and then for independently testing its performance. LLOCV means that the cross-validation folds are made of the observations left out of complete locations, assuring spatial independence (Meyer et al., 2018). More specifically, the data were divided into 5 outer folds, where one was kept for testing and the remaining were split again into 5 nested folds to iteratively train and select the optimal tuning using a standard LLOCV. Five optimal models were obtained for each outer partition and the accuracy metrics (described in the section below) were then calculated based on the collection of predictions from all the outer folds. The same procedure was repeated 100 times with different data splits (that is, 500 independent validations), and the overall predictive power metrics were the mean of all repetitions.

Using this method, I assessed MP over 5 different model combinations with the entire set of variables, with the variables selected during the FFS, and with/without applying the $NDVI_{CV}$ filter. $NDVI_{CV}$ was treated as an additional hyperparameter and implemented in both the whole dataset (training and test) and only to the training partition. The five models consisted of: (1) all variables without filters; (2) all variables with $NDVI_{CV}$ filters on the whole dataset; (3) FFS-selected variables without filters; (4) FFS-selected variables with $NDVI_{CV}$ filters on the whole dataset; and (5) the best of all/selected variables with the $NDVI_{CV}$ filter only applied to the training partition. This method of evaluation provides an appropriate estimate of model reliability since the reported metrics are not a function of a specific model calibration, and many alternative independent datasets (outer folds) are used for testing (Krstajic et al., 2014). Thus, models

1-4 allowed to examine the effectiveness of the $NDVI_{CV}$ filter on the model performance, and the predictive improvement achieved by using only the selected features along with different parameter combinations than the fixed ones in the FFS process. With model 5 I tested how well a model optimized for homogeneous sites (defined by the selected $NDVI_{CV}$ value threshold) predicted on independent sites that represent both homogeneous and heterogeneous pixels. The best alternative was employed in the subsequent calibration and validation strategies.

After selecting the best approach, I evaluated the predictions by first calibrating the model with LFMC samples from 2000 to 2014 (~80% of the total dataset) and then validated it using the samples collected in 2015-2019 (~20% of the total dataset). That is, I first determined the optimal hyperparameter values for a single model using the samples collected during 2000-2014 by training the algorithm iteratively on one-fifth of the sampling sites and tested on the remaining ones using a standard LLOCV. This process was repeated over 25 random site-resamples for each of the model candidates to stabilize the error rate and eliminate the effect of a particular data partition (Krstajic et al., 2014). The model with the lowest average predictive error was selected and calibrated again to obtain predictions on the whole 5 cross-validation folds. The respective accuracy metrics (called CAL) referred to estimates within the sample period but are not independent from model calibration, as they are the outer-fold metrics in MP. I then evaluated how well the model extrapolates outside the sample period using the samples collected in 2015-2019. This validation phase (named EXT) included some new locations (3 sites) not used in CAL, which means validating future predictions also at unknown points in space.

The final model was used to compare the RF predictions against the RTM estimations produced by Quan et al. (2021). To be a fair comparison, both estimates were contrasted over the same ground-truth samples separately for the $LFMC_{RF}$ predictions inside (CAL) and outside (EXT) the training period.

The optimal hyperparameters for model calibration were chosen from an initial set of possible inputs performing a grid-search scheme (Krstajic et al., 2014). I considered a wider range of possible values (Table A4) of the grid-search scheme for the MP, and then I limited the range according to the results obtained from all fitted models. For CAL, each parameter combination in the grid was iteratively assessed. In the MP, a random subset of combinations (e.g., 50) was implemented at each training process to be more computationally effective. In this case, the choice of hyperparameters was not so important since the cross-validation estimates were a generalization of the model performance.

In all cases, models were optimized to predict on new locations, which is the interest of remote sensing (that is, to estimate LFMC over areas without available ground data), and it prevents spatial over-fitting (Meyer et al., 2018). For MP and CAL grid-search steps, these locations were randomly selected using the method of Meyer et al. (2018), which benefits splitting diversity. In the final model adjustment, prior to predictions, sample-site splitting was conducted by means of their coordinates and the *K*-means algorithm to ensure equal spatial distribution (Pejović et al., 2018).

2.2.5. Validation Methods and Map Production

The predictive capabilities of the model were characterized by means of the root mean square error (RMSE), the mean absolute error (MAE), the mean bias error (MBE), and the unbiased RMSE (ubRMSE), as well as the variance explained by predictive models based on cross-validation (VEcv) (Li, 2016) and the Lin's Concordance Correlation Coefficient (CCC) (Lin, 1989). RMSE, MAE, and MBE measure, respectively, squared, absolute, and mean departures between the estimated $(\hat{y}_i)$ and observed (y_i) test values of LFMC in the same units of the outcomes. RMSE was the statistic used as a criterion for parameter tuning and variable selection processes. I included the ubRMSE following Zhu et al. (2021), which shows the error after removing the tendency to over- or under-predict in the model:

$$ubRMSE = \sqrt{\frac{1}{n}\sum_{1}^{n}(y_i - \hat{y}_i)^2 - \left(\frac{1}{n}\sum_{1}^{n}(y_i - \hat{y}_i)\right)^2} \quad (1)$$

Here, n is the number of observations in a validation dataset. VE_{CV} is similar to the coefficient of determination R^2, but it measures the predictive accuracy of a model by comparing observations and predictions derived from cross-validation and not the square correlation between observed and fitted values. It is defined as:

$$VE_{CV} = 1 - \frac{\sum_1^n (y_i - \hat{y}_i)^2}{\sum_1^n (y_i - \bar{y})^2} \quad (2)$$

where $\bar{y}$ is the mean of the observed values. Otherwise, CCC provides a measure of correlation relative to the line of agreement, which is expected to be unbiased with a slope of 1 and apply a penalty (Cb) if the relationship is far from this line. From CAL, EXT and the RTM, I also obtained the slope and intercept from the linear regression between observed against predicted to assess general deviation trends.

Spatiotemporal analyses were additionally made through land cover types (Rao et al., 2020). More specifically, I calculated general performance metrics from the CAL and EXT estimates for each land cover class, and I decomposed the mean RMSE by land cover and the month of the year to determine the temporal variability of the predictions over each vegetation functional type.

After the validations, I recalibrated the model using the whole dataset in order to consider all the available information to train the algorithm. The readjusted $LFMC_{RF}$ was then used to produce the collection of maps of the reported LFMC database.

2.2.6. Marginal Effects of the Predictors

I used partial dependence plots derived from the fitted model to evaluate the contribution of each variable to the LFMC estimations. The partial dependence function represents the average effect of a given variable on the predicted response marginalized over the effects of the rest of model inputs (Friedman, 2001). Mainly, I divided the distribution of values of the variable of interest into equal steps (e.g., 50). At each step, I calculated the average of all possible predictions made on the data holding the value of the step constant, and I drew a line joining all average points. Resulting plots allowed for the examination of the functional relationships between the most relevant features and the LFMC estimates.

2.2.7. Software and Data Availability

Model building and statistical analysis were made with the statistical software R version 4.2.0 (R Core Team, 2022) and their base package for generic operations. RF was principally implemented with the R package '*ranger'* (Wright & Ziegler, 2017) but also with the '*randomForest'* library (Liaw & Wiener, 2002) to extract the partial dependence plots. The R packages '*raster'* (Hijmans, 2022) and '*sf'* (Pebesma, 2018) were used for remote sensing and spatial data manipulation, and '*doParallel'* (Microsoft-Corp & Weston, 2022) for parallel computing. An adaptation of the *stratfold3d* function of the '*sparsereg3D'* package (Pejović et al., 2018) was used to make the equally spatially distributed LLOCV folds, while the spatially random splits were created with the *CreateSpacetimeFolds* function from the '*CAST'* package (Meyer, 2022).

All the scripts and datasets generated and analyzed during the current study are publicly available in the Github repository, https://github.com/fuegologos/lfmc_rf. The final product of the weekly LFMC maps is fully available at https://doi.org/10.5281/zenodo.6784663.

3. Results

3.1. Selected Variables

Results of the FFS indicated that the most important predictors of LFMC, in terms of error reduction, were the combination of LST and DOY_SIN followed by VARI, NDTI, and DOY_COS (Figure 3). These five variables alone led to an RMSE of 20.1%. I also considered NR3 and NR5 because each one represented on average an additional improvement of ~0.1% in RMSE from the previous stepwise selection, which was greater than the corresponding RMSE standard error (~0.05%) calculated from the 25 FFS step realizations. Selected variables for the subsequent developments reached an overall RMSE of 19.9%.

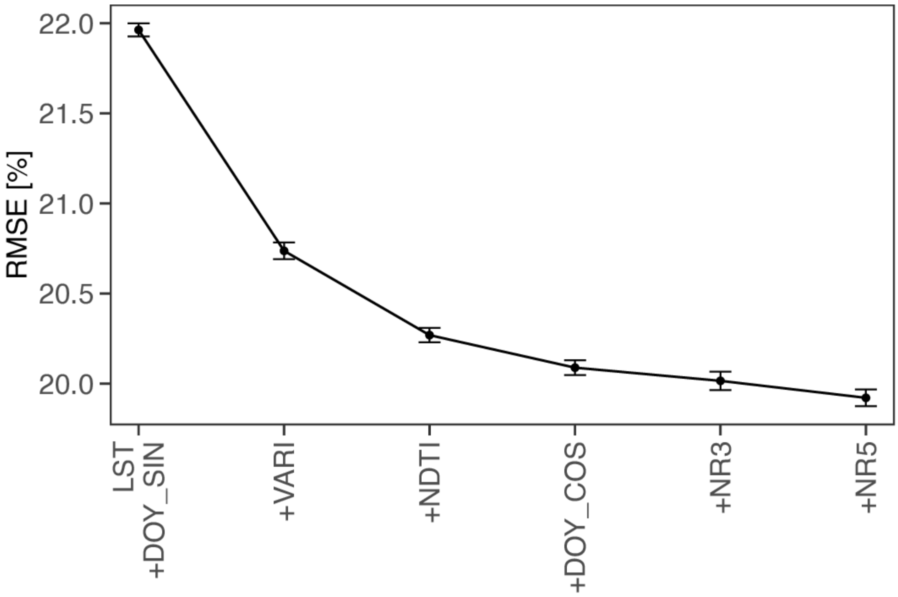

Figure 3. Selected variables derived from the combination of the Forward Feature Selection (FFS) process and the leave-location-out cross-validation (LLOCV). Black dots and vertical segments represent, respectively, the average LLOCV error and the standard error calculated from the 25 RF computed at each FFS step.

3.2. Statistical Performance of the $\mathrm{LFMC}_{\mathrm{RF}}$

Calibrated and evaluated models within the general model performance assessment (MP) achieved similar results among them, with overall RMSEs (that is, from all separate iterations of each MP alternative in conjunction) ranging from 19.1% to 21.4% and $\mathrm{VE}_{\mathrm{CV}}$ ranging from 0.28 to 0.43. Average performance statistics (Table 2) showed that all MP alternatives tended towards a slight overprediction (MBE: 0.9-1.5%). Nonetheless, the ubRMSE values were close to the RMSE (max. difference ~0.07%), further indicating a relatively low bias of the $\mathrm{LFMC}_{\mathrm{RF}}$ estimates. In general, models with all the initial predictors (All*p*) showed worse performance than those with only the selected ones (Sel*p*) (Table 2). The latter benefited from the elimination of the spatially dependent variables, and were used in the successive validation strategies.

Application of the $\mathrm{NDVI}_{\mathrm{CV}}$ filter did not show significant effects on the general model performance (Table 2; Figures B4-B5). For example, applying the optimal filter in Sel*p* to the entire dataset (F1) led to a small improvement in RMSE (<2%) and $\mathrm{VE}_{\mathrm{CV}}$ (~0.01), but also to an increase of MBE (~0.15%) with respect to no filter application. Moreover, comparing MP with no filter and with the filter only applied to the training data (F2) resulted in increases in RMSE and MBE by 0.02% and 0.2%, respectively. In addition, the application of the filter led to the elimination of 26-28% of the dataset. It is worth noting that only a very small percentage of the data (2-4%) was deleted with $\mathrm{NDVI}_{\mathrm{CV}}$ application because they were above the optimal filter threshold (0.3-0.35 in this case). The rest of the data was removed because of missing rows in $\mathrm{NDVI}_{\mathrm{CV}}$, which were derived from poor-quality pixels in Landsat products. The model with no filter was thus used in subsequent analyses.

3.3. Prediction Assessment and Intercomparison

Accuracy metrics from the calibrated model (CAL) were consistent with the general performance (MP) of the $\mathrm{LFMC}_{\mathrm{RF}}$ (Table 2). These results were expected because CAL was developed with 80% of the data employed in MP, but they proved that the adjusted model was not overfitted to the particular data or by the current hyperparameter optimization (Table A4). In contrast, the EXT validation showed smaller RMSE (~3.5%) and higher $\mathrm{VE}_{\mathrm{CV}}$ (~0.15) than CAL, probably due to differences in the validation samples.

According to the previous section (section 3.2), I did not observe any significant

Table 2. Evaluation metrics from predicted and observed values of the model performance (MP), spatial cross-validation (CAL), and the time extrapolation (EXT) assessment. Different methods based on the NDVI_{CV} filter application and the complete (All*p*) or selected (Sel*p*) predictive variables. Predictions from CAL and EXT broken down by fuel type. RTM extractions and LFMC_{RF} were validated on the same ground-truth observations separately if they were used in CAL or EXT.

Method	**Fuel type**	**Variables**	**Filter***	**MBE (%)**	**MAE (%)**	**RMSE (%)**	**ubRMSE (%)**	**CCC**	**VE_{CV}**	**#Testing Samples/ Sites**
MP	All	All*p*	NF	1.10	15.70	20.57	20.54	0.53	0.32	10,374 / 118
	All	All*p*	F1	1.43	15.47	20.29	20.24	0.55	0.35	7633 / 103
	All	Sel*p*	NF	0.86	15.18	19.90	19.88	0.56	0.37	10,374 / 118
	All	Sel*p*	F1	1.00	15.07	19.74	19.71	0.57	0.38	7633 / 103
	All	Sel*p*	F2	1.06	15.18	19.92	19.89	0.57	0.39	7887 / 109
CAL	All	Sel*p*	NF	0.47	15.10	19.93	19.93	0.56	0.37	8983 / 115
	Forests			0.87	14.49	18.32	18.30	0.54	0.33	2633 / 27
	Savannas			1.94	15.22	19.74	19.65	0.51	0.33	4330 / 46
	Shrublands			-7.76	16.20	20.98	19.50	0.53	0.31	442 / 9
	Grasslands			-1.94	15.48	22.57	22.49	0.57	0.36	1578 / 43
EXT	All	Sel*p*	NF	2.75	13.05	16.35	16.12	0.69	0.52	1391 / 43
	Forests			7.40	13.57	16.87	15.16	0.62	0.40	456 / 17
	Savannas			1.63	13.18	16.46	16.38	0.69	0.55	730 / 22
	Shrublands			-4.62	12.08	15.27	14.56	0.72	0.54	166 / 3
	Grasslands			0.86	8.56	12.04	12.01	0.72	0.55	39 / 2
LFMC_{RF} (CAL)	All	Sel*p*	NF	0.86	14.54	18.74	18.73	0.54	0.34	1152 / 68
RTM (CAL)	All	-	-	65.10	66.56	77.78	42.58	0.04	-10.31	1152 / 68
LFMC_{RF} (EXT)	All	Sel*p*	NF	3.88	14.15	17.32	16.88	0.66	0.46	157 / 41
RTM (EXT)	All	-	-	61.87	63.10	74.41	41.33	0.07	-8.98	157 / 41

*NF: no filter; F1: NDVI_{CV} filter applied to the entire dataset (training and test); F2: filter applied only to the training data.

bias in the LFMC_{RF} estimations, as the *y*-intercepts and slopes were close to 0 and 1 in the fitted line between measured and predicted values of LFMC, respectively (Figure 4a,d). However, the residuals between predictions and observations revealed a linear pattern along the range of LFMC in both CAL and EXT (Figure B6). For example, the model highly underestimated values above 120% (MBE CAL=-33.97% and MBE EXT=-22.58%) and overestimated values below 30% (MBE CAL = 45.7%; no data in EXT). This explained the aforementioned better outcomes from EXT, because the range of the actual LFMC for testing (31-209%) excluded values where the model performed worst. Within the LFMC values where live fuels transition from flammable to non-flammable (30-120%), the model attained a smaller RMSE (MAE) of 16.75% (13.35%) for CAL and 15.10% (12.19%) for EXT relative to the overall performance of the corresponding estimates, with a small propensity to overestimate (MBE of 3.30% and 5.24% for CAL and EXT, respectively). It is worth noting that 92% of the data was within the range of 30-120%, and data below 30% may have represented curing.

LFMC$_{RF}$ had better performance than RTM-based estimates when comparing against the same validation samples. In fact, poor correlation and large errors between observed and predicted values occurred in RTM simulations (Table 2). RTM systematically overpredicted LFMC when LFMC exceeded ~76% (Figure 4c,f). Negative values of VE$_{CV}$ (-10.15 and -8.98) indicated that these LFMC estimates were less accurate

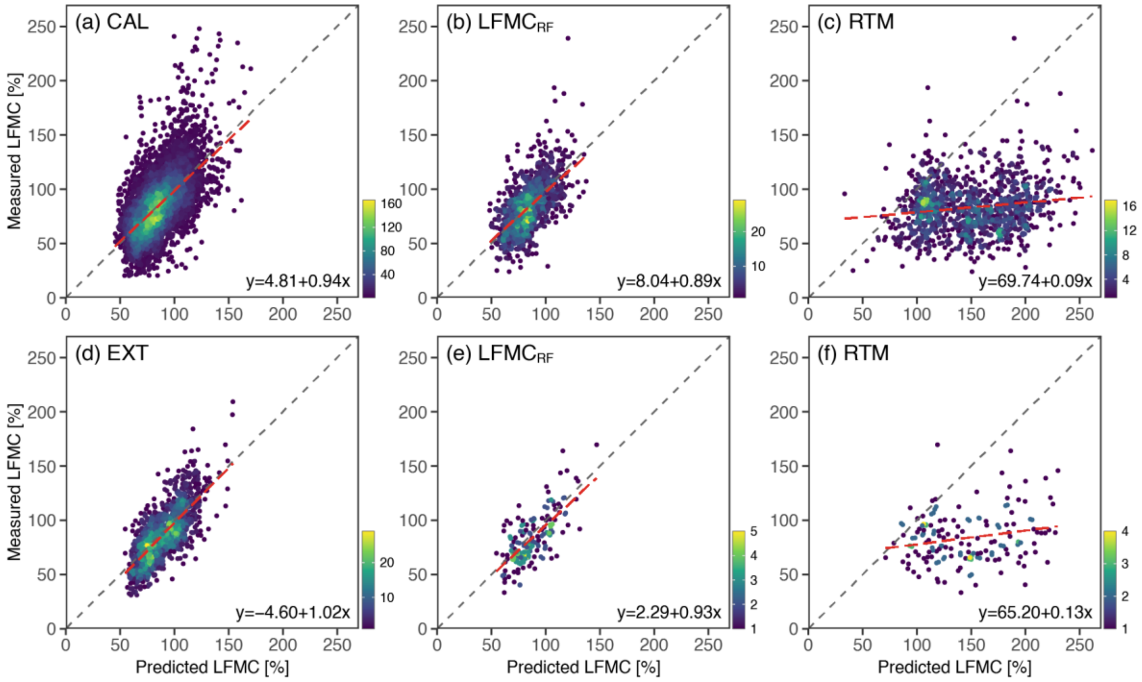

Figure 4. LFMC field measurements versus predictions from CAL (upper plots) and EXT (lower plots): all predictions (a; d), LFMC$_{RF}$ predictions made on the same data points available in RTM (b; e), and the corresponding RTM (c; f). Dashed black line and red line indicates the expected 1:1 relationship and the fitted linear model, respectively. Color scale indicates point density.

than using the mean of observations as predictions. Otherwise, the LFMC$_{RF}$ estimations used for this comparison showed the same level of accuracy as in the previous sections (Figure 4b,e), given that they were subsets of predictions from CAL and EXT.

3.4. Evaluation Across Vegetation Types

Assessing the performance across vegetation types, LFMC$_{RF}$ reached better results in EXT (RMSE: 12-17%; CCC: 0.6-0.7) than in CAL (RMSE: 18-23%; CCC: 0.5-0.6) for all fuel types (Table 2). This coincides with previous results and may be because of the greater range in LFMC variation observed in the CAL dataset (Figure B2). Forests showed the smallest errors in CAL (MBE = 0.87%; RMSE = 18.32%), but the largest in EXT (MBE = 7.40%; RMSE = 16.87%). Grasslands obtained the best performance within the EXT validation (Table 2). However, they represented <3% of the validation records and were mainly concentrated (~80% of the total) in Jul-Aug, where the model performed better (Figure 5c,d). In both cases, LFMC$_{RF}$ significantly underpredicted LFMC in shrublands (MBE -7.76 to -4.62). Temporally, the smallest errors (RMSE: 16-19%) were obtained during the hottest months (Jul-Aug), where field samples were primarily collected (Figure 5a,b), and also in winter months (Jan-Feb), matching with the lowest

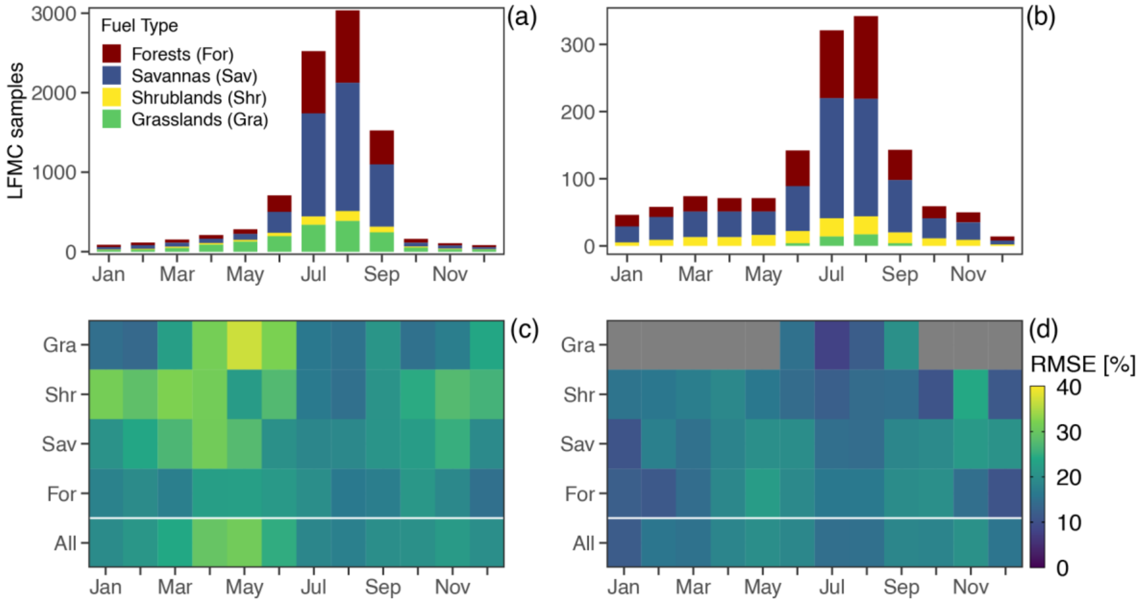

Figure 5. Number of testing samples and RMSE from CAL (a, c) and EXT (b, d) by vegetation type and month of the year. Gray cells in c and d indicate no data available. The IGBP classes from MCD12Q1 were aggregated by the vegetation functional type to which they belong.

LFMC variability (Figure B2d). Forests showed larger stability during the entire year in both RMSE (Figure 5c) and LFMC measurements (Figure B2c). Contrarily, the performance of the model greatly fluctuated in grasslands. Grasslands reported the largest RMSE (36.7%) in May, one of the wettest months of the Mediterranean region, when fires are scarce, declining to an RMSE of 14.9% (Figure 5c) during the driest month.

3.5. Marginal Effects of the Predictors

Partial dependence plots exposed different patterns on the variation of LFMC estimates (Figure 6). VARI and DOY_SIN exerted the strongest effects on predictions. LFMC$_{RF}$ estimates monotonically increased as the VARI values increased. Conversely, LFMC generally monotonically decreased with increases of DOY_SIN, indicating that the

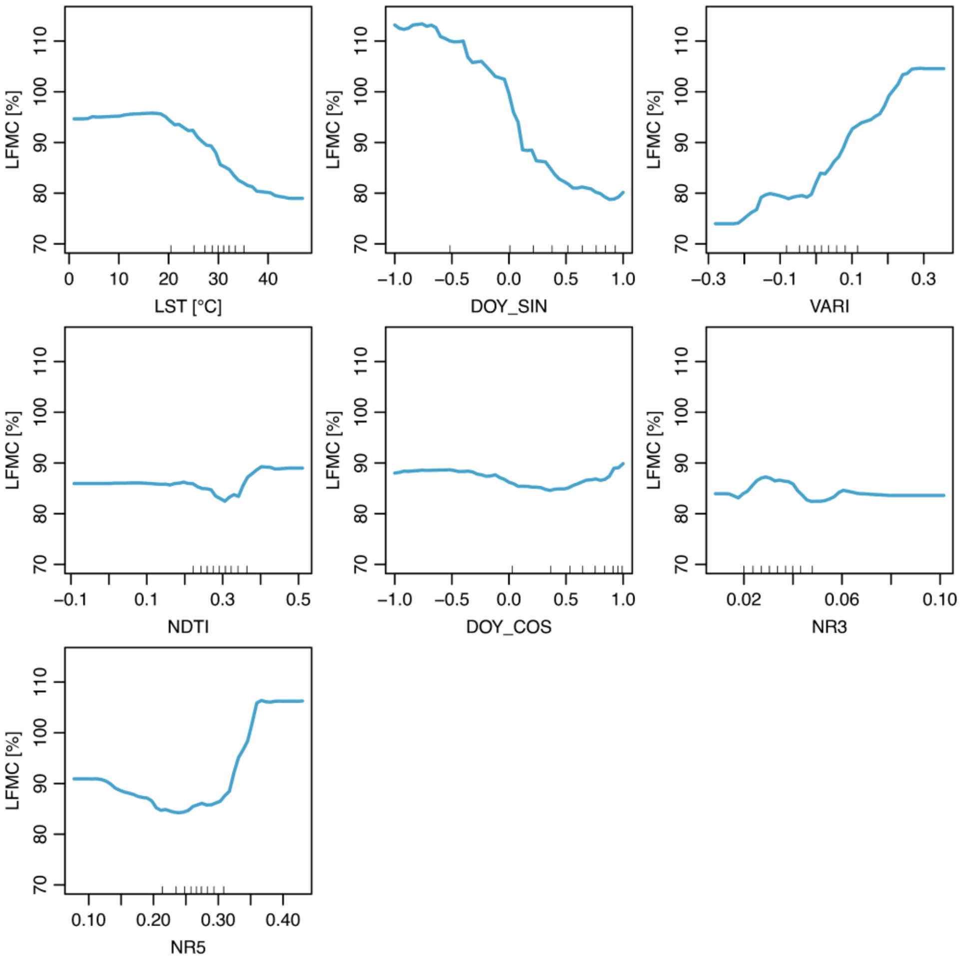

Figure 6. Partial dependence plots from the fitted model. Blue lines describe the average effect of a given predictor in the LFMC estimates. Small lines in the x axis indicate the deciles of the predictor values.

highest LFMC values occurred in spring (-1) and the lowest in late summer (1). LST had non-significant effects on the LFMC estimations up to 20°C but then presented a clear negative relationship. NR5 showed a concave shape, with marked increases at higher values of NR5 (>0.3; last decile). NDTI, NR3, and DOY_COS showed little effects on the predictions of LFMC, but they were still considered informative. The partial dependence of DOY_COS on the LFMC prediction may have been masked by the marginal effects of LST, as they were highly correlated (Figure B3).

4. Discussion

I propose a novel method to estimate LFMC from remote sensing at the subcontinental scale by means of a selected set of remote sensing predictors and the RF algorithm. $LFMC_{RF}$ outperforms current approaches used in the Mediterranean basin in terms of validation errors and provides a solid alternative to predict LFMC over a wide range of environmental conditions using a simple but robust model with a unique formulation. In the next sections, I discuss the contribution of each selected predictor, the general and the spatiotemporal performance of the model, as well as their potential applicability and future improvements.

4.1. Selected Predictors

The key explanatory features resulting from the FFS process were the variables derived from the day of the year (DOY_COS, DOY_SIN), LST, VARI, and NDTI, along with nadir reflectance bands 3 (blue) and 5 (NIR) to a minor degree.

DOY_SIN and DOY_COS had a significative influence on the LFMC estimates due to the seasonal variation in LFMC. In general, LFMC dynamics follow the distribution of the balance between evapotranspiration and rainfall in the Mediterranean region (Argañaraz et al., 2016; Marino et al., 2020; Zhu et al., 2021). DOY_SIN partly reflects the average annual pattern in soil water availability and acts as a complement of the SI, maintaining the periodicity of LFMC within the year. Similarly, DOY_COS reflects changes in the temperature and is more related to vegetation surface temperature, which is measured by the LST (Chuvieco et al., 2004; García et al., 2008).

As I expected, LST was a key factor explaining LFMC, and it showed a negative relationship with LFMC when temperatures were above ~20ºC (Chuvieco et al., 2004; García et al., 2008; McCandless et al., 2020). LST is a key determinant of the energy balance of the vegetation, and its difference with air temperature is related to evapotranspiration and water losses (Vidal et al., 1994). Such differences between air temperature and LST depend on the density of vegetation cover, and previous works have shown strong relationships when combining LST and a vegetation index (e.g., García et al., 2008), as was done here. DOY_COS and LST are complementary because the former keeps the inter-annual variation of LFMC trends, while the latter provides better spatial information (that is, local deviations from the average trends) (García et al., 2008). The partial dependence of LFMC on LST was similar to that reported in previous studies in that LST only affected LFMC after a certain temperature threshold (McCandless et al., 2020). LST is related to vapor pressure deficit (VPD) (Hashimoto et al., 2008), which is a variable that can also affect plant water content as a primary driver of evapotranspiration (Balaguer-Romano et al., 2022). The importance of LST may thus be related to the fact that VPD significantly acts on leaf moisture content after a certain threshold is reached. Therefore, it is also possible that LST could be reflecting local differences in surface temperature and VPD. Further work is needed to fully understand the mechanisms by which LST affects LFMC.

VARI combines different visible wavelength bands (blue-green-red), and it has the ability to detect chlorophyll content and leaf structure variations, which are indirectly associated with changes in canopy moisture (Yebra et al., 2008). Several authors (Caccamo et al., 2012; Fan et al., 2018; Peterson et al., 2008; Roberts et al., 2006; Yebra et al., 2008) have shown that VARI is one of the best indices for predicting LFMC on different vegetation types, and I also demonstrated a notable dependency of the $\mathrm{LFMC_{RF}}$ estimations (Figure 6). Other authors found stronger correlations with indices that include SWIR (Costa-Saura et al., 2021) and NIR (Argañaraz et al., 2016; García et al., 2020) bands predicting LFMC at local scales.

Reductions in chlorophyll content can result from water shortage but also from changes in leaf age, nutrient deficiency, health, and phenological stages (Ceccato et al., 2001; Wang et al., 2013). Introducing NDTI from SWIR bands, in the spectral region greatly sensitive to plant water content (Ceccato et al., 2001; Chuvieco et al., 2002), was necessary to correct for VARI changes not driven by the moisture status of plants.

Moreover, Wang et al. (2013) described a connection (*r* = 0.45) between NDTI and dry matter content of vegetation. Dry matter weight is the denominator of the LFMC equation and could lead to variations in the spectral response and LFMC due to plant seasonal growth, independently of drought changes (Bowyer & Danson, 2004; Fan et al., 2018).

On the other hand, NIR (NR5, centered at 1240 nm) is partly influenced by water content, but also by leaf internal structure and dry biomass (Bowyer & Danson, 2004; Ceccato et al., 2001). This particularity may explain the concave effect that this variable had on predictions. Water loss produces an increment of NR5 as a result of lower water absorption (Chuvieco et al., 2002). However, at certain species and LFMC levels, water stress leads to leaf cell structure changes (reducing reflective areas by wilting) and leaf curling, which cause a decrease in NR5 (Chuvieco et al., 2002; Yebra et al., 2013).

I acknowledge that topography could have affected the results as it alters micro-climatic variables influencing LFMC, such as solar radiation. However, a previous study that used reflectance bands as main explanatory variables (Zhu et al., 2021) indicated a rather small effect on LFMC estimations with an RMSE improvement of ~1%.

4.2. Model Performance Assessment

Generalization errors of the LFMC_{RF} (RMSE: 16-20%; MAE: 13-15%) were lower than in other studies attempting to model LFMC at large spatial scales. For instance, Zhu et al. (2021) reported an overall RMSE of 27.9% using a similar spatial validation strategy but for the contiguous US. They also achieved an RMSE of 22.7% performing a standard cross-validation, which normally results in higher accuracy because the training and testing sets are not spatially independent. LFMC_{RF} also showed smaller RMSE than did Rao et al. (2020) (25%), who used the same spatial approach as Zhu et al. (2021) but ignored multi-species sites with high LFMC seasonal variation, where predictions tend to be more uncertain.

The proposed model tended to underestimate large values and overestimate small values of LFMC (Figure B6). Poor performance of the RF-based model towards the extremes is a well-known problem within RF models (Kuhn & Johnson, 2013). Nonetheless, similar problems were also reported in previous works based on ML (Adab et al., 2016; Zhu et al., 2021), classical regression (Marino et al., 2020; Peterson et al., 2008) and RTM simulation (Yebra et al., 2018) methods. One reason for the systematic bias at high moisture levels can be the lower sensitivity of optical spectra to capture changes in water content while the vegetation gets wet (Quan et al., 2021; Yebra & Chuvieco, 2009). The strategy I adopted to address this problem was to assess LFMC over a very wide range, such that extreme values, those where LFMC estimation is problematic, are largely out of range. The lower level of LFMC in this study was 20%, but fuel moisture below 30% often corresponds to dead fuel (e.g., cured grass) and is thus beyond the scope of the model, since it was designed for LFMC (Yebra et al., 2013). Similarly, the higher LFMC values (above 200%) may be related to harvested samples with the presence of primary tissues from a new vegetative period (Argañaraz et al., 2016), plant parts other than leaves (e.g., fruits, flowers), or the inadequately inclusion of samples collected after rain or dew events (Yebra et al., 2018).

The $LFMC_{RF}$ showed a better performance that RTM predictions from Quan et al. (2021). The RTM-based estimates were highly biased with a strong tendency to overpredict beyond 76% LFMC. This coincided with the results reported by Marino et al. (2020), who found an identical pattern starting at the threshold of 65% using the RTM developed by Yebra et al. (2018). This demonstrates a better predictive power for the $LFMC_{RF}$ approach, even though physically-based approaches are expected to be more precise when applied to sites not used for calibration (Yebra et al., 2013). At any rate, I acknowledge that comparing a regional dataset like this one against a global dataset is not entirely fair, given the scale gap, but the reported results highlight that the RMSE of the global RTM hinders any local application for operational purposes.

The critical LFMC level associated with fire occurrence in the Mediterranean forests, and other parts of the world, occurs around 100% (Dennison & Moritz, 2009; Luo et al., 2019; Nolan et al., 2016). The $LFMC_{RF}$ model improves current products, but MAE around the critical threshold of 100% LFMC is still ~13%. Differences of 10% in LFMC estimation from field measurements are generally acceptable for fire management (Marino et al., 2020). However, these results indicate that there is still room for further improvements, particularly towards the critical threshold, so as to avoid reporting of false fire alerts or omission of danger situations (Yebra et al., 2008).

4.3. Evaluation Across Vegetation Types

The predictive errors obtained by the $LFMC_{RF}$ within the training period for forests/savannas (RMSE 18-20%), shrublands (RMSE ~21%), and grasslands (RMSE ~23%) were similar between them and comparable to those reported by other studies for the same vegetation types (forests/savannas 22-32%, shrublands 14-29%, grasslands 29-49%) (Jurdao et al., 2013; Quan et al., 2021; Rao et al., 2020; Yebra et al., 2018; Zhu et al., 2021). Despite the methodological differences, this comparison demonstrates that a single model can be as accurate or even better than formalizing a model for each fuel class separately. This could be due to the RF architecture that allows using the spectral and thermal information itself to discriminate between vegetation functional types. Furthermore, misclassification problems of the land cover products used to differentiate between fuel classes can further increase the uncertainty of the LFMC estimates (Yebra et al., 2018; Zhu et al., 2021).

In general, I observed that the uncertainty of the LFMC predictions (Figure 5) depended on the range of LFMC values for testing and their local and temporal variability (Figure B2). For example, forests showed more stable behavior in both LFMC dynamics and prediction agreement. Deep root systems in trees reduce the seasonal LFMC variation (Resco de Dios, 2020). On the contrary, grasslands reported the highest errors in spring (the wettest part of the year) and the lowest in the driest periods (summer, when fires are more likely). These patterns overlapped with the monthly maximum and minimum values of LFMC, that is, larger LFMC errors under higher LFMC values and smaller LFMC errors under lower values. Shrublands instead had a low temporal variability but presented a significant bias (MBE from -5 to -8%), likely because of the high proportion of large LFMC values (>120%) in their ground-truth sample distributions (16.4% of the total ground-truth samples). The error associated with predictions outside the training period (EXT) was similar to that from the CAL dataset (Figure 5). However, RMSE was slightly lower with the EXT dataset because of the lower LFMC variability in the EXT dataset relative to CAL. I thus conclude that the fitted model with historical data can be safely applied in future situations without the need for frequent readjustment, but with careful interpretation in the wettest months and for LFMC values below 30%.

4.4. Applicability and Potential Improvements

The relatively coarse resolution (~500 m) of the final product is appropriate for landscape-scale use and does not guarantee smaller-scale applications. Each individual pixel normally contains information from a mixture of vegetation canopy layers, species, surface litter, and soil elements with different properties that cannot be unambiguously interpreted (Chuvieco et al., 2020; Yebra et al., 2018). I acknowledge that a limitation to this study is that I did not explicitly assess the representativeness of the samples within the site. I therefore took into account small-scale heterogeneity by implementing an $NDVI_{CV}$ filter, as in Quan et al. (2021). However, I did not observe any significant improvement after applying this filter, likely indicating that sample areas were relatively homogenous. In any case, sub-pixel variation and the scale mismatch between sample-plot size and pixel resolution hinder establishing relationships between field observations and satellite-derived variables, introducing uncertainties into the predictions. The latter could be solved using higher spatial resolution data (e.g., Sentinel-2 or Landsat) (Adab et al., 2016; Costa-Saura et al., 2021; Marino et al., 2020), but these satellites usually have lower revisit frequency disabling near-real-time usage and introducing a time lag between the images and the sampling date (Marino et al., 2020). Future work should extend the use of the $LFMC_{RF}$ methods to these newer satellites because historical LFMC field data currently available is not yet sufficient to achieve this goal.

Further progress will come from joining the $LFMC_{RF}$ approach with microwave remote sensing data. Microwave observations (active and passive) can also detect changes in vegetation water content but are less sensitive to atmospheric conditions (e.g., clouds) than optical wavelengths (Fan et al., 2018) and have the ability to penetrate deeper into the canopies (Rao et al., 2020). The recently available non-commercial radar data supplied by the Sentinel-1 A/B Synthetic Aperture Radar (SAR) may represent a great opportunity to infer the improvement of LFMC models at the operational level (Rao et al., 2020; Wang et al., 2019).

Sample representativeness is a general constrain in the empirical models (Yebra et al., 2013). In this study, field samples were not evenly distributed across the whole Mediterranean basin. They could be considered representative of the Western Mediterranean conditions since they were abundant in number (space and time) within their specific biome, as well as in species and environmental conditions. However, LFMC predictions in temperate zones should be interpreted with caution as they were underrepresented. Thus, the application of the $LFMC_{RF}$ should be limited to areas with similar characteristics, and LFMC estimates must be interpreted with caution in underrepresented areas (e.g., temperate zones). Despite that, the generated maps extend to the entire Mediterranean biome included in the Mediterranean basin, as well as some meridional areas of the temperate biomes of Europe (e.g., northern Spain) (Figure 2).

5. Conclusions

I successfully tested an RF algorithm as an approach to predict large-scale LFMC using the spectral and thermal information of MODIS and two static variables representing seasonal patterns. The $LFMC_{RF}$ is applicable to a wide variety of vegetation types, and the performance of the fitted model (MBE = 0.47%, RMSE = 19.9%, VE_{CV} = 0.37, CCC = 0.56) was comparable to that of other studies with similar purposes but with a higher degree of complexity than $LFMC_{RF}$, including the RTM-based methods with applications in the Mediterranean basin. The architecture of RF allows the introduction of new explanatory variables that would help to reduce the uncertainty in the predictions. LFMC maps were produced at 8-day intervals from 2001 to 2021. The final product provides a complete asset for studying the relationships between LFMC and the influencing factors that promote wildfire activity and fire regimes in the Mediterranean basin. Furthermore, after the imminent MODIS decommission, the new Visible Infrared Imaging Radiometer Suite (VIIRS) is expected to provide long-term continuity with better spatial resolution (Yebra et al., 2013). Continuous retrievals, either with MODIS or VIIRS, might be a valuable tool for quasi near-real-time fire risk assessment and for operational applications such as the mobilization of resources and people or the planning of preventive actions for fire mitigation (e.g., fuel reduction or prescribed burns).

6. References

Adab, H., Devi Kanniah, K., & 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*(11), 961. https://doi.org/10.3390/rs8110961

Argañaraz, J. P., Landi, M. A., Bravo, S. J., Gavier-Pizarro, G. I., Scavuzzo, C. M., &

Bellis, L. M. (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. https://doi.org/10.1109/JSTARS.2016.2575366

Balaguer-Romano, R., Díaz-Sierra, R., De Cáceres, M., Cunill-Camprubí, À., Nolan, R. H., Boer, M. M., Voltas, J., & Resco de Dios, V. (2022). A semi-mechanistic model for predicting daily variations in species-level live fuel moisture content. *Agric. For. Meteorol.*, *323*(June), 109022. https://doi.org/10.1016/j.agrformet.2022.109022

Boer, M. M., Nolan, R. H., Resco De Dios, V., Clarke, H., Price, O. F., & Bradstock, R. A. (2017). Changing Weather Extremes Call for Early Warning of Potential for Catastrophic Fire. *Earth's Futur.*, *5*(12), 1196–1202. https://doi.org/10.1002/2017EF000657

Bowman, D. M. J. S., Kolden, C. A., Abatzoglou, J. T., Johnston, F. H., van der Werf, G. R., & Flannigan, M. (2020). Vegetation fires in the Anthropocene. *Nat. Rev. Earth Environ.*, *1*(10), 500–515. https://doi.org/10.1038/s43017-020-0085-3

Bowyer, P., & Danson, F. M. (2004). Sensitivity of spectral reflectance to variation in live fuel moisture content at leaf and canopy level. *Remote Sens. Environ.*, *92*(3), 297–308. https://doi.org/10.1016/j.rse.2004.05.020

Bradstock, R. A. (2010). A biogeographic model of fire regimes in Australia: current and future implications. *Glob. Ecol. Biogeogr.*, *19*(2), 145–158. https://doi.org/10.1111/j.1466-8238.2009.00512.x

Breiman, L. (2001). Random Forests. *Mach. Learn.*, *45*, 5–32. https://doi.org/10.1201/9780367816377-11

Caccamo, G., Chisholm, L. A., Bradstock, R. A., Puotinen, M. L., & Pippen, B. G. (2012). Monitoring live fuel moisture content of heathland, shrubland and sclerophyll forest in south-eastern Australia using MODIS data. *Int. J. Wildl. Fire*, *21*(3), 257. https://doi.org/10.1071/WF11024

Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S., & Grégoire, J.-M. (2001). Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sens. Environ.*, *77*(1), 22–33. https://doi.org/10.1016/S0034-4257(01)00191-2

Chuvieco, E., Aguado, I., Salas, J., García, M., Yebra, M., & Oliva, P. (2020). Satellite

Remote Sensing Contributions to Wildland Fire Science and Management. *Curr. For. Reports*, *6*(2), 81–96. https://doi.org/10.1007/s40725-020-00116-5

Chuvieco, E., Cocero, D., Riaño, D., Martin, P., Martínez-Vega, J., de la Riva, J., & Pérez, F. (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. https://doi.org/10.1016/j.rse.2004.01.019

Chuvieco, E., Riaño, D., Aguado, I., & Cocero, D. (2002). Estimation of fuel moisture content from multitemporal analysis of Landsat Thematic Mapper reflectance data: Applications in fire danger assessment. *Int. J. Remote Sens.*, *23*(11), 2145–2162. https://doi.org/10.1080/01431160110069818

Congalton, R. G., & Green, K. (2019). *Assessing the Accuracy of Remotely Sensed Data*. CRC Press. https://doi.org/10.1201/9780429052729

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), 3726. https://doi.org/10.3390/rs13183726

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. Wildl. Fire*, *18*(8), 1021. https://doi.org/10.1071/WF08055

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., Hansen, M., Locke, H., Ellis, E. C., Jones, B., Barber, C. V., Hayes, R., Kormos, C., Martin, V., Crist, E., … Saleem, M. (2017). An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *Bioscience*, *67*(6), 534–545. https://doi.org/10.1093/biosci/bix014

Duane, A., Castellnou, M., & Brotons, L. (2021). Towards a comprehensive look at global drivers of novel extreme wildfire events. *Clim. Change*, *165*(3–4), 43. https://doi.org/10.1007/s10584-021-03066-4

Dupuy, J., Fargeon, H., Martin-StPaul, N., Pimont, F., Ruffault, J., Guijarro, M., Hernando, C., Madrigal, J., & Fernandes, P. (2020). Climate change impact on future wildfire danger and activity in southern Europe: a review. *Ann. For. Sci.*, *77*(2), 35. https://doi.org/10.1007/s13595-020-00933-5

Fan, L., Wigneron, J.-P., Xiao, Q., Al-Yaari, A., Wen, J., Martin-StPaul, N., Dupuy, J.-L., Pimont, F., Al Bitar, A., Fernandez-Moran, R., & Kerr, Y. H. (2018). Evaluation of microwave remote sensing for monitoring live fuel moisture content in the Mediterranean region. *Remote Sens. Environ.*, *205*(November 2017), 210–223. https://doi.org/10.1016/j.rse.2017.11.020

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Ann. Stat.*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Gabriel, E., Delgado-Dávila, R., De Cáceres, M., Casals, P., Tudela, A., & Castro, X. (2021). Live fuel moisture content time series in Catalonia since 1998. *Ann. For. Sci.*, *78*(2), 44. https://doi.org/10.1007/s13595-021-01057-0

García, M., Chuvieco, E., Nieto, H., & Aguado, I. (2008). Combining AVHRR and meteorological data for estimating live fuel moisture content. *Remote Sens. Environ.*, *112*(9), 3618–3627. https://doi.org/10.1016/j.rse.2008.05.002

García, M., Riaño, D., Yebra, M., Salas, J., Cardil, A., Monedero, S., Ramirez, J., Martín, M. P., Vilar, L., Gajardo, J., & Ustin, S. (2020). A live fuel moisture content product from landsat TM satellite time series for implementation in fire behavior models. *Remote Sens.*, *12*(11). https://doi.org/10.3390/rs12111714

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Hashimoto, H., Dungan, J., White, M. A., Yang, F., Michaelis, A., Running, S. W., & Nemani, R. (2008). Satellite-based estimation of surface vapor pressure deficits using MODIS land surface temperature data. *Remote Sens. Environ.*, *112*(1), 142–155. https://doi.org/10.1016/j.rse.2007.04.016

Hijmans, R. J. (2022). *raster: Geographic Data Analysis and Modeling*. R Package Version 3.5-15. https://cran.r-project.org/package=raster

IPCC. (2022). *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (H.-O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, & B. Rama (eds.)).

Jolly, W., & Johnson, D. (2018). Pyro-Ecophysiology: Shifting the Paradigm of Live

Wildland Fuel Research. *Fire*, *1*(1), 8. https://doi.org/10.3390/fire1010008

Jurdao, S., Yebra, M., Guerschman, J. P., & Chuvieco, E. (2013). Regional estimation of woodland moisture content by inverting Radiative Transfer Models. *Remote Sens. Environ.*, *132*, 59–70. https://doi.org/10.1016/j.rse.2013.01.004

Karavani, A., Boer, M. M., Baudena, M., Colinas, C., Díaz-Sierra, R., Pemán, J., de Luis, M., Enríquez-de-Salamanca, Á., & Resco de Dios, V. (2018). Fire-induced deforestation in drought-prone Mediterranean forests: drivers and unknowns from leaves to communities. *Ecol. Monogr.*, *88*(2), 141–169. https://doi.org/10.1002/ecm.1285

Krstajic, D., Buturovic, L. J., Leahy, D. E., & Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *J. Cheminform.*, *6*(1), 10. https://doi.org/10.1186/1758-2946-6-10

Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. Springer New York. https://doi.org/10.1007/978-1-4614-6849-3

Li, J. (2016). Assessing spatial predictive models in the environmental sciences: Accuracy measures, data variation and variance explained. *Environ. Model. Softw.*, *80*, 1–8. https://doi.org/10.1016/j.envsoft.2016.02.004

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*(3), 18–22. https://cran.r-project.org/doc/Rnews/

Lin, L. I.-K. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics*, *45*(1), 255. https://doi.org/10.2307/2532051

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), 1–17. https://doi.org/10.3390/f10100887

Marino, E., Yebra, M., Guillén-Climent, M., Algeet, N., Tomé, J. L., Madrigal, J., Guijarro, M., & Hernando, C. (2020). Investigating Live Fuel Moisture Content Estimation in Fire-Prone Shrubland from Remote Sensing Using Empirical Modelling and RTM Simulations. *Remote Sens.*, *12*(14), 2251. https://doi.org/10.3390/rs12142251

Martin-StPaul, N., Pimont, F., Dupuy, J. L., Rigolot, E., Ruffault, J., Fargeon, H., Cabane, E., Duché, Y., Savazzi, R., & Toutchkov, M. (2018). Live fuel moisture content (LFMC) time series for multiple sites and species in the French Mediterranean area since 1996. *Ann. For. Sci.*, *75*(3), 70. https://doi.org/10.1007/s13595-018-0744-4

McCandless, T. C., Kosovic, B., & Petzke, W. (2020). Enhancing wildfire spread modelling by building a gridded fuel moisture content product with machine learning. *Mach. Learn. Sci. Technol.*, *1*(3), 035010. https://doi.org/10.1088/2632-2153/aba480

Meyer, H. (2022). *CAST: "caret" Applications for Spatial-Temporal Models*. Package Version 6.0-92. https://cran.r-project.org/package=CAST

Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., & Nauss, T. (2018). Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environ. Model. Softw.*, *101*, 1–9. https://doi.org/10.1016/j.envsoft.2017.12.001

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. Modell.*, *411*(March), 108815. https://doi.org/10.1016/j.ecolmodel.2019.108815

Microsoft-Corp, & Weston, S. (2022). *doParallel: Foreach Parallel Adaptor for the "parallel" Package*. Package, R package version 1.0.17. https://cran.r-project.org/package=doParallel

Nelson, R. M. (2001). Water Relations of Forest Fuels. In E. E. A. Johnson and K. Miyanishi (Ed.), *Forest fires: behavior and ecological effects* (pp. 79–149). Academic Press.

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. https://doi.org/10.1002/2016GL068614

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *R J.*, *10*(1), 439–446. https://doi.org/10.32614/RJ-2018-009

Pejović, M., Nikolić, M., Heuvelink, G. B. M., Hengl, T., Kilibarda, M., & Bajat, B. (2018). Sparse regression interaction models for spatial prediction of soil properties in 3D. *Comput. Geosci.*, *118*, 1–13. https://doi.org/10.1016/j.cageo.2018.05.008

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. https://doi.org/10.1016/j.rse.2008.07.012

Quan, X., Yebra, M., Riaño, D., He, B., Lai, G., & Liu, X. (2021). Global fuel moisture content mapping from MODIS. *Int. J. Appl. Earth Obs. Geoinf.*, *101*(December 2020), 102354. https://doi.org/10.1016/j.jag.2021.102354

R Core Team. (2022). *R: A language and environment for statistical computing* (4.2.0). R Foundation for Statistical Computing. https://www.r-project.org/

Rao, K., Williams, A. P., Flefil, J. F., & Konings, A. G. (2020). SAR-enhanced mapping of live fuel moisture content. *Remote Sens. Environ.*, *245*(March), 111797. https://doi.org/10.1016/j.rse.2020.111797

Resco de Dios, V. (2020). *Plant-Fire Interactions: Applying Ecophysiology to Wildfire Management, Managing Forest Ecosystems* (Vol. 36). Springer International Publishing. https://doi.org/10.1007/978-3-030-41192-3

Roberts, D. A., Dennison, P. E., Peterson, S., Sweeney, S., & Rechel, J. (2006). Evaluation of Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and Moderate Resolution Imaging Spectrometer (MODIS) measures of live fuel moisture and fuel condition in a shrubland ecosystem in southern California. *J. Geophys. Res. Biogeosciences*, *111*(G4), 1–16. https://doi.org/10.1029/2005JG000113

Ruffault, J., Martin-StPaul, N., Pimont, F., & Dupuy, J.-L. 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*(June), 391–401. https://doi.org/10.1016/j.agrformet.2018.07.031

Schaaf, C. B., & Wang, Z. (2015). *MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global - 500m V006 [Data set]*. NASA EOSDIS Land Processes DAAC. https://doi.org/https://doi.org/10.5067/MODIS/MCD43A4.006

Soler Martin, M., Bonet, J. A., Martínez De Aragón, J., Voltas, J., Coll, L., & Resco De Dios, V. (2017). Crown bulk density and fuel moisture dynamics in Pinus pinaster stands are neither modified by thinning nor captured by the Forest Fire Weather Index. *Ann. For. Sci.*, *74*(3), 51. https://doi.org/10.1007/s13595-017-0650-1

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. https://doi.org/10.3390/rs5062617

Sulla-Menashe, D., Gray, J. M., Abercrombie, S. P., & Friedl, M. A. (2019). Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product. *Remote Sens. Environ.*, *222*, 183–194. https://doi.org/10.1016/j.rse.2018.12.013

Van Wagner, C. E. (1987). Development and structure of the Canadian forest fire weather index system. Canadian Forestry Service. *For. Tech. Rep.*, *35*, 37.

Vidal, A., Pinglo, F., Durand, H., Devaux-Ros, C., & Maillet, A. (1994). Evaluation of a temporal fire risk index in mediterranean forests from NOAA thermal IR. *Remote Sens. Environ.*, *49*(3), 296–303. https://doi.org/10.1016/0034-4257(94)90024-8

Wan, Z. (2014). New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote Sens. Environ.*, *140*, 36–45. https://doi.org/10.1016/j.rse.2013.08.027

Wang, L., Hunt, E. R., Qu, J. J., Hao, X., & Daughtry, C. S. T. (2013). Remote sensing of fuel moisture content from ratios of narrow-band vegetation water and dry-matter indices. *Remote Sens. Environ.*, *129*, 103–110. https://doi.org/10.1016/j.rse.2012.10.027

Wang, L., Quan, X., He, B., Yebra, M., Xing, M., & Liu, X. (2019). Assessment of the Dual Polarimetric Sentinel-1A Data for Forest Fuel Moisture Content Estimation. *Remote Sens.*, *11*(13), 1568. https://doi.org/10.3390/rs11131568

Wright, M. N., & Ziegler, A. (2017). ranger : A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *J. Stat. Softw.*, *77*(1). https://doi.org/10.18637/jss.v077.i01

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. https://doi.org/10.1016/j.rse.2009.07.001

Yebra, M., Chuvieco, E., & Riaño, D. (2008). Estimation of live fuel moisture content from MODIS images for fire risk assessment. *Agric. For. Meteorol.*, *148*(4), 523–536. https://doi.org/10.1016/j.agrformet.2007.12.005

Yebra, M., Dennison, P. E., Chuvieco, E., Riaño, D., Zylstra, P., Hunt, E. R., Danson, F.

M., Qi, Y., & Jurdao, S. (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. https://doi.org/10.1016/j.rse.2013.05.029

Yebra, M., Quan, X., Riaño, D., Rozas Larraondo, P., van Dijk, A. I. J. M., & Cary, G. J. (2018). A fuel moisture content and flammability monitoring methodology for continental Australia based on optical remote sensing. *Remote Sens. Environ.*, *212*(April), 260–272. https://doi.org/10.1016/j.rse.2018.04.053

Yebra, M., Scortechini, G., Badi, A., Beget, M. E., Boer, M. M., Bradstock, R., Chuvieco, E., Danson, F. M., Dennison, P., Resco de Dios, V., Di Bella, C. M., Forsyth, G., Frost, P., Garcia, M., Hamdi, A., He, B., Jolly, M., Kraaij, T., Martín, M. P., … Ustin, S. (2019). Globe-LFMC, a global plant water status database for vegetation ecophysiology and wildfire applications. *Sci. Data*, *6*(1), 155. https://doi.org/10.1038/s41597-019-0164-9

Zhu, L., Webb, G. I., Yebra, M., Scortechini, G., Miller, L., & Petitjean, F. (2021). Live fuel moisture content estimation from MODIS: A deep learning approach. *ISPRS J. Photogramm. Remote Sens.*, *179*(August), 81–91. https://doi.org/10.1016/j.isprsjprs.2021.07.010

Appendix A. Supplementary Methods

A1. Land Surface Temperature

I performed preliminary tests to specify the LST product that would be used in the development of the model. The comparison was made using the Forward Feature Selection (FFS) process described in section 2.2.3. Specifically, I tested the ability of the MODIS LST daily (MOD11A1) and the 8-day average composite (MOD11A2) products to predict LFMC along with the other variables in a variable selection process, where its usefulness is also evaluated. MOD11A1 data were obtained for each sample site by their sampling date. MOD11A2 were extracted from the composite layer which includes, on their averaged days, the corresponding sampling date. The number of missing data was much greater in MOD11A1 than in MOD11A2 (34.7% and 7.7%, respectively). In order to be a fair comparison, I formed a single dataset of equal size by eliminating missing values in both variables. The two LST variables were selected as predictors in the FFS. The final models reached an RMSE of 20.38% and 20.27% for the daily LST and 8-day composite product, respectively. Given that RMSEs were very similar between both models, and that the 8-days LST composite showed smaller data gaps, I used the latter.

A2. Data Extraction Method

I previously tested the method for remote-sensing data extraction at the sampling sites. In particular, I compared the performance of models with all predictor variables obtained from a simple pixel extraction or the average value from the 3×3 pixels window closest to each field sampling location as performed in Quan et al. (2021). To do so, I used a nested 5-fold leave-location-out cross-validation and 30 iterations, such as I described in section 2.2.4. Average mean bias error (MBE), root mean square error (RMSE) and variance explained by cross-validation (VE_{CV}) are shown in **Table A1**. Additionally, I applied a Chi-square test using the distributions of RMSE and VE_{CV} values from the 30 repetitions to statistically check differences. I concluded that there were no significant differences between both methods (RMSE *p*-value = 0.9; VE_{CV} *p*-value = 0.88). Averaging of adjacent pixels to avoid positioning errors may be considered advantageous at higher spatial resolutions (e.g., Sentinel-2) than those treated here (Congalton & Green, 2019) and I used simple pixel extraction for simplicity and shortest computation time.

Table A1. Performance metrics from the focal mean and simple pixel extraction comparison.

Method	**MBE [%]**	**RMSE [%]**	**VE$_{CV}$**
Focal mean	1.33	20.52	0.33
Simple pixel	1.05	20.53	0.33

A3. Spectral Vegetation Indices

Table A2. Spectral vegetation indices used to estimate LFMC based on the MCD43A4 Collection 6 reflectance bands: B1, Red; B2, NIR1; B3, Blue; B4, Green; B5, NIR2; B6, SWIR1; B7, SWIR2. MODIS formulations extracted from the literature cited in the main text.

Index	**Formulation**	**Reference**
Normalized Difference Vegetation Index	$NDVI = \frac{B2 - B1}{B2 + B1}$	Rouse et al. (1974)
Normalized Difference Water Index	$NDWI = \frac{B2 - B5}{B2 + B5}$	Gao (1996)
Normalized Difference Infrared Index	$NDII6 = \frac{B2 - B6}{B2 + B6}$	Hardisky et al. (1983)
Normalized Difference Infrared Index (with band 7)	$NDII7 = \frac{B2 - B7}{B2 + B7}$	Hardisky et al. (1983)
Global Vegetation Moisture Index	$GVMI = \frac{(B2 + 0.1) - (B6 + 0.02)}{(B2 + 0.1) + (B6 + 0.02)}$	Ceccato et al. (2002)
Enhanced Vegetation Index	$EVI = \frac{2.5 \times (B2 - B1)}{B2 + 6 \times B1 - 7.5 \times B3 + 1}$	Huete et al. (2002)
Soil Adjusted Vegetation Index	$SAVI = \frac{(1 + 0.5)(B2 - B1)}{B2 + B1 + 0.5}$	Huete (1988)
Visible Atmospherically Resistant Index	$VARI = \frac{B4 - B1}{B4 + B1 - B3}$	Gitelson et al. (2002)
Vegetation Index — Green	$VI_{green} = \frac{B4 - B1}{B4 + B1}$	Tucker (1979)
Normalized Difference Tillage Index	$NDTI = \frac{B6 - B7}{B6 + B7}$	van Deventer et al. (1997)
Simple Tillage Index	$STI = B6/B7$	van Deventer et al. (1997)
Moisture Stress Index	$MSI = B6/B2$	Rock et al. (1986)
Greenness index	$G_{ratio} = B4/B1$	Zarco-Tejeda et al. (2005)

A4. Land Cover Definitions

Table A3. Land cover classes from samples used in the study. International Geosphere-Biosphere Programme (IGBP) definitions and corresponding grouped classes for the analyses.

IGBP class	Definition	Grouped class
Evergreen needleleaf forests	Dominated by evergreen conifer trees (canopy >2 m). Tree cover >60%.	Forests
Evergreen broadleaf forests	Dominated by evergreen broadleaf and palmate trees (canopy >2 m). Tree cover >60%.	Forests
Mixed forests	Dominated by neither deciduous nor evergreen (40-60% of each) tree type (canopy >2 m). Tree cover >60%.	Forests
Woody savannas	Tree cover 30-60% (canopy >2 m).	Savannas
Savannas	Tree cover 10-30% (canopy >2 m).	Savannas
Open shrublands	Dominated by woody perennials (1-2 m height) 10-60% cover.	Shrublands
Closed shrublands	Dominated by woody perennials (1-2 m height) >60% cover.	Shrublands
Grasslands	Dominated by herbaceous annuals (<2 m).	Grasslands

A5. Model Parametrization

The Random Forest algorithm requires specification of some hyperparameters for model calibration. The parameters considered here were the number of variables randomly selected at each split (*mtry*), the total number of trees in the forest (*ntree*), the minimal terminal node size (*min. node size*), and the ratio of observations sampled for each decision tree (*sample fraction*). Samples selection was made without replacement except for the FFS process. In this case, *mtry* was set to 2, as suggested by Meyer et al. (2019). The number of trees was fixed to 250 to reduce the computational time, as I showed no significant increase of performance by using more trees. The rest of parameters were left as configured by default in the RF algorithm. Table A4 shows the range of parameters used in model development and the optimal parameters chosen after model tuning.

Table A4. Boundaries of the RF hyperparameters grid-search space, adjusted parameters for the Forward Feature Selection (FFS) process and optimized hyperparameters for the final model.

Step	**Type**	**$NDVI_{CV}$**	**ntree**	**mtry**	**min. node size**	**sample fraction**
Grid-Search	Start	0.20	250	2	5	0.2
	End	0.60	1000	p × 0.40 p min = 4	30	0.95
	Step	0.05	250	1	1	0.05
FFS	Selected	-	250	2	5	0.632*
CAL	Optimal	-	500	2	5	0.3

*Sample with replacement.; p, predictor variables.

Appendix B. Supplementary Analyses

B1. Data Description

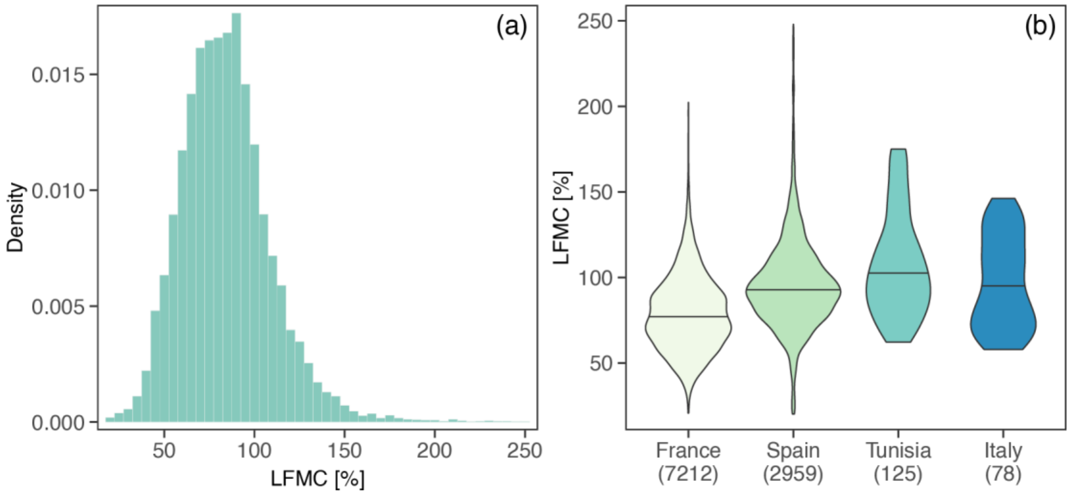

Figure B1. LFMC ground samples overall (a) and by country (b) distributions. Numbers between parenthesis under country names are the number of samples in that country.

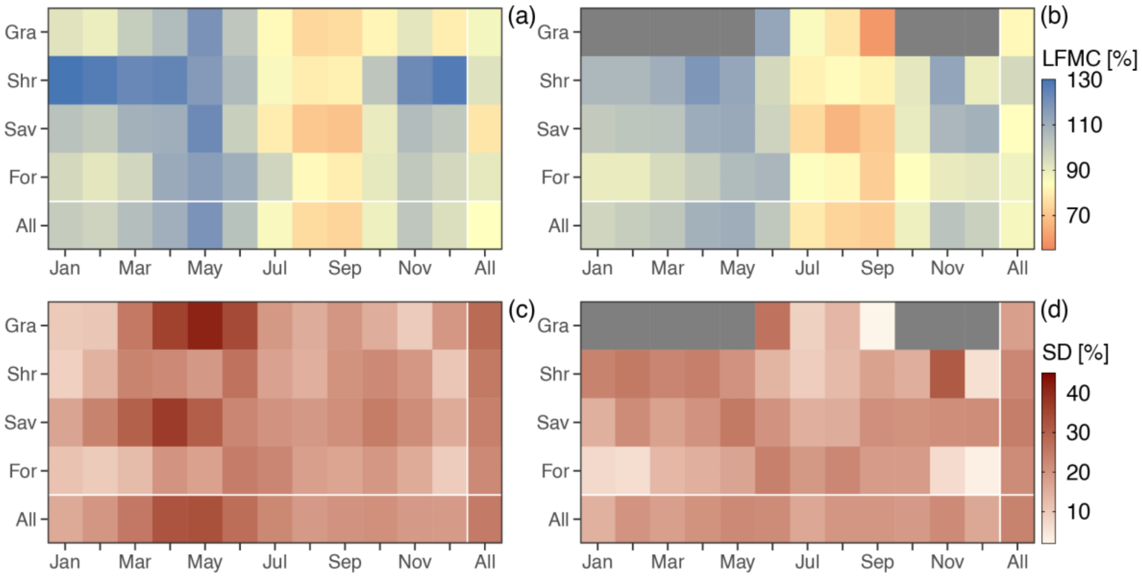

Figure B2. Mean and standard deviation (SD) matrices from CAL (a, c) and EXT (b, d) of the LFMC field measurements showed by fuel type and month of the year, and the overall of each one. Gray cells indicate no data availability.

B2. Features Correlation

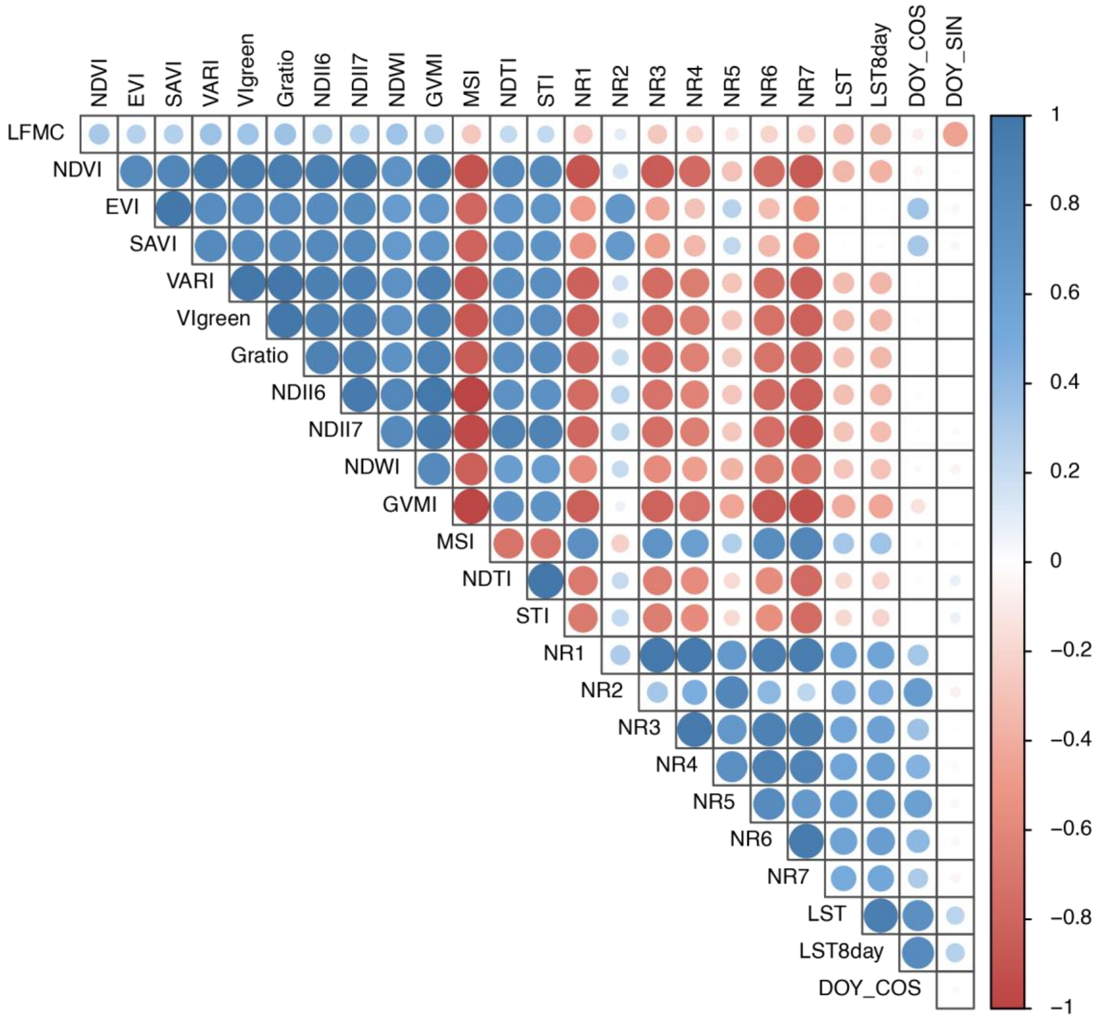

Figure B3. Correlation matrix between LFMC and predictive variables.

B3. NDVI$_{CV}$ Filter

The optimal values for the application of the NDVI$_{CV}$ filter were in the range of 0.3-0.4 (Figure B4). I additionally examined LFMC predictions made with the calibrated LFMC$_{RF}$ model against observations that were discarded by a NDVI$_{CV}$ threshold value of 0.3 (Figure B5). The most error estimates were around the mean absolute error of the model (MAE = 15.10%). So, the filter did not discriminate bad quality predictions at all.

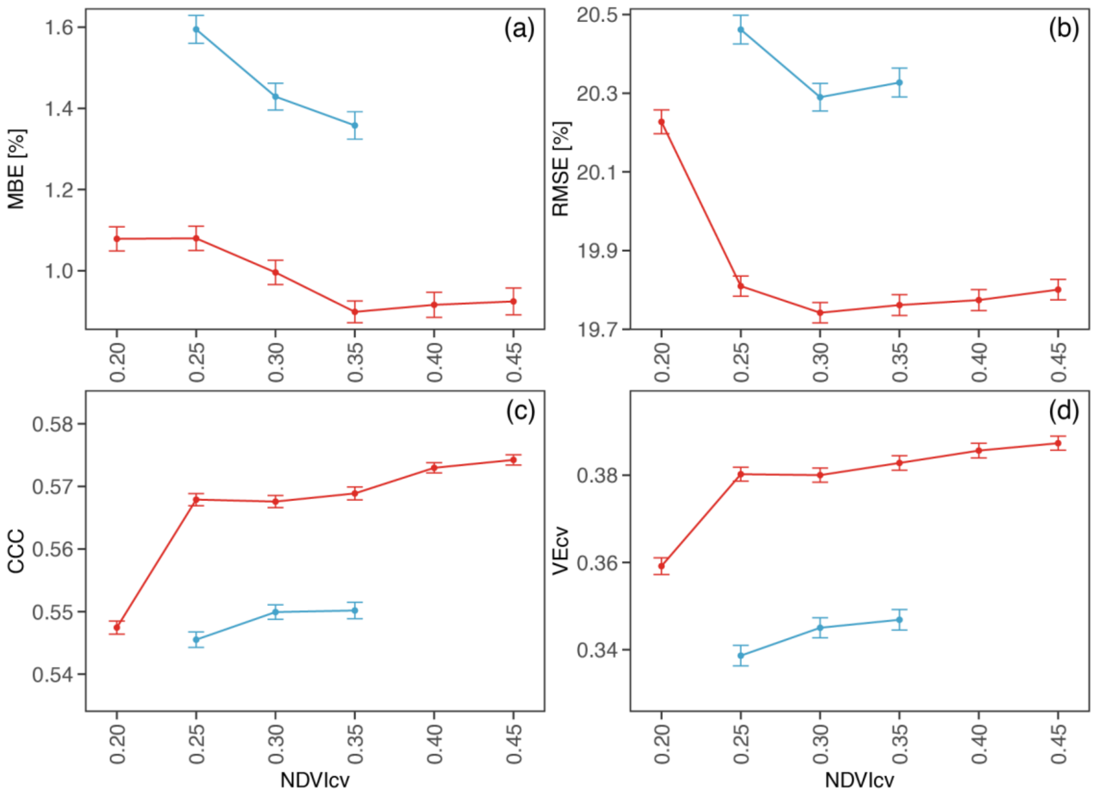

Figure B4. Performance metrics profiles from the general model performance assessment (MP) alternative with the selected variables and the NDVI$_{CV}$ filter applied to the entire dataset (blue line) and only to the training partition (red line). Dots and vertical segments represent the average value and ±1 standard error obtained from the 100 nested LLOCV repetitions, respectively.

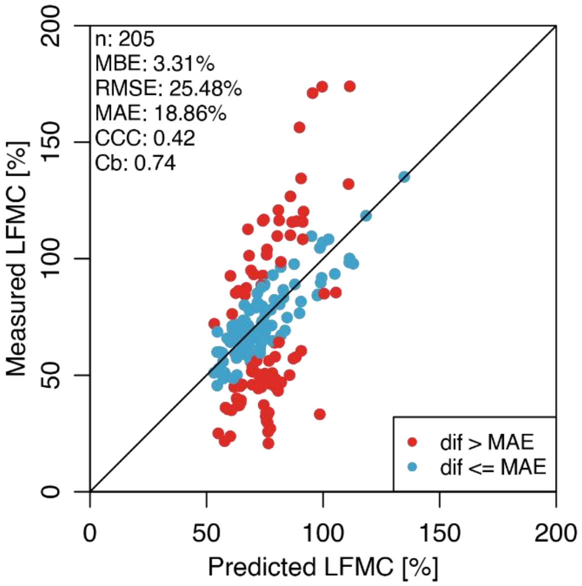

Figure B5. LFMC field observations versus predictions from the CAL validation theoretically rejected by the 0.3 NDVI$_{CV}$ threshold.

B4. Additional Prediction Analysis

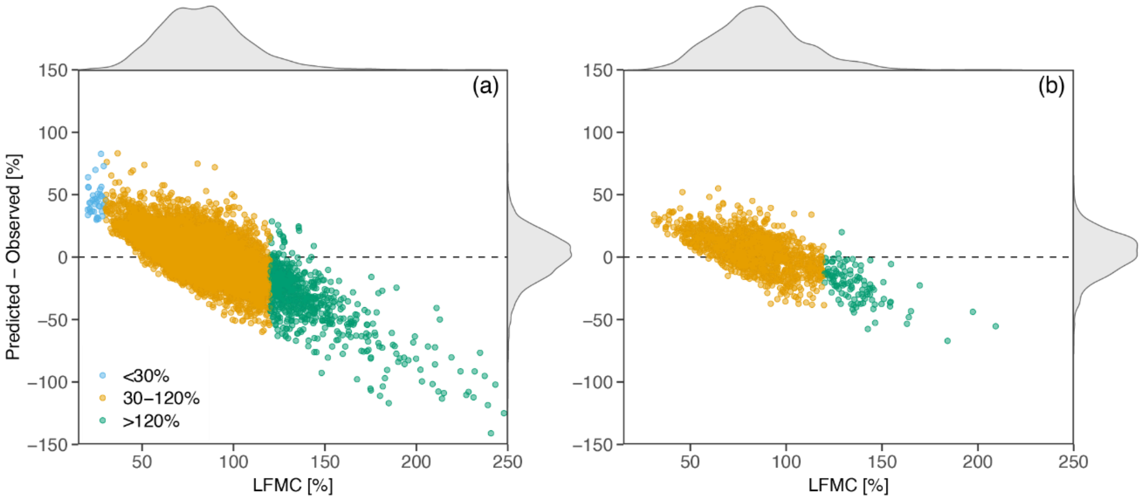

Figure B6. Residuals between predictions and observations against the LFMC observations and their marginal density distributions for CAL (a) and EXT (b). Point colors highlights LFMC observations below, within and above the critical interval for live fuel flammability.