# SMOS-based algorithm to predict potential fire propagation in Europe

Afonso Castela
Instituto Superior Técnico
University of Lisbon
Lisbon, Portugal
afonsocastela@tecnico.ulisboa.pt

***Abstract*— Climate change raises global temperatures, as well as the duration of droughts in some regions of the globe. These shifts may increase the propagation and frequency of wildfires. In this context, Earth Observation (EO) satellites permit to track conditions related to fire risk on a regular basis and with global coverage. In this study, the focus is put on the water content of soils, a crucial factor on forest fires propagation. Here, soil moisture data from the Soil Moisture and Ocean Salinity (SMOS) mission –as well as information on surface temperature, land cover, ecoregions and the time of fires burning- are applied to estimate the potential risk of fire propagation in Europe. The main objectives are: (i) to study the relationship between soil moisture (SM), land surface temperature (LST) and the wildfires in the European Union (EU) for the period 2010-2018, and (ii) to develop a model to predict the maximum fire spread of wildfires in the EU.**

**Different models are proposed and tested. Firstly, regressions between (i) burned area and soil moisture and (ii) burned area and surface temperature, are computed. Secondly, SM and LST are combined in a bilinear model. Thirdly, a more complex model is built including interactions of moisture and temperature with land cover, European ecoregions and the month of the fire outbreak.**

***Keywords-*** Wildfires; Soil moisture; Land surface temperature; SMOS; Remote sensing; Maximum spread.

## I. INTRODUCTION

Human activity increases greenhouse gas (GHGs) emissions, consequently raising global temperatures and the duration of droughts in some regions (Oppenheimer et al. 2015). This leads to more dangerous conditions in terms of the risk of ignition and propagation of wildfires and, in turn, these changes on fire patterns could increment gas emissions reinforcing global warming (e.g., the $CO_2$ emitted from these events can exceed 50% of fossil fuel combustion emissions; Jolly et al., 2015).

Moreover, human factors such as roads and population density play an important role in fire ignition. For instance, in the Mediterranean region it is estimated that 95% of wildfires are human-caused. This high percentage is also observed in other regions, for example in southern Asia (90%), South America (85%) and north eastern Asia (80%; FAO, 2006). Also, recent data analyses from the European Forest Fires Information System (EFFIS) show that approximately 95% of fires in Europe are human-induced (San-Miguel-Ayanz et al., 2012). Although in some areas natural fires are necessary to maintain the ecosystem's dynamics, most wildfires burning directly or indirectly due to human actions cause huge damage in environmental and economic terms, affecting millions of hectares of vegetation and threatening human lives (FAO, 2006).

In this context, it is essential to study the prior conditions igniting wildfires and favouring their propagation. It is essential to take into account that dry weather (i.e., high temperatures, low precipitation amount, and dry soil and vegetation conditions), as well as wind speed, influence fire spread (Jolly et al. 2015).

Earth Observation (EO) satellites permit to track such conditions worldwide and on a regular basis. In the case of this paper, the focus is put on the water content of soils, as well as on surface temperature, in order to determine drought conditions posing risk of fire propagation. To that purpose, passive microwave EO satellites have proved to be very helpful as their measurements are sensitive to water content in soils. Specifically, the first space mission committed to soil moisture monitoring has L-band microwave sensors on board. This is the Soil Moisture and Ocean Salinity (SMOS, launched by the European Space Agency in November 2009).

This study will focus on the applicability of SMOS data, complementarily to other variables, as the basis to estimate the potential extension of wildfires in Europe.

## II. MATERIALS AND METHODS

### A. Data

The present study was carried out on European Union during the period June 2010 to December 2018. The set of variables included fire burned area, soil moisture and land surface temperature conditions prior to fire occurrence, land cover, ecoregions, and the month of the fire outbreak.

The European Space Agency's (ESA) Soil Moisture and Ocean Salinity (SMOS) provides global surface moisture with an estimation from the soil top 5 cm, with a target accuracy of 0.04 $m^3/m^3$, and a spatial resolution of ~40 km Kerr et al. (2010). The product used in this study is the L4 with a resolution of 1 km (BEC, 2018).

Land surface temperature (LST) at midday is obtained from the ERA5 dataset of climatic reanalysed data (Albergel et al., 2018; ERA5-LST) provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). Its resolution is approximately of 31km. In order to match the resolution of the SM dataset, the LST has been linearly interpolated to the SMOS L4 1km grid.

Wind speed (W) is obtained from the system UERRA-HARMONIE (UERRA. (2019)), it is provided in different height levels, and the 15 m height data have been used in this study. Its resolution is approximately of 11 km. It is worth noting that wind information was included as a complementary variable

in order to explore its relationship with fire propagation, not being a factor included in the models.

The land cover map used was obtained from the CORINE (Bossard et al., 2000; LAND COVER. (2019), Figure 1) and allows classifying the fires depending on the type of vegetation burned. The CORINE land cover has a resolution of 250m and provides 44 categories of land cover classified in 5 groups: artificial surfaces, agricultural areas, forest and semi-natural areas, wetlands, and water bodies.

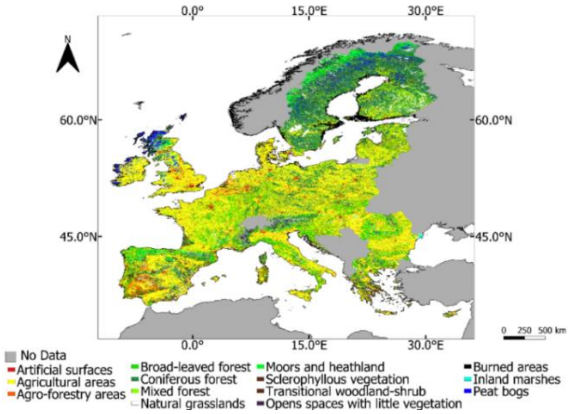


Fig. 1. Land cover in the European Union, derived from the CORINE land cover map (LAND COVER.).

Ecoregions (ER, Metzger, 2005) have been used to classify fire locations in relatively homogeneous regions in terms of their climate, soil, and vegetation. This dataset divides Europe in 12 regions: Alpine North, Boreal, Nemoral, Atlantic North, Alpine South, Continental, Atlantic Central, Pannonian-Pontic, Lusitanian, Mediterranean Mountains, Mediterranean North and Mediterranean South (Fig. 2.).

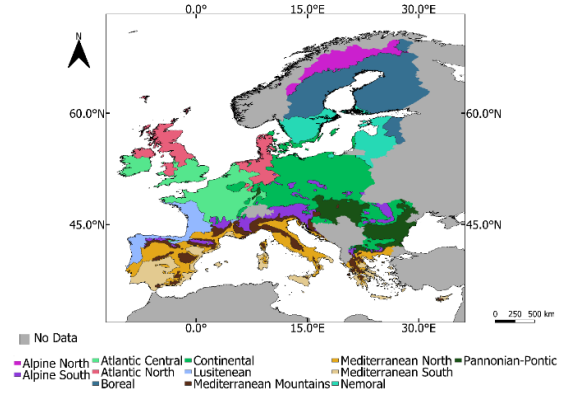


Fig. 2. Ecoregions in the European Union.

Fires information is obtained from the European Forest Fires Information System (EFFIS; San-Miguel-Ayanz et al., 2012) database, which registers fires larger than 10 ha occurring in the EU. This dataset includes the date of fire ignition, the burned area, and the location of the wildfire, as well as its georeferenced perimeter. The database registered 7,569 fires larger than 10 ha in the study area between June 2010 and December 2018 (Figure 3), although the final dataset is only composed by 7496 fires (see Section 2-C). Most fires burned less than 1,000 ha, although near 500 fires affected larger extensions (see Table I).

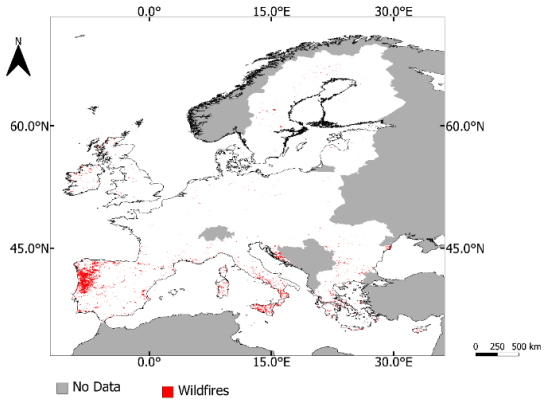


Fig. 3. Wildfires in Europe from 2010 to 2018 larger than 10ha.

TABLE I SAMPLE OF FIRES

| Burned areas (ha) | Number of fires | Percentage (%) |
|---|---|---|
| 10-100 | 3622 | 48,33 |
| 100-1,000 | 3387 | 45,19 |
| 1,000-10,000 | 462 | 6,16 |
| >10,000 | 24 | 0,32 |
| Total | 7496 | 100 |

### *B. Study area*

The study area is the European Union (73.33ºN, 31.36ºN; -12.29ºW, 35.04ºW; excluding Madeira, Azores and Canary Islands) and it spans through more than 10 million km². As Europe is a vast continent, it presents contrasting climate patterns which determine the vegetation types and abundance, and consequently the fire ignition and spread. Figure 5 shows different temperature and precipitation regimes in Europe, with cold northern and Atlantic zones, continental regions with marked seasonal contrasts, and the hot and dry Mediterranean basin. Typically, in the north, as it is a subarctic region, the climate is defined by long and very cold winters, and short, cool to mild summers (Metzger et al.; 2012). The European mountainous regions, the inner continental lands, and the Pannonian-Pontic zones, present a continental climate, with a significant annual variation in temperature: hot summers contrast with very cold winters. Concerning to the Atlantic Central zone, it includes the south west and south of the British Islands, Northern France and the Netherlands: the climate is typically oceanic in the west, with cool wet winters and mild moist summers, but becomes more continental in the east (Metzger et al.; 2012). In the southern part of Europe there are two main regions: the Lusitanian and the Mediterranean. The Lusitanian zone includes west and south west France, northern Spain, and most of Portugal. The climate is warm and wet, with precipitation concentrated in the winter months. The Mediterranean zones present warm and dry summers and precipitation concentrated in the autumn months (Metzger et al.; 2012).

The wet climate of the Lusitanian region leads to a high amount of fuel, which in warm conditions (especially in summer, in southern countries such as Portugal) brings to

dangerous prone to fire conditions: more than one third of fires (37.66%) burn in this region. Also, the warm temperatures in the Mediterranean regions lead to almost 40% of fires burning in these areas (specifically, 19.58% burn in the Mediterranean South region, and 19.63% in the Mediterranean North region). In other areas, wildfires are less frequent such as in the case of Alpine North, Boreal and Nemoral regions, which are colder. However, exceptional dry and warm conditions have led to important burning episodes in countries such as Sweden or the United Kingdom in summer 2018 (see Table III).

Figures 4 and 5 present the prior conditions to the fire occurrences in terms of soil moisture and land surface temperature in relation to the ecoregion and land cover. The land covers transitional woodland-shrub, sclerophyllous vegetation, broad-leaved forest and agricultural areas presented the lowest values of soil moisture, while peat bogs, coniferous forest and inland marshes, the highest humidity values. Regarding regions such as Alpine south, Lusitanian, and Mediterranean zones registered the lower moisture values and also the warmest temperatures. There is a link between the wildfires in the regions with lower values of soil moisture and higher temperatures.

Figures 4 and 5 are coherent with Tables II and III, showing that in regions where most wildfires occurred the temperatures were high and the soils were dry (i.e., in Lusitanian and Mediterranean regions). Concerning to land cover, transitional woodland-shrub and agricultural areas presented similar relationships.

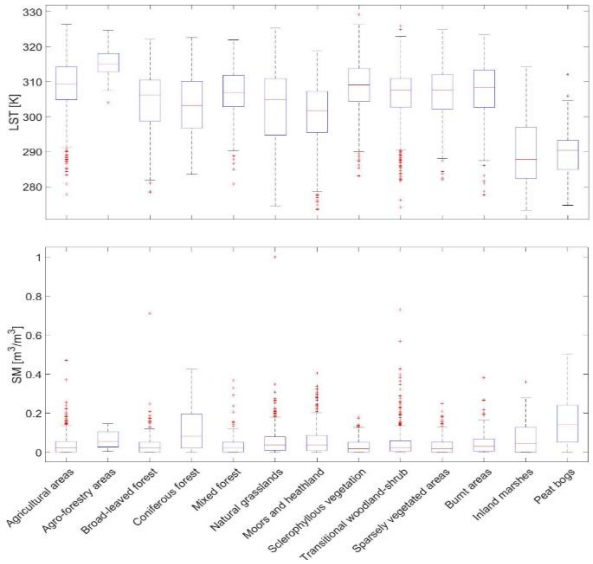


Fig. 4. Relationship between temperature (top) and moisture (bottom) with land covers where fires burned.

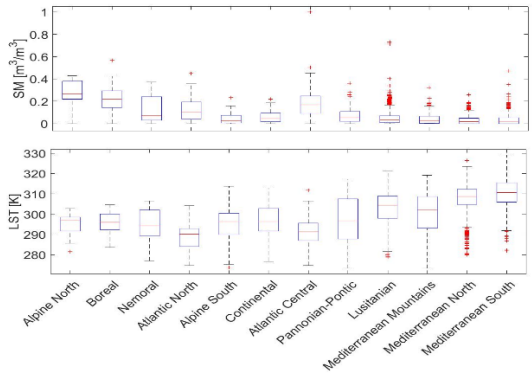


Fig. 5Relationship between temperature (top) and moisture (bottom) with ecoregions where fires burned.

TABLE II Wildfires classified by ecoregion.

| Ecoregion Classification | Number of wildfires | Percentage (%) |
|---|---|---|
| Alpine North | 17 | 0,23 |
| Boreal | 89 | 1,19 |
| Nemoral | 25 | 0,33 |
| Atlantic North | 203 | 2,71 |
| Alpine South | 211 | 2,82 |
| Continental | 111 | 1,48 |
| Atlantic Central | 112 | 1,49 |
| Pannonian-Pontic | 208 | 2,78 |
| Lusitanian | 2579 | 34,41 |
| Mediterranean Mountains | 528 | 7,04 |
| Mediterranean North | 1684 | 22,47 |
| Mediterranean South | 1728 | 23,06 |

TABLE III Wildfires classified by land cover.

| Land cover classification | Number of wildfires | Percentage (%) |
|---|---|---|
| Agricultural areas | 1223 | 16,32 |
| Agro-forestry areas | 22 | 0,29 |
| Broad-leaved forest | 467 | 6,23 |
| Coniferous forest | 218 | 2,91 |
| Mixed forest | 163 | 2,17 |
| Natural grasslands | 729 | 9,73 |
| Moors and heathland | 1721 | 22,96 |
| Sclerophyllous vegetation | 608 | 8,11 |
| Transitional woodland-shrub | 1356 | 18,09 |
| Sparsely vegetated areas | 513 | 6,84 |
| Burnt areas | 218 | 2,91 |
| Inland marshes | 97 | 1,29 |
| Peat bogs | 161 | 2,15 |

### *C. Data Preparation*

In order to build the database used in this study, firstly the 7569 wildfires have been classified in terms of land cover, ecoregion, soil moisture and land surface temperature. To that purpose, the geographic positions of the burned pixels have been linked to all data maps using the minimum distance between the centres of the pixels. For each fire, the ER and LC modal class among the affected pixels has been computed. For quantitative

variables (LST and SM) the mean has been computed. Note that SM and LST data from the day of fire occurrence to three days before have been included in the database.

After that they were filtered under certain conditions, so from the initial data of 7569 wildfires, to the final dataset composed by 7496 fires, as said before, fires in the islands and fires before July 2010 were eliminated; then some land covers were not considered (artificial surfaces, maritime wetlands water bodies and open spaces with little or no vegetation), because this study will focus only in forest wildfires so 23 wildfires were removed from the dataset.

## III. LINEAR MODEL TO PREDICT POTENTIAL EXTENSION OF WILDFIRES

### A. *Development of a model to predict the potential extension of wildfires*

Since the goal of this work is to develop a model to predict the potential extension of wildfires, in this section several models will be presented and compared. The different models tested are ordered in this section from the simplest to the most complex. These models are essentially based on the SMOS-derived soil moisture and on the ERA5 land surface temperature, but also explore the influence of the land cover, the ecoregions, and the month of fire burnings.

However, as a previous step to the model development, the availability of data for operational purposes must be considered. In that sense, data of LST for the day of fire ignition at noon cannot be used to produce operational propagation risk maps for obvious reasons: most fires burn during morning and at noon, so risk maps could not reach the fire prevention services on time. Also, it must be considered that L4 SM products from ascending passes are not available the same day of data acquisition. Additionally, the possibility of acquiring LST forecasts for the study day is not affordable on economic terms (as such data is delivered after payment to the Spanish meteorological agency, AEMET, on behalf of ECMWF). For these reasons, model calibration has been based on the most recent SM and LST data from the day previous to fire occurrences to three days before. To evaluate the differences between the fire date and the day before fire burning, LST and SM have been compared for both cases. Results are shown in Figure 6 and demonstrate that mean differences peak is approximately at zero. Hence, similarities between both days suggest that moisture and temperature data for days before fires can be used for model calibration.

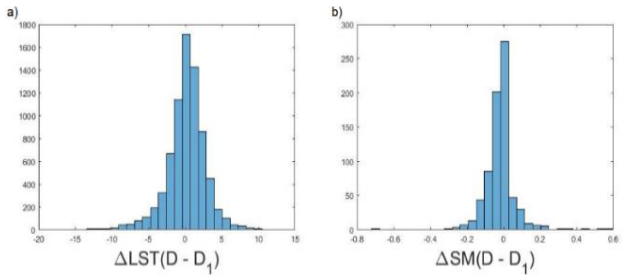


Fig. 6. Differences between the fire date and one day before fires for (a) land surface temperature and (b) soil moisture.

### B. *Preliminary analysis of SM and LST relationship with wildfires propagation*

To build the model 70% of the registered fires (training dataset) have been used. The remaining 30% has been used for validation purposes (validation dataset).

The distribution in these two groups have been done with a random sample. Soil moisture and temperature of the training set are plotted against burned area, which has been logarithmically transformed (decimal logarithm), following Chaparro et al. (2016a) and Chaparro (2018). Figure 7 shows the resulting triangle shaped plots, where both moisture and temperature limit the fire extent.

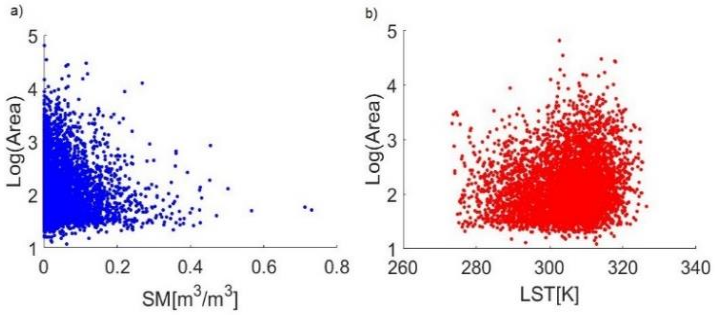


Fig. 7. Decimal logarithm of burned area (originally in hectares) for the training dataset is plotted against (a) soil moisture and (b) land surface temperature.

From the previous plots, an "imaginary hypotenuse" of each triangle delimits the maximum burned area at different values of SM or LST. Hence, higher temperatures lead to larger burned areas, and drier soils are linked, also, with larger wildfires. Obviously, multiple factors (e.g., extinction capacity, lack of fuel, changing meteorological conditions…) lead in most cases to smaller fires than which could be expected from the SM and LST conditions.

Hence, the approximation based on triangular-shaped relationships between moisture, temperature, and burned area pursues to predict the maximum or potential propagation of fires at particular SM and/or LST conditions, using a linear regression to approximate the hypotenuse of the triangle.

To this purpose, the variables SM and LST have been binned, and the largest fires for each bin have been chosen in order to fit these linear regressions. This method has been used previously in Chaparro et al. (2016a) for the Iberian Peninsula. Different modelling approaches have been tested and are described hereafter:

### C. *SM-burned area and LST-burned area models*

Different SM and LST bins have been tested: from 0.01 $m^3/m^3$ to 0.05 $m^3/m^3$, and from 1K to 4 K, respectively. Firstly, a linear regression to estimate Log(Area) has been performed as a function of SM (1):

$$Log_{10}(Area) = \alpha_{SM} + \beta_{SM} \cdot SM, \qquad (1)$$

where $Log_{10}(Area)$ corresponds to the maximum of the decimal logarithm of burned area for each SM bin, $\alpha_{SM}$ is the intercept, SM is soil moisture, and its coefficient is $\beta_{SM}$.

The same approach has been used for LST:

$$Log_{10}(Area) = \alpha_{LST} + \beta_{LST} \cdot LST\,, \qquad (2)$$

where $Log_{10}(Area)$ corresponds to the maximum of the decimal logarithm of burned area for each LST bin , $\alpha_{LST}$ is the intercept, LST is the land surface temperature, and its coefficient is $\beta_{LST}$. Samples for models (1) and (2) are 47 and 48 fires, respectively, in the case of the binnings chosen (see Section 4-A).

After such models have been tested, an important number of wildfires exceed the potential propagation predicted (see Section 4-A).

To solve this underestimation, an additive factor (equivalent to one standard deviation of LogArea) was added to the intercept of both equations. Then, (1) and (2) changed to:

$$Log_{10}(Area) = \alpha'_{SM} + \beta_{SM} \cdot SM \qquad (3)$$

$$Log_{10}(Area) = \alpha'_{LST} + \beta_{LST} \cdot LST\,, \qquad (4)$$

where:

$$\alpha'_{SM} = \alpha_{SM} + \sigma_{\log(area)} \text{ and}$$

$$\alpha'_{LST} = \alpha_{LST} + \sigma_{\log(area)'}.$$

These regressions are parallel to (1) and (2) (i.e., same slope), but represent more reliably the maximum extension of fires at each SM or LST value.

### *D. Models combining SM and LST*

A second step consisted on applying both variables in the same model.

The same procedure (i.e., binning the variables) has been applied. However, now the variables have been binned together: soil moisture bins are 0.01 $m^3/m^3$, 0.02 $m^3/m^3$and 0.03 $m^3/m^3$ and LST bins are 1K, 2K and 3K. All combinations of SM-LST bins have been tested. In this case, the equation linking SM and LST to the burned area is:

$$Log_{10}(Area) = \alpha + \beta \cdot SM + \varphi \cdot LST, \qquad (5)$$

where $Log_{10}(Area)$ corresponds to the largest (over 90th percentile) decimal logarithm of burned area per each SM-LST bin, $\alpha$ is the intercept, $SM$ is the soil moisture, $LST$ is the land surface temperature, and their coefficients are $\beta$ and $\varphi$ , respectively.

The sample for this equation using the binning which provides the best fit (see Section 4-B) is 639.

As an attempt to improve (5), the model has been finally complemented with auxiliary variables: land cover, ecoregions, and the month when the wildfire burned.

Only those variables and those interactions which are significant at 95% confidence level ($p<0.05$) have been chosen and included in the model defined by (6):

$$Log_{10}(Area) = \alpha + \beta_{(M,LC)} \cdot SM + \varphi_{(ER,LC)} \cdot LST + M + LC + ER, \qquad (6)$$

where $Log_{10}(Area)$ corresponds to the largest (over 90th percentile) decimal logarithm of burned area per each SM-LST bin , $\alpha$ is the intercept, $SM$ is the soil moisture, $LST$ is the land surface temperature, $M$ refers to the moth of the wildfire, $LC$ refers to the land cover classification, $ER$ refers to the ecoregion and the coefficients are: $\beta_{(M,LC)}$ for SM (which changes depending on the interaction with month and land cover) and $\varphi_{(ER,LC)}$ for LST (which changes depending on the interaction with ecoregion and land cover).

In this model, the classification of ecoregions has been changed from that in Section 2-A in order to fit the best model and to reach enough sample in each area. The best reclassification is found when grouping the original ecoregions in three large areas as follows:

1. North: Alpine North, Boreal and Nemoral.
2. Central: Atlantic North, Alpine South, Continental and Atlantic Central.
3. South: Alpine South, Pannonian-Pontic, Lusitanian, Mediterranean Mountains, Mediterranean North and Mediterranean South.

In addition, agricultural areas and agro-forestry areas have been grouped also in a single category to ensure enough sample in this group.

### *E. Relationship between wind and burned area*

Later it was thought that to the model described by (6) it could be added the wind speed.

So, the influence of the wind in the wildfire propagation has been also studied since it is an important factor which usually increases the fire spread, turning extinction tasks difficult.

This task has been performed with exploratory purposes using part of the dataset (specifically fires between July 2010 and December 2013).

### *F. Validation of the models and development of potential fire propagation maps*

The models applied have been validated using the validation dataset. To this purpose, the burned area of each wildfire has been estimated using the different models, and compared to the real value of fire extent.

Validations for the different models have been compared in order to choose the best model. Then, this model has been used to provide European maps of potential fire propagation. The best solution is presented at results section.

## IV. RESULTS

### *A. SM-burned area and LST-burned area models*

Fitting for models in (1) and (2) changes depending on the bin width chosen.

After analysing results with several binnings, it has been decided to select the smallest bin since bigger bins excessively reduce the sample ($n<20$). So, for (1) the binning chosen is $\Delta SM = 0.01$ $m^3/m^3$ and for (2) it is $\Delta LST = 1$ K. Table 6 shows goodness of fit and sample for these models.

TABLE IV Bin widths for SM and LST.

| | Δ (bin width) | $R^2$ | Sample |
|---|---|---|---|
| Δ SM [$m^3/m^3$] | 0,01 | 0,73 | 47 |
| Δ LST [K] | 1 | 0,40 | 48 |

The regression between the logarithm of burned area and soil moisture is presented in Figure 8, which represents the models described by (1) and (3).

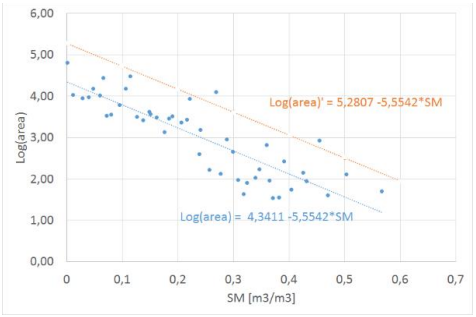


Fig. 8. Regressions of LogArea as a function of SM for (1) (blue) and (3) (orange).

The relationship between the logarithm of burned area and the land surface temperature is presented in Figure 9, both for (2) and for (4).

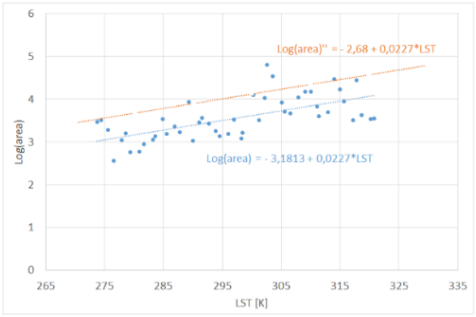


Fig. 9. Regressions of LogArea as a function of LST, both for (2) (blue) and (4) (orange).

### *B. Models combining SM and LST*

The linear regression proposed by (5) applies SM and LST as explanatory variables. The best model in terms of $R^2$ has been chosen from the different binning possibilities reported in Table 7.

TABLE V $R^2$ values for different bins of SM and LST in (5). Bold font highlights the binning finally chosen for the SM-LST model.

| ΔLST [K] / ΔSM [$m^3/m^3$] | 1 | 2 | 3 |
|---|---|---|---|
| 0,01 | 0,37 | 0,45 | **0,48** |
| 0,02 | 0,45 | 0,47 | 0,45 |
| 0,03 | 0,47 | 0,46 | 0,43 |

The best regression has been obtained for the combination ΔLST = 3K and ΔSM= 0.01 $m^3/m^3$, with a $R^2$ = 0,48 ($p<0.0001$). LST and SM range from 274 to 321 K and from 0 to 0.57 $m^3/m^3$, respectively.

The relationship between these variables and LogArea is shown in Figure 10, where the largest fires are generally found in the driest and warmest soils.

For instance, most of the bins with maximum spread ⩾1,000 ha (i.e., $Log_{10}(Area)3$) present SM⩽0.10 $m^3$ /$m^3$ and LST⩾ 300 K. Finally, coefficients α , β and γ for the (5) are -3,331, -2,708 and 0,0193, respectively.

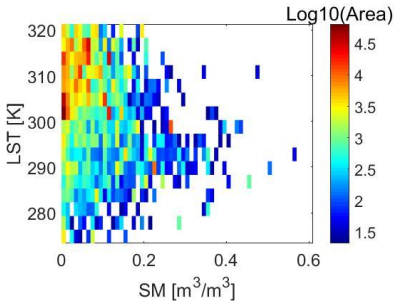


Fig. 10. Relationship among SM, LST and the percentile 90 of the decimal logarithm of burned area (colorbar) for each SM-LST bin. Binning: ΔLST = 3K and ΔSM= 0.01 $m^3/m^3$. This binning led to the model in (5), with $R^2$=0.48.

Finally, the last model proposed (6) includes: SM, LST, the month of fire occurrence, the land cover, the ecoregion, and the interactions SM × month, SM × land cover, LST × land cover, and LST × ecoregion.

The adjusted $R^2$ of the model is 0,57. The variance explained by this model is reported in Table 8, showing that SM was the variable explaining most of the variance (25%), followed by LST (24%).

TABLE VI Variance explained by each variable in (6). The variance explained by the model is 0,61 ($R^2$=0.61). Nevertheless, the adjusted $R^2$ (0,57) is used hereafter as it is unaffected by possible redundancy of information.

| Variable | Variance model (%) | Variance total (%) |
|---|---|---|
| SM | 41,01 | 25,08 |
| LST | 38,83 | 23,74 |
| LC | 5,24 | 3,21 |
| SM×M | 4,21 | 2,58 |
| SM×LC | 4,12 | 2,52 |
| M | 3,78 | 2,31 |
| LST×LC | 1,62 | 0,99 |
| LST×ER | 1,15 | 0,70 |
| ER | 0,03 | 0,02 |
| Total | 100,00 | 61,15 |

Figures 11 and 12 show the prediction of burned area as a function of land surface temperature and soil moisture. The

potential propagations predicted report how SM and LST limit the maximum burned area up to a maximum of ~3100 ha

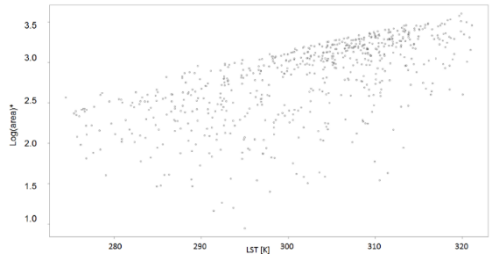


Fig. 11. Prediction of the decimal logarithm of burned area (originally in hectares) for the model described by (6), plotted against land surface temperature. Note: Log(area)* corresponds to the predicted area.

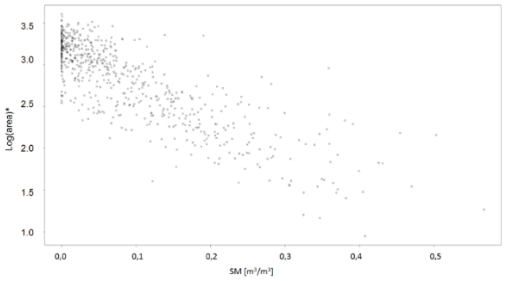


Fig. 12. Prediction of the decimal logarithm of burned area (originally in hectares) for the model described by (6), plotted against soil moisture. Note: Log(area)* corresponds to the predicted area.

Figures 13, 14 15 and 16 show the interactions between the variables. The interactions between the temperature and the different regions are shown in Figure 13, where the slope is larger in the northern region showing larger impact of temperature in this area. Concerning LST and SM interactions with land cover (Figure 14 and 15), most effects of these variables are consistent with the overall pattern (i.e., higher risk in warmer and drier conditions). Nevertheless, some exceptions are found: (i) mixed forests present an opposite effect (although with wide confidence intervals probably due to low sample); (ii) an almost flat trend is found in humid land covers (inland marshes and peat bogs); and (iii) broadleaved forests do not show a clear trend for SM (Figures 15 and 16). Concerning to the SM × months interaction, a coherent pattern is found for most months (Figure 16).

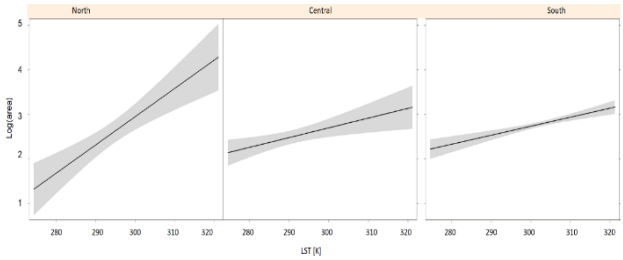


Fig. 13. Interaction between land surface temperature and ecoregions, plotted as function of the decimal logarithm of burned area (originally in hectares) from the model described by (6).

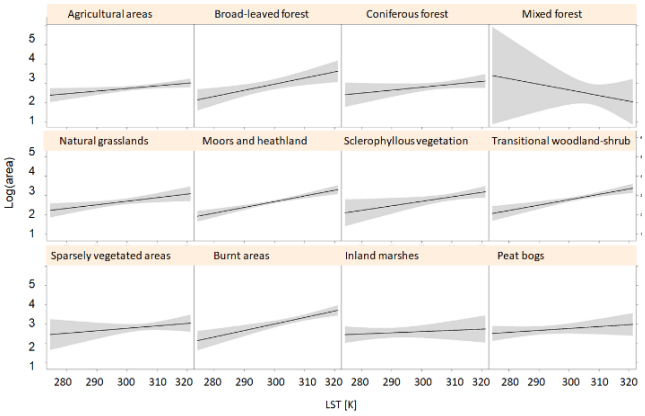


Fig. 14. Interaction between land surface temperature and land cover, plotted as function of the decimal logarithm of burned area (originally in hectares) from the model described by (6).

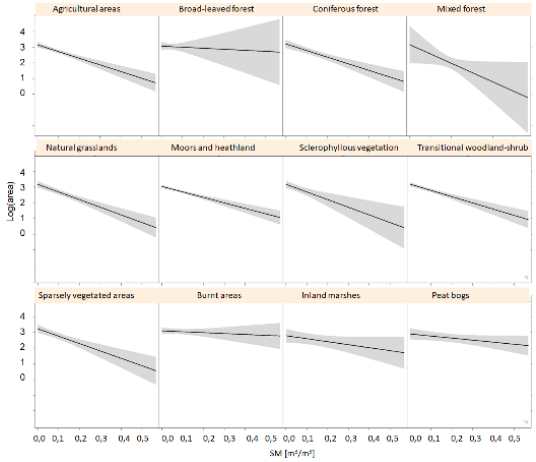


Fig. 15. Interaction between soil moisture and land cover, plotted as function of the decimal logarithm of burned area (originally in hectares) from the model described by (6).

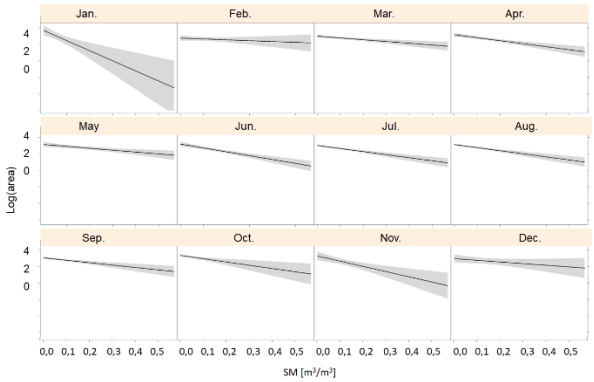


Fig. 16. Interaction between soil moisture and month, plotted as function of the decimal logarithm of burned area (originally in hectares) from the model described by (6).

## *C. Summary of all models*

All the models have been validated using the validation dataset. Results are shown in Table VI. Equation (3) presents the

lowest errors and will be applied for mapping potential fire propagation.

TABLE VII Comparison between all models. *Note: Models (3) and (4) are similar to (1) and (2), respectively (same slope, but different intercept). Bold shows the model which will be finally applied.

| Equation | $R^2$ | Correct validation | Correct validation (%) | Max. excess (max. error) [ha] | Fires >10k ha [correctly classified] (%) | Fires >10k ha [incorrectly classified] (%) |
|---|---|---|---|---|---|---|
| (1) – SM | 0,73 | 1890 | 99,21% | 27917,60 | 16,67 | 83,33 |
| (2)- LST | 0,40 | 1885 | 90% | 29773,69 | 0,00 | 100,00 |
| **(3) – SM*** | ***** | **1904** | **99,94%** | **215,15** | **100,00** | **0,00** |
| (4) – LST* | * | 1902 | 99,84% | 18959,78 | 66,67 | 33,33 |
| (5) – SM & LST | 0,48 | 1750 | 91,87 | 33520,91 | 0,00 | 100,00 |
| (6)– SM & LST & interactions | 0,57 | 1727 | 90,65 | 33655,11 | 0,00 | 100,00 |

### D. *Maps of predicted potential burned area*

Results show how the suggested methodology provides good approximation to the maximum expected fire spread. In that sense, it is possible to obtain maps of predicted potential burned area based on the best model (3). Figure 17 presents an example of these maps with the predicted maximum potential area being classified in different risk categories (Table VII).

TABLE VIII Classification of the predicted maximum potential burned area in the risk categories of Figure 17.

| Predicted area (ha) | <100 | 100 - 1000 | 1000-10.000 | 10.000-50.000 | >50.000 |
|---|---|---|---|---|---|
| Risk category | Low | Moderate | High | Very high | Extreme |

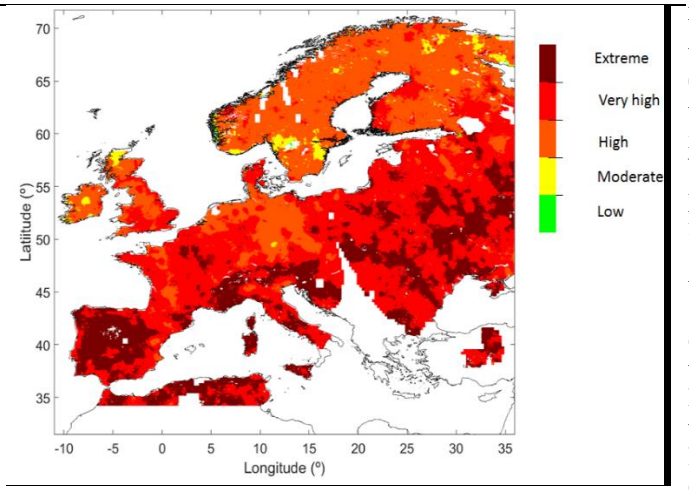


Fig. 17. Example of a maximum potential area map (20th August 2015) derived from (3).

### E. *Wind influence*

In this study the wind speed absolute value (no direction) has been considered. It has been found scarce relationship between the wind speed absolute value and the logarithm of burned area ($R^2$ = 0,03). This is further discussed in Section 5.

## V. Discussion

### A. *Role of soil moisture and land surface temperature in fire propagation*

SMOS-derived soil moisture has been found to be an important explanatory variable in evaluating fire propagation in the European Union. Results show that wet soils strongly limit burned area and, in an opposite way, drier soils increase the potential risk of large forest fires. Concerning surface temperature, the largest wildfires happen under high temperatures, while propagation risk decreases in colder situations. These results are consistent with similar studies in the Iberian Peninsula (Chaparro et al., 2016 a, and b) and also in other regions of the world (Bartsch et al., 2009). There is a strong link between large burned areas, high temperatures and low values of soil moisture, and a large percentage of fires larger than >1000 ha have burned under these conditions (Figure 10). This results allow determining a maximum or potential burned area as a function of SM and/or LST, either used separately in specific models, or combined in a single regression. This approach has explained between 40% and 73% of the potential burned area (Table 7). This is coherent with previous models developed in the Iberian Peninsula (Chaparro et al., 2016a).

The model based on soil moisture (3) has shown the best results once an additive factor (σLogArea) has been included ($R^2$ = 0.73; maximum error = 215 ha). Similarly, soil moisture has shown also important capacity to explain fire propagation (25% of explained variance; (6)) when combined with LST. Concerning to LST, the variance explained by this variable is lower (40% in (2); 24% when combined with SM in (6)). These effects are consistent with the research performed for the Iberian Peninsula by Chaparro et al. (2016 a, and b), where SM explained 33% and LST 20% of the modelled variance. However, the present results surprisingly report an improved performance of the simplest models (based on SM) if compared to the model combining SM and LST. In this case, we suggest that the scarce correlation found between SM and LST (r = -0.31) may explain that including both variables could add noise to our model. In particular, the relationship found in Figure 10 is not perfect and may be decoupled because, although large fires burn in high temperatures and dry soils, both LST and SM preceding these fires show important variability (approximately from 300 to 320 K and from 0 to 0.10 $m^3/m^3$ in the largest fires).

### B. *Land cover, ecoregions, and month of fire burning*

Including complementary variables, such as land cover, ecoregions and month, and their interactions; model (6); has led to a poor improvement ($R^2$ = 0,57) with respect to the basic SM-LST model ((5); $R^2$ = 0,48). There might be some reasons for that. Firstly, the application of the method used for the Iberian Peninsula in Chaparro et al. (2016a) might not be similar for the entire European Union (as it is a larger study area with many different contrasting patterns in terms of climate and vegetation). Secondly, the significance and applicability of the proposed variables may differ among regions of Europe and, importantly, maybe other variables than those tested in the Iberian Peninsula should be tested now (e.g., synoptic situations, fuel amount, vegetation indices, etc…).

In that sense, the influence of the ecoregions is almost negligible when compared to the other variables (<1%), possibly due to the fact that SM and LST do capture most climate variability through the temperature and moisture gradients in the continent (see Metzger et al., 2005), leading ecoregions to be a redundant source of information.

In addition, the effect of land covers on fire propagation has shown some different patterns if compared to the land cover – burned area relationships found in Chaparro et al. (2016a), but most of their effects and interactions show coherent patterns.

For instance, in northern regions the role of LST is more marked than in southern and warmer areas (Figure 13), similarly to what is reported in Chaparro et al. (2016a) for the Iberian Peninsula. Also, as expected, larger wildfires are predicted in the south of Europe rather in the north and central regions, which is coherent with the main climate patterns and with number of fires per region.

Regarding the months, the summer months (July, August), but some autumn months as well (September and October) present higher predicted burned areas.

### C. *Validation and application of a model to predict potential burned area*

The model described by (3), which applies SM + σLogArea has been selected for operational applications. The validation shows an agreement of 99.94% with a maximum error of 215,15ha, and reports that no fires larger of 10,000 ha have been underestimated. It is important to note that this high accuracy is reached due to the addition of the σLogArea factor, which avoids underestimation of most fires. Although this could also lead to an overestimation, the intercept of the model ($\alpha$ = 5,28) shows that the maximum fire propagation would be of ~100.000 ha. This is consistent with: (i) the largest fires detected in the EU during the study period (which have reached ~70.000 ha and have burned in Portugal); (ii) the intercept for the model proposed by Chaparro et al. (2016a), which reaches similar – even higher- values (depending on the interaction of the variables in that model); and (iii) the fact that under the present climate change conditions, a new type of wildfires, called "megafires", are increasingly frequent (e.g., recently >200.000 ha have burned in a single fire in Canada in 2019, and a similar situation happened in 2018 when four fires >100.000 ha burned in the country).

The map of potential burned area (Figure 17) is an example of the first application for fire risk assessment based on SMOS data for Europe. Now, thanks to the availability of SMOS-derived 1 km soil moisture maps over Europe (BEC, 2018), this product can be put on operational applications.

### D. *Future work*

With the purpose of enhancing the present results, the influence of the wind has been studied. However, the relationship between burned area and wind speed is scarce (see Section 4-E) probably due to the fact that the resolution from the wind model applied is too coarse. Instead, wind effects are more complex and involve local factors (e.g., orography) and the change of wind direction. Also, synoptic situations are needed to link wind and meteorological conditions to wildfire regimes (Duane and Brotons (2018).

Finally, other approaches could be studied to improve the reliability of the present model. First, remotely-sensed LST data could be derived from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) sensor on board of Meteosat Second Generation satellites, in order to reduce the dependency on meteorological stations and to the get LST data every 15 minutes. However, note that this approach presents the limitation of cloud-masking in the infrared band. Secondly, the application of soil moisture and temperature anomalies to estimate fire propagation should be studied (Chaparro et al., 2016b). Thirdly, the applicability of microwave-derived L-band vegetation optical depth (VOD) as an estimate of biomass (Brandt et al., 2018; Rodríguez-Fernández et al., 2018; Chaparro et al., 2019) should be used to infer fuel availability, and the application of VOD to study vegetation water content conditions (e.g., Rao et al., 2019), should also be addressed.

## VI. CONCLUSIONS

Pre-fire conditions in the European Union have been analysed for the period 2010–2018 using the SMOS L4 high resolution soil moisture product, the ERA-5 land surface temperature data, CORINE land cover, ecoregions and month of the fire outbreak. Low temperatures and wet soils limited wildfires spread, while dry soils and high temperatures were linked to large burned areas.

Several models have been tested with different approaches. Firstly, estimates on burned area based on SM and LST have been performed using these variables in different linear regressions. Also, an additive term (σLogArea) has been summed to the intercept to avoid underestimation of fire propagation. Secondly, moisture and surface temperature have been joined in a bilinear regression, where SM has explained 25% of maximum fire spread, and LST 24% of the variance. In this case, the largest wildfires, ≥1,000 ha (i.e. $Log_{10}(Area) \geq 3$), are triggered by a moisture level below 0.10 m3 /m3 and temperatures higher than 300K. Thirdly, this model has been extended including land cover, ecoregions, and the month when fires burn. As expected, the largest wildfires take place in the southern region of Europe. Transitional woodland-shrub and burnt areas have been found as the land cover presenting a higher risk of fire spread. It is during the summer months that mores wildfires occur and are also the largest ones.

The comparison among the different models, and their validation, report that the model described in (3) (i.e., burned area as a function of soil moisture plus σLogArea) provides the best result ($R^2$ = 0.73; maximum error = 215 ha). Adding the standard deviation of the logarithm of burned area avoids underestimation and does not lead to overestimation effects, if compared to previous studies and to recent wildfires burning ~100.000 ha. From this model, a new operational application will arise: the SMOS Barcelona Expert Center will take profit of the new European SM maps at 1 km resolution to provide fire propagation risk maps in the European Union, based on the present work. Future work could enhance this model approach (e.g., including satellite-derived LST, moisture and temperature anomalies, VOD data, or wind information and synoptic situations) leading to enhanced estimates.

In conclusion, the proposed approach is valid for estimating the risk of propagation of wildfires in the European Union and has permitted to demonstrate and apply the high potentiality of L-band soil moisture data in fire risk assessment.

## ACKNOWLEDGMENT

Fires data were provided by the European Forest Fire Information System—EFFIS (http://effisjrc.ec.europa.eu) of the European Commission Joint Research Centre.

## REFERENCES

[1] Albergel, C., Dutra, E., Munier, S., Calvet, J., Munoz-Sabater, J., De Rosnay, P., & Balsamo, G. (2018). ERA-5 and ERA-Interim driven ISBA land surface model simulations: Which one performs better? Hydrology and Earth System Sciences, 22(6), 3515-3532.

[2] Aubrecht, C., Elvidge, C. D., Baugh, K., Hahn, S., & Jorge, N. (2011). Identification of wildfire precursor conditions: Linking satellite based fire and soil moisture data.

[3] Bartsch, A., Balzter, H., & George, C. (2009). The influence of regional surface soil moisture anomalies on forest fires in Siberia observed from satellites. Environmental Research Letters, 4(4), 45021.

[4] BEC (Barcelona Expert Centre). (2018). Land Variables. Retrieved from: http://bec.icm.csic.es/land-datasets. (s.f.).

[5] Bossard, M., Feranec, J., Otahel, J., & Steenmans, C. (2000). CORINE land cover technical guide-Addendum 2000.

[6] Brandt, M., Wigneron, J., Chave, J., Tagesson, T., Penuelas, J., Ciais, P., . . . Fensholt, R. (2018). Satellite passive microwaves reveal recent climate-induced carbon losses in African drylands. Nature Ecology and Evolution, 2(5), 827-835.

[7] Chaparro, D., Duveiller, G., Piles, M., Cescatti, A., Vall-llossera, M., Camps, A., Entekhabi, D. (in press). Sensitivity of L-band vegetation optical depth to carbon stocks in tropical forests: comparison to higher frequencies and optical indices. Remote Sensing of Environment.

[8] Chaparro Danon, D. "Applications of L-band missions for environmental research". Tesi doctoral, UPC, Departament de Teoria del Senyal i Comunicacions, 2018

[9] Chaparro, D., Piles, M., Vall-Llossera, M., & Camps, A. (2016). Surface moisture and temperature trends anticipate drought conditions linked to wildfire activity in the Iberian Peninsula. European Journal of Remote Sensing, 49, 955-971.

[10] Chaparro, D., Vall-Llossera, M., Piles, M., Camps, A., Rudiger, C., & Riera-Tatche, R. (2016). Predicting the Extent of Wildfires Using Remotely Sensed Soil Moisture and Temperature Trends. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(6), 2818-2829.

[11] Duane, A., & Brotons, L. (2018). Synoptic weather conditions and changing fire regimes in a Mediterranean environment. Agricultural and Forest Meteorology, 253-254, 190-202.

[12] ECMWF. European Centre for Medium-Range Weather Forecasts. (2019). https://www.ecmwf.int/en/forecasts/datasets

[13] Entekhabi, D., Njoku, E., O'Neill, P., Kellogg, K., Crow, W., Edelstein, W., . . . Van Zyl, J. (2010). The soil moisture active passive (SMAP) mission. Proceedings of the IEEE, 98(5), 704-716.

[14] ERA5-LST. (s.f.). https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5

[15] FAO, 2006. FireManagement-Global Assessment 2006. A Thematic Study Prepared in the Framework. (s.f.).

[16] Forkel, M., Thonicke, K., Beer, C., Cramer, W., Bartalev, S., & Schmullius, C. (2012). Extreme fire events are related to previous-year surface moisture conditions in permafrost-underlain larch forests of Siberia. Environmental Research Letters, 7(4).

[17] Jolly, W., Cochrane, M., Freeborn, P., Holden, Z., Brown, T., Williamson, G., & Bowman, D. (2015). Climate-induced variations in global wildfire danger from 1979 to 2013. Nature Communications, 6.

[18] Kerr, Y., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., . . . Gruhier, C. (2010). The SMOS mission: new tool for monitoring key elements of the global water cycle The SMOS Mission: New Tool 2 for Monitoring Key Elements of. (5), 666-687.

[19] LAND COVER. (s.f.). https://land.copernicus.eu/

[20] Le Vine, D., Lagerloef, G., & Torrusio, S. (2010). Aquarius and remote sensing of sea surface salinity from space. Proceedings of the IEEE, 98(5), 688-703.

[21] McArthur, A., & Australia. Forestry and Timber Bureau. (1967). Fire behaviour in eucalypt forests.

[22] Metzger, M., Bunce, R., Jongman, R., Mücher, C., & Watkins, J. (2005). A climatic stratification of the environment of Europe. Global Ecology and Biogeography, 14(6), 549-563.

[23] Metzger, M., Shkaruba, A., Jongman, R., & Bunce, R. (2012). EBONE EUROPEAN BIODIVERSITY OBSERVATION NETWORK.

[24] Oppenheimer, M., Campos, M., Warren, R., Birkmann, J., Luber, G., O'Neill, B., ... & Hsiang, S. (2015). Emergent risks and key vulnerabilities.

[25] Piles, M., Sánchez, N., Vall-Llossera, M., Camps, A., Martínez-Fernandez, J., Martinez, J., & Gonzalez-Gambau, V. (2014). A downscaling approach for SMOS land observations: Evaluation of high-resolution soil moisture maps over the Iberian peninsula. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(9), 3845-3857.

[26] Portal, G., Vall-Llossera, M., Piles, M., Camps, A., Chaparro, D., Pablos, M., & Rossato, L. (2018). A spatially consistent downscaling approach for SMOS using an adaptive moving window. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(6), 1883-1894.

[27] Rao, K., Anderegg, W., Sala, A., Martínez-Vilalta, J., & Konings, A. (2019). Satellite-based vegetation optical depth as an indicator of drought-driven tree mortality. Remote Sensing of Environment, 227, 125-136.

[28] Rodríguez-Fernández, N. J., Mialon, A., Mermoz, S., Bouvet, A., Richaume, P., Al Bitar, A., ... & Kerr, Y. H. (2018). An evaluation of SMOS L-band vegetation optical depth (L-VOD) data sets: High sensitivity of L-VOD to above-ground biomass in Africa. Biogeosciences, 15(14), 4627-4645.

[29] San-Miguel-Ayanz, J., Schulte, E., Schmuck, G., Camia, A., Strobl, P., Liberta, G., . . . Amatulli, G. (2012). Comprehensive Monitoring of Wildfires in Europe: The European Forest Fire Information System (EFFIS). En J. San-Miguel-Ayanz, E. Schulte, G. Schmuck, A. Camia, P. Strobl, G. Liberta, . . . G. Amatulli, Approaches to Managing Disaster - Assessing Hazards, Emergencies and Disaster Impacts. InTech.

[30] Schmugge, T., O'Neill, P. E., & Wang, J. (1986). Passive microwave soil moisture research. IEEE Transactions on Geoscience and Remote Sensing, (1), 12-22.

[31] UERRA. (2019). https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-uerra-europe-height-levels?tab=overview

[32] Ulaby, F. T., Moore, R. K., & Fung, A. (1981). Microwave remote sensing: Active and passive. volume 1-microwave remote sensing fundamentals and radiometry.

[33] Van Wagner, C. (1987). STRUCTURE OF THE CANADIAN FOREST FIRE WEATHER INDEX.