Predicting burnt areas during the summer season in Portugal by combining wildfire susceptibility and spring meteorological conditions

Rafaello Bergonse\textsuperscript{a}, Sandra Oliveira\textsuperscript{a}, Ana Gonçalves\textsuperscript{a}, Sílvia Nunes\textsuperscript{b}, Carlos DaCamar\textsuperscript{b} and José Luis Zezere\textsuperscript{a}

\textsuperscript{a}Centre for Geographical Studies, Universidade de Lisboa, Lisbon, Portugal; \textsuperscript{b}Instituto Dom Luiz (IDL), Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal

ABSTRACT
Wildfire susceptibility maps are a well-known tool for optimizing available means to plan for prevention, early detection, and wildfire suppression in Portugal, especially regarding the critical fire season (1 July — 30 September). These susceptibility maps typically disregard seasonal weather conditions on each given year, being based on predisposing variables that remain constant on the long-term, such as elevation. We employ logistic regression for combining wildfire susceptibility with a meteorological index representing spring conditions (the Seasonal Severity Rating), with the purpose of predicting, for any given year and ahead of the critical fire season, which areas will burn. Results show that the combination of the index with wildfire susceptibility slightly increases the capability to predict which areas will burn, when compared with susceptibility alone. Spring meteorological context was found better suited for predicting if the following summer wildfire season will be more severe, rather than predicting where wildfires will effectively occur. The model can be updated yearly after the critical wildfire season and can be applied to optimize the allocation of human and material resources regarding the prevention, early detection and suppression activities, required to reduce the severity of wildfires in the country.

ARTICLE HISTORY
Received 2 December 2020
Accepted 23 March 2021

KEYWORDS
Burnt area model; logistic regression; meteorological wildfire index; seasonal severity rating; wildfire susceptibility

1. Introduction
Wildfire is a major natural hazard in Portugal where, according to data for the period 1980-2018, the average annual number of fires exceeds those of the other Southern European countries (Spain, France, Italy and Greece), and annual burnt area per surface area of the country is second only to Spain. During this period, mean annual burnt area has been slightly over 116 000 ha, with much higher values on particularly severe years, such as 2003 (425 726 ha) or 2017 (540 630 ha) (San-Miguel-Ayanz et al.)
Most of the damage takes place during the summer months (M. G. Pereira et al. 2005) as result of a relatively small number of large fires (J. M. C. Pereira et al. 2006; M. G. Pereira et al. 2011).

Wildfire is controlled by numerous, interrelated factors. Meteorological conditions, both prior and simultaneous to ignition, are a major factor, influencing fuel availability, the fuel’s moisture content, and the conditions for propagation (e.g., Aldersley et al. 2011; Hernandez et al. 2015). Viegas et al. (2004) observed that the Drought Code (a sub-index of the Canadian Forest Service Fire Weather Index, FWI) relative to a single weather station explained 69% of the inter-annual variation in burnt area in Portugal between 1987 and 2000. Similarly, Carvalho et al. (2008) were able to explain up to 80% of the variance in monthly area burned in 11 Portuguese districts between 1980 and 2004, using the FWI and its individual components, along with other meteorological variables. M. G. Pereira et al. (2005) combined rainfall values between May and August with a descriptor of atmospheric circulation in the middle troposphere to produce estimates of annual burnt area, that showed a correlation of 0.82 with the actual values registered between 1980 and 1999. These authors defined two major atmospheric controls over annual burnt area in Portugal: the amount of precipitation during spring, and the occurrence of synoptic conditions that induce very hot and dry conditions over western Iberia during the summer. Late-spring atmospheric conditions were also used by S. A. Nunes et al. (2014) to predict the severity of yearly wildfire damage, and more recently by S. A. Nunes et al. (2019) to predict the severity of wildfire damage with respect to the summer months of July and August.

In parallel to meteorological controls, there are numerous properties inherent to the terrain itself that exert important effects over the ignition and propagation of wildfires. Examples are topographical variables such as roughness, slope inclination or elevation, or the nature and spatial arrangement of land cover, including both vegetation and human-made structures (e.g., roads). Moreover, land management practices and human density and behaviour exert great influence over the availability of fuel, the capacity of fire to propagate rapidly, the frequency and location of ignitions or the ease of access for firefighters (e.g., J. M. C. Pereira et al. 2006; Verde and Zêzere 2010; Oliveira et al. 2012; A. N. Nunes et al. 2016; Fernandes et al. 2016).

Modelling wildfire is complicated by the fact that controlling factors can influence one another and occasionally have simultaneous effects. For example, higher slope angle has a direct effect by promoting quicker fire propagation, but also an indirect effect by influencing land cover and therefore the type and availability of fuel. A relatively dense human population will likely imply the availability of means to detect and combat wildfire (promoting less severe events), but also a higher frequency of ignitions (e.g., Catry et al. (2008). Areas with alternating patches of agricultural and forest land uses (i.e., low fuel connectivity) will be relatively less prone to the occurrence of large fires (Fernandes et al. 2016), but will likely have more ignitions (A. N. Nunes et al. 2016).

The set of factors that influence wildfire occurrence and that constitute properties of the landscape define the territory’s susceptibility to wildfires. Verde and Zêzere (2010) defined this concept as “the terrain propensity to suffer a wildfire or to
support its spreading, given by the terrain’s intrinsic characteristics”. Beverly et al. (2009) defined it simply as the “the likelihood that an area will be burned by a fire”. More recently, Leuenberger et al. (2018) defined susceptibility as “an estimation of the probability that fire occurs in a specific area without considering a temporal scale, assessed on the basis of predisposing factors related to the terrain’s intrinsic characteristics”. Even when one or more of the characteristics used is meteorological, such as rainfall, it is expressed using mean annual values, which implies that the susceptibility assessment reflects averaged long-term conditions and will be of limited practical interest during any short period marked by specific meteorological conditions. In this context, Tonini et al. (2020) consider meteorological factors to be triggering factors and not predisposing factors, excluding them altogether from susceptibility analysis.

It is well known that, in Portugal, as well as in other countries with Mediterranean-type climate, the most severe wildfires occur in specific meteorological contexts (M. G. Pereira et al. 2005; San-Miguel-Ayanz et al. 2013; Fernandes et al. 2016; Ruffault et al. 2018) and that most wildfire damage takes place in the summer months (M. G. Pereira et al. 2013). Considering the long-term perspective of a wildfire susceptibility assessment, it is unsurprising that Verde and Zézere (2010) concluded that the inclusion of (mean annual) climatic variables did not produce any relevant increase in the predictive capacity of their wildfire susceptibility model, or that A. N. Nunes et al. (2016) excluded mean annual precipitation and temperature as factors for explaining annual burnt areas at the municipal level. Indeed, for the case of Portugal, weather conditions are found particularly relevant for models focusing on specific periods within one year, or to explain the interannual variability of wildfire occurrence in certain areas, but lose predictive power when averaged over longer time intervals (Oliveira et al. 2021). In contrast, meteorological wildfire indexes, such as the Canadian FWI and its components, are calculated daily by state agencies, such as the Portuguese Institute for Sea and Atmosphere (IPMA, the acronym in Portuguese), and are extremely relevant during the critical fire season (summer). Although they can be used to produce seasonal maps of wildfire hazard, because of their meteorological focus these maps ignore the spatial variation in wildfire susceptibility inherent to the terrain’s characteristics, which will inevitably play a role during each year’s critical season.

Given the above, this article has two main objectives. The first is to combine the wildfire susceptibility model for mainland Portugal adopted by the Portuguese Institute for the Conservation of Nature and Forests (ICNF) (Pahl and and IGOT 2020; Oliveira et al. 2021), with a seasonal meteorological index representing conditions in spring, in order to build a regression model capable of predicting, in each year and at the end of spring, which areas are most likely to burn during the subsequent summer season across mainland Portugal. Due to its predictive character covering a seasonal timescale, our approach precludes the inclusion of summer meteorological conditions in the model, which are also a well-known control over fire occurrence in Mediterranean areas (M. G. Pereira et al. 2005; Turco et al. 2017).

The second objective is to compare this combined model with similar models built with the individual components separately (susceptibility and meteorological index),
thus assessing if the combined approach significantly increases the capacity to predict summer burnt areas. Logistic regression was used due to its adequacy to model phenomena represented in a binary format (presence vs. absence), here interpreted in terms of burnt vs. unburnt. This technique has often been applied to wildfire; Vega Garcia et al. (1995) used it to predict days with fires vs. days without fires. Gudmundsson et al. (2014) used it to predict the occurrence of above-normal monthly burnt areas, whereas Chang et al. (2013), De Vasconcelos et al. (2001) and Vilar del Hoyo et al. (2011) employed it to predict the spatial location of ignitions. Besides its wide use in wildfire models, this technique is straightforward to apply, which facilitates the interpretation of the results obtained and the possible replication of the methodology in other regions or countries.

2. Data and methods

The study area corresponds to mainland Portugal, with a total area of approximately 89,015 km$^2$. Annual burnt areas for the 24 years between 1995 and 2018 were obtained in vector format from the Portuguese Institute for Nature Conservation and Forests (ICNF), with a minimum polygon size of 5 ha. These annual maps only include the date of each wildfire from 2012 onward, which implies that burnt areas mapped from 1995 to 2011 are not limited to the summer period (June 15 – September 15), which is the focus of this work. Despite this, it is well known that most of the yearly wildfire damage occurs in summertime (e.g., M. G. Pereira et al. 2005; 2013), as the analysis of the mean percentage of annual burnt area associated to each month during the period 2012-2018 shows (Figure 1). This issue was therefore not considered a limitation for this work. The original vector maps were converted to a binary raster, with yearly burnt areas having a value of 1 and the remaining areas 0.

2.1. Wildfire susceptibility

Wildfire susceptibility was calculated for the period 1995-2018 using the methodology employed by the Institute for Nature and Forests Conservation, described in detail in Pahl and IGOT (2020) and (Oliveira et al. 2021). For each pixel, susceptibility values are the result of the sum of the likelihood ratios associated to its classes in the variables landcover, slope, and elevation. The likelihood ratios of the different classes are obtained by crossing each classified variable with annual burnt areas. For each class $i$ of each variable, the score $L_{ri}$ is calculated as:

$$L_{ri} = \frac{S_i}{N_i/N}$$

Where $S_i$ is the number of burnt terrain units (pixels) corresponding to class $i$, $S$ is the number of burnt terrain units, $N_i$ is the number of terrain units associated to class $i$ and $N$ is the total number of terrain units. For a total of $n$ predisposing variables, the total LR score of each terrain unit ($L_{rj}$) is calculated as:
\[ L_{rj} = \sum_{i=1}^{n} X_{ij} \cdot L_{ri}, \]  

where \( X_{ij} \) equals 1 for the classes of the variables that are present and 0 for all others.

Topographic data was obtained from the European Environmental Agency’s Digital Surface Model, with a 25 m cell size (https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-eu-dem) (Figure 2). Land-cover data was obtained from the Portuguese General-Directorate of the Territory (Direcção-Geral do Território) (Figure 3 and Table 1).

As available landcover/use maps for the modelling period include 1995, 2007, 2010, 2015 and 2018, LR scores were calculated for each of these maps using only the years that are covered by the corresponding map, up until the year prior to the next map. Likelihood ratio scores were, therefore, calculated for the 1995 map using annual burnt areas for the years 1995-2006 (12 years), for the 2007 map using burnt areas for the years 2007-2009 (3 years), for the 2010 map using burnt areas for the years 2010-2014 (5 years), for the 2015 map using burnt areas for the years 2015-2017 (3 years) and for the 2018 map using burnt areas from that same year. The final LR

Figure 1. Mean monthly percentage of annual burnt area during the period 2012-2018. Source: yearly wildfire reports published by the Portuguese Institute for Nature Conservation and Forests (ICNF), available at: http://www2.icnf.pt/portal/florestas/dfci/relat/rel-if. Due to the inexistence of data for Nov-Dec 2017 and Oct-Dec 2018 within the reports, values for these five months were obtained from the annual burnt area maps published by the same institution. It is noteworthy that the mean value for October (23.49%) was strongly influenced by the exceptional year of 2017, in which 50.47% of all burnt area was concentrated in this month. If this year is disregarded, the mean percentage of annual burnt area corresponding to October is 1.73%, and the months of July, August and September represent on average 84.29% of yearly burnt area.
scores for each class were calculated as the mean of the scores associated to the successive landcover maps, weighted by the number of years covered by each map.

The methodology is schematized in Figure 4.

### 2.2. Seasonal severity rating

The meteorological index used, here called Seasonal Severity Rating (SSR), is calculated as the sum of the values of Daily Severity Rating (DSR), obtained from April 1 to June 15. The DSR is obtained through a simple transformation of the Fire Weather Index (FWI), which is the last component of the Canadian Forest Fire Weather Index System (CFFWIS) (Van Wagner 1987), and is considered more suitable than the FWI to be cumulated or averaged (S. A. Nunes et al. 2019). Although it is meant to express the difficulty in controlling wildfire, an elevated cumulative value in the months preceding the summer season implies a prevalence of relatively high temperatures and low rainfall. These, in turn, will promote water and thermal stress in vegetation in summertime, making it more prone to burning. The advantage of this predictive use of cumulative spring DSR values is that it allows knowing, ahead of the summer, if the vegetation will be more fire-prone because of the effect of weather on water content. This approach has been previously used by M. G. Pereira et al. (2013), S. A. Nunes et al. (2014) and S. A. Nunes et al. (2019) to anticipate the severity of summer wildfires.

**Figure 2.** Elevation (A) and slope angle (B) in mainland Portugal. Classes of elevation represent successive steps of 100 m, until the 800 m threshold, when classes become larger but occupy less area. Slope angle is classified in steps of 5 degrees until the 20-degree threshold, with the last class containing all values above. Source of data: https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-eu-dem.
Annual SSR maps representing spring weather conditions were computed from gridded daily values at 12:00 UTC of temperature, relative humidity, wind speed and 24 h-cumulated precipitation measured at 2 meters height, obtained from the ERA-Interim reanalysis dataset (Dee et al., 2011), and issued by the European Centre for Medium-Range Weather Forecasts (ECMWF). The ERA-Interim data were then reprojected onto the normalized geostationary projection (NGP) of Meteosat Second

Figure 3. Land cover/use distribution in mainland Portugal in 2018. Source: Directorate-General of the Territory (DGT, https://www.dgterritorio.gov.pt.).

Table 1. Percentage of Portuguese mainland territory occupied by each land cover/use class between 1995 and 2018. Source of the data: directorate-general of the territory (DGT, https://www.dgterritorio.gov.pt.).

| Land cover class                                | 1995 | 2007 | 2010 | 2015 | 2018 |
|------------------------------------------------|------|------|------|------|------|
| Agriculture and agro-forestry                   | 44.27| 39.78| 39.65| 38.89| 40.47|
| Eucalyptus forests                              | 7.81 | 9.14 | 9.42 | 9.91 | 10.40|
| Coniferous forests                              | 16.00| 16.34| 16.00| 14.48| 14.13|
| Other forests                                   | 12.75| 14.19| 13.65| 14.56| 14.33|
| Forests of invasive species                    | 0.01 | 0.14 | 0.14 | 0.17 | 0.18 |
| Scrub                                           | 12.87| 13.05| 13.34| 13.97| 12.41|
| Unvegetated or sparsely vegetated              | 0.71 | 0.84 | 0.83 | 1.03 | 0.68 |
| Artificialized territory                       | 3.81 | 4.85 | 5.02 | 4.87 | 5.21 |
| Water bodies                                    | 1.77 | 1.66 | 1.96 | 2.10 | 2.20 |
| Total                                          | 100  | 100  | 100  | 100  | 100  |
Generation (MSG) (Wolf 1999), with a mean cell size of about 4 km × 4 km over Portugal. Details about the procedure may be found in Pinto et al. (2018). Finally, the gridded values of SSR were interpolated into the 25 m grid cells of the study area.

In parallel to the employment of the annual SSR values, we experimented expressing each annual SSR value as an anomaly, calculated in relation to the mean SSR between the beginning of the period analyzed (1995) and 2018, using the following formula:

\[
\text{Anomaly } [\text{SSR}_{\text{year}}] = \frac{[\text{SSR}_{\text{year}} - \text{Mean } (\text{SSR}_{1995} - \text{SSR}_{2018})]}{\text{Mean } (\text{SSR}_{1995} - \text{SSR}_{2018})}
\]

In this way, the annual value in each cell expresses the degree to which that year’s SSR is exceptional in relation to that cell’s mean annual SSR, be it in positive or negative terms.

### 2.3. Construction of the point dataset

For each of the 24 years, we randomly generated 400 points, half over burnt cells, and half over unburnt ones, totalling 9600 points. The points were thus generated in the form of a simple random sample per year, without stratification. We then overlayed each annual point dataset to the maps of the different variables: the condition of burnt/unburnt (1/0) (the dependent variable), as well as wildfire susceptibility, SSR and the SSR anomaly (independent variables), extracting the values of the cell corresponding to each point to a table. Because the susceptibility map does not include urban or aquatic areas, all random points falling into these locations were eliminated, resulting in a total of 9263 valid data points (from the initial 9600). This number included 6800 data points within the period 1995-2011, for which no wildfire dates
are available, and 2778 data points within the period 2012-2018, explicitly related to summer wildfires (15 Jun-15 Sep). It is important to note that the annual burnt area maps do not contemplate the possibility of overlapping burnt areas in the same year, being expressed simply as burnt versus unburnt. In addition, in order to focus on the effects of meteorological conditions and terrain susceptibility, we are not considering other factors that may influence the size of area burned, such as suppression efforts or prevention strategies. Therefore, similar conditions conducive to area burnt were assumed each year and, as such, a simple random sampling was applied instead of a stratified approach.

In order to validate the use of the SSR as a wildfire factor in later modelling procedures, either in the form of absolute values (SSRAbs) or anomaly values (SSRAnom), we tested whether there were statistically significant differences between SSRAbs and SSRAnom values in burnt and unburnt areas. We performed the same procedure separately for the 1995-2011 data points (non-summer-specific) and the 2012-2018 data points (limited to summer wildfires). The Kolmogorov-Smirnov normality test was performed for each period and each variable, showing significantly (sig. < 0.001) that neither SSRAbs nor SSRAnom values follow a normal distribution in any of the two periods. We subsequently adopted the non-parametric Mann-Whitney test to compare the means. Test results showed that SSRAbs values are significantly different (p < 0.001) between burnt and unburnt areas, both for the 1995-2011 and the 2012-2018 periods. Conversely, however, it was not possible to reject the null hypothesis that SSRAnom values for burnt and unburnt areas belong to the same distribution, either for the 1995-2011 (p = 0.147) or the 2012-2018 period (p = 0.152). Given the impossibility of statistically discriminating between SSRAnom values in burnt and unburnt areas, we discarded it as an independent variable in the subsequent modelling procedures.

### 2.4. Modelling procedure

Following the standard procedure in modelling approaches, we separated the full dataset into independent modelling and validation datasets. We wanted to minimize the influence of non-summer wildfires in the evaluation of results, and therefore ensured that the validation dataset included only summer wildfires and that a part of the modelling dataset also corresponded to known summer fires. For this, we randomly selected a third of the summer-specific data points within the period 2012-2018 (926 points) and merged it with the 6800 data points for the period 1995-2011. The resulting 7411 data points (80% of all data points) were used as modelling dataset, with the remaining two-thirds of the summer-specific data points (1852 points, 20% percent of the total dataset) being used as validation dataset.

These operations for pre-processing the variables were carried out with ArcMAP 10.7.1 software (ESRI Inc.). The modelling dataset was then imported into IBM SPSS 24 software (IBM Corp.), which was used for all subsequent statistical analyses.

The logistic function calculates the probability that the desired instance of the dependent variable (i.e., \( y = 1 \)) will occur as a function of \( p \) independent variables, as formalized in (3).
This function is linearized using the logistic transformation, as shown in (4).

\[
P(y = 1) = \frac{\exp (\beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p)}{1 + \exp (\beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p)}
\]

(3)

\[
\ln \left( \frac{P(y = 1)}{1 - P(y = 1)} \right) = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p
\]

(4)

Logistic regression differs from linear regression regarding the method of estimation of regression coefficients, as it uses a Maximum Likelihood procedure as opposed to the Least Squares method. It does not require independent variables to be normally distributed, but only that they are not collinear (Maroco 2007).

Three regression models were built, using different sets of independent variables: i) using as independents only susceptibility (Susc); ii) using only the absolute values of Seasonal Severity Rating (SSRAbs), and iii) using both Susc and SSRAbs. The derived regression equations were then applied to the validation dataset, allowing to estimate the condition of burnt vs. unburnt for each cell. For each regression equation, results were then compared with the observed values to quantify the percentages of correct and incorrect estimations.

3. Results and discussion

3.1. Wildfire susceptibility

The likelihood ratio scores obtained for the different classes of elevation, slope and land cover are shown in Tables 2–4. The wildfire susceptibility map based on the period 1995-2018 is shown in Figure 5.

Likelihood ratio scores increase progressively for elevation classes up to a maximum of 3.2786 between 1000 and 1500 m, decreasing significantly afterwards (Table 2). The relation with slope is somewhat similar, with the likelihood ratios increasing from the lowest to the highest slope class (Table 3), although without decreasing in the highest class. Regarding the mean Likelihood Ratios associated to the various landcover classes, Table 4 shows that shrub/scrub and sparse vegetation have the highest scores, with 2.8573 and 3.6347, respectively. Regarding forested areas, chestnut forests show the highest mean LR score (1.7682), with Eucalyptus, Pinus Pinaster
and a subclass of oak forests (Other Oak forests) also having LR scores above 1.5. On the contrary, all agricultural and agroforestry areas show relatively low favorability, with LR scores below 1. It is recognized that topography, in particular convoluted terrain and higher slopes, exerts a strong influence over wildfire distribution (Calviño-Cancela et al. 2017; Carmo et al. 2011; A. N. Nunes et al. 2016; Oliveira and Zêzere 2020). Regarding landcover, shrubland (scrub) is a fire-prone landcover type in Mediterranean environments and it has the ability to rapidly colonize burned areas or abandoned farmland, thus promoting fuel accumulation (Carmo et al. 2011; Moreira et al. 2011; Barros and Pereira 2014; Oliveira et al. 2014). The spatial distribution of wildfire susceptibility (Figure 5) reflects the distribution of the most fire-prone landcover and topography classes in mainland Portugal. Values are highest in the central and northern regions, especially inland, with a secondary concentration of high values further south in the Algarve region.

Table 3. Likelihood Ratio scores (LR) obtained for each slope angle class.

| Slope angle (°) | LR    |
|----------------|-------|
| [0–5]          | 0.4459|
| [5–10]         | 1.2286|
| [10–15]        | 1.9400|
| [15–20]        | 2.3824|
| >20            | 2.5817|

Table 4. Likelihood Ratio scores (weighted mean) obtained for each landcover class.

| Landcover class                                                                 | LR    |
|--------------------------------------------------------------------------------|-------|
| Temporary irrigated + dryland crops                                             | 0.1983|
| Temporary crops and/or improved pastures + vineyards                          | 0.1682|
| Temporary crops and/or improved pastures + orchards                           | 0.3749|
| Temporary crops and/or improved pastures + olive groves                        | 0.3676|
| Olive groves                                                                   | 0.2613|
| Orchards                                                                       | 0.1509|
| Vineyards                                                                      | 0.1256|
| Protected agriculture/nurseries                                                | 0.0395|
| Complex cultivation and allotment mosaics                                      | 0.3261|
| Agriculture + natural and semi-natural spaces                                   | 0.7283|
| Holm oak-based agroforestry                                                    | 0.0779|
| Pinus pinea based agroforestry                                                 | 0.0728|
| Cork oak-based agroforestry                                                    | 0.1370|
| Cork oak and holm oak-based agroforestry                                      | 0.0960|
| Agroforestry based on other species                                            | 0.5741|
| Agroforestry based on other combinations                                       | 0.1812|
| Agroforestry based on other oaks                                               | 0.6688|
| Pinus Pinaster forests                                                         | 1.5656|
| Pinus Pinea forests                                                            | 0.2609|
| Cork-oak forests                                                               | 0.5304|
| Holm oak forests                                                               | 0.3880|
| Chestnut forests                                                               | 1.7682|
| Invasive species forests                                                       | 1.3856|
| Eucalyptus forests                                                             | 1.5305|
| Other broadleaved florests                                                     | 1.1995|
| Other coniferous forests                                                       | 1.0880|
| Other oak forests                                                              | 1.6969|
| Natural pastures                                                               | 0.9107|
| Improved pastures                                                              | 0.3030|
| Shrub - Scrub                                                                  | 2.8573|
| Sparse vegetation                                                              | 3.6347|
3.2. Seasonal severity rating

Annual spatial patterns of SSR values are marked by highest values in the south and in the inland portions of the territory, sometimes extending to the central region. Apart from the spatial patterns, the magnitude of the values varies significantly from year to year, as it can be exemplified by comparing the maximum values for the years 1998 (ca. 249) and 2017 (ca. 805) (Figure 6). The seasonal and year-to-year variability in weather conditions is a major factor in the annual differences in burned area extent (Jolly et al. 2015; Trigo et al. 2016). Exceptional years, marked by higher temperatures, lower humidity, drought conditions and/or stormy winds, have been responsible for recent wildfire disasters around the world, including Portugal (San-Miguel-Ayanz et al. 2013; Gómez-González et al. 2018; Turco et al. 2019). In these circumstances, landcover-specific favorability is reduced and nearly all landcover types can be affected (Barros and Pereira 2014). Moreover, the intensity of fires can reach such extreme levels as to overcome existing suppression abilities (Fernandes et al. 2016; Moreira et al. 2020). The role of weather conditions in fire activity thus justifies their integration with the purpose of complementing long-term terrain susceptibility maps.
Figure 6. Some examples of the variation in spatial patterns and value ranges defined by annual SSR values during the study period, namely for the years 1996, 1998, 2001, 2003, 2016 and 2017.
3.3. Logistic regression models

The results of the Omnibus test of model coefficients show that the independent variables influence the logit very significantly ($\alpha < 0.001$) in all cases (Table 5). A percentage threshold of 0.5 was used to distinguish unburnt from burnt cells (i.e., a cell was classified as burnt if the logit $\geq 0.5$). The results of Wald’s test, testing the null hypothesis that each regression coefficient equals 0 (i.e., that is does not contribute to the model) show that in all cases, the intercept and the independent variables contribute very significantly to the models. Despite the significant influence of the independent variables over the logits, in all three models Hosmer-Lemeshow’s test reveals a low adjustment between the model and the data, as confirmed by the relatively low Pseudo-$R^2$ values. The Pseudo-$R^2$ values also show that the model with both $Susc$ and $SSRAbs$ has the best adjustment to the data, followed by that with $Susc$ only. The small difference between both values indicates from the start that $SSRAbs$ has a minor role, as its inclusion leads only to an increase of 0.004 in Nagelkerke’s Pseudo-$R^2$.

Regarding the accuracy shown by the models (Tables 6 and 7), results demonstrate that the model with $SSRAbs$ as single predictor has the worst predictive capacity, with only 58.6% of correct predictions (Table 7). In contrast, the model with both $Susc$ and $SSRAbs$ has the best results, correctly predicting the condition of burnt/unburnt for 69.8% of the validation cells. All three models show a better capacity to predict burnt cells (Sensitivity) than unburnt ones (Specificity). The ROC (Relative Operating Characteristic) coefficients obtained by the models with $Susc$ and both $Susc$ and $SSRAbs$ (Table 7) are within the range proposed by Hosmer and Lemeshow (2000), cited by (Maroco 2007) as defining an “acceptable” discrimination ([0.70-0.80]).

Regarding the contribution of $SSRAbs$ for predicting burnt cells, its combination with $Susc$ is responsible for very small increases in model Specificity (+0.3%) and overall predictive accuracy (+0.1%), as well as a very slight increase of 0.001 in the ROC coefficient when compared with a model using $Susc$ only (Table 7). These values

---

**Table 5.** Parameters and test results for the three regression models constructed.

| Independent variables | Regression parameters | $\beta$ | Exp($\beta$) | Wald | $\alpha$ | Omnibus test of model coefficients | Hosmer-Lemeshow test | Nagelkerke’s Pseudo-$R^2$ |
|-----------------------|-----------------------|--------|-------------|-------|---------|-----------------------------------|---------------------|-------------------------|
| Susc                  | Intercept             | -2.259 | 0.104       | 1218.781 | 0.000   | 0.000                            | 0.000            | 0.311                   |
|                       | Susc                  | 0.573  | 1.774       | 1496.208 | 0.000   | 0.000                            | 0.000            | 0.104                   |
| SSRAbs                | Intercept             | 0.839  | 2.314       | 410.266   | 0.000   | 0.000                            | 0.000            | 0.315                   |
|                       | SSRAbs               | -0.004 | 0.996       | 513.340   | 0.000   | 0.000                            | 0.000            | 0.104                   |
| Susc + SSRAbs         | Intercept             | -1.948 | 0.143       | 453.771   | 0.000   | 0.000                            | 0.000            | 0.315                   |
|                       | Susc                  | 0.542  | 1.719       | 1131.609  | 0.000   | 0.000                            | 0.000            | 0.104                   |
|                       | SSRAbs               | -0.001 | 0.999       | 21.950    | 0.000   | 0.000                            | 0.000            | 0.104                   |

**Table 6.** Accuracy of the three regression models when applied to the validation subset: number of predicted cells in each group (P1-burnt; P0-unburnt).

| Predicted/ Observed | Model Susc |          | Model SSRAbs |          | Combined model (Susc, SSRAbs) |          |
|---------------------|------------|----------|--------------|----------|-------------------------------|----------|
|                     | P1 ($p \geq 0.5$) | P0 ($p < 0.5$) | Total       | P1 ($p \geq 0.5$) | P0 ($p < 0.5$) | Total       | P1 ($p \geq 0.5$) | P0 ($p < 0.5$) | Total       |
| O1                  | 657        | 261      | 918          | 579      | 339                           | 918      | 657        | 261      | 918          |
| O0                  | 301        | 633      | 934          | 428      | 506                           | 934      | 298        | 636      | 934          |
| Total               | 958        | 894      | 1852         | 1007     | 845                           | 1852     | 955        | 897      | 1852         |
indicate that, albeit statistically significant, the role played by \textit{SSRAbs} is a marginal one, with an overwhelming majority of the predictive capacity of the model being dependent on \textit{Susc}. Regardless of successful past applications of spring SSR values for predicting the severity of the following wildfire summer season (S. A. Nunes et al. 2014, 2019), an analysis of the structure of the FWI (Van Wagner 1987), of which the SSR is a transformation, suggests that better results could possibly be obtained if the DC or the BUI components were adopted instead of the final FWI value. Employing the final FWI implies taking into consideration wind speed (as a control over fire spread in the ISI component), detracting from our purpose of predicting summer water and thermal stress in vegetation. In contrast, using the DC or BUI components would narrow the input variables to those relevant for our purpose: relative humidity, precipitation, and air temperature.

Taking into consideration that the susceptibility values used have resulted from the weighted combination of slope, altitude and land cover, we can conclude that so-called “bottom-up variables” (Fernandes et al. 2016) exert a much larger effect over summer burnt area than spring meteorological conditions, as expressed by the Seasonal Severity Rating. This is in accordance with the results obtained by Fernandes et al. (2016), who observed that fuel-related variables (fuel composition and connectivity plus pyrodiversity) accounted for 81.8% of the variability in the size of large fires (\(>100\) ha) in Portugal, whereas weather/climate related variables only accounted for 14.5%. Late-spring weather conditions have been shown to be valuable in predicting if the following critical wildfire season will be severe (Gudmundsson et al. 2014; S. A. Nunes et al. 2014, 2019), as well as to exert a significant influence over annual burnt area (Viegas and Viegas 1994), but they seem to be less relevant when it comes to predicting which areas will effectively burn. Wildfire susceptibility maps based on intrinsic terrain properties are better suited for this latter purpose, with meteorological indexes such as the SSR providing a valuable, albeit minor contribution. The use of SSR allowed for the inclusion in our models of only one of the two atmospheric controls over burnt area during the critical fire season (M. G. Pereira et al. 2005), which is spring meteorological context. The other one (synoptic conditions inducing very hot and dry weather during summer) was shown by Turco et al. (2017) to play a more important role in fire occurrence in most Mediterranean areas than antecedent climate conditions. Although this control was not considered here, as its inclusion would invalidate the predictive purpose of the model, its relevance suggests an avenue for future research. This is to investigate the possibilities of (a) accurately predicting the occurrence of these synoptic conditions ahead of the summer season (e.g., Turco et al. 2017), and (b) integrating these predictions into a multivariate model such as the one we tested.

| Model | Susc | SSRAbs | Combined Susc, SSRAbs |
|-------|------|--------|------------------------|
| Sensitivity (P1O1/Total O1*100) (%) | 71.6 | 63.1 | 71.6 |
| Specificity (P0O0/Total O0*100) (%) | 67.8 | 54.2 | 68.1 |
| Correctly predicted (%) | 69.7 | 58.6 | 69.8 |
| ROC Coefficient | 0.755 | 0.619 | 0.756 |
Regarding the results obtained, it should be taken into consideration that the majority of the data points used for model building (6485 out of 7411) do not discriminate between fires occurring during the summer season and those that take place over the remaining of the year. This was necessary because state-produced wildfire maps in Portugal only include the date of each event since 2012. We considered this to be a minor issue due to the typical higher concentration of damage in the critical wildfire season. Nevertheless, it is expectable that if only summer fires were used for modelling, model performance could eventually be higher.

The question of the scale of analysis used also deserves consideration. Although we employed a country-wide, and therefore generalized approach in this work, the spatial variability regarding wildfire controls within the Portuguese territory gives rise to distinct fire regimes (e.g., Costa et al. 2011; Meneses et al. 2018). It is likely that spring meteorological conditions, as represented here using the SSR, will play a more important role in certain areas that in others, given the spatial variations in variables such as the type of fuel and its dependence on humid spring conditions to ensure fuel continuity during the summer (Turco et al. 2017). It would therefore be interesting in a future approach to apply the same methodology separately to a set of contrasting study areas within the country, and to compare the predictive capacities of the resulting models.

Although logistic regression was adopted in this work and produced valuable results, other data modelling techniques could be also applied, such as Classification Trees or Discriminant Analysis. It would be interesting in future approaches to compare different techniques as to their capacity to combine these same two variables into a predictive burnt area model.

4. Conclusions

We have shown that combining a conventional wildfire susceptibility map based on intrinsic properties of the territory with absolute values of the Seasonal Severity Rating (SSR), a meteorological index descriptive of late spring weather conditions based on the Canadian Forest Service Fire Weather Index, only very slightly increases our capacity to predict which areas will burn across the Portuguese territory during any given summer, compared with susceptibility alone. A model using the SSR as sole predictor variable revealed a markedly lower predictive capacity regarding where wildfires will occur. It seems that the spring meteorological context, seen in the spatial distribution and values range of the SSR index, appears as more suitable to analyse the potential severity of the following wildfire summer season, as previously done by S. A. Nunes et al. (2014) and S. A. Nunes et al. (2019), than to predict the spatial location of burnt areas.

The simple methodology used can be easily applied to update the predictive model at the end of each spring, contributing to improve the effectiveness of wildfire prevention and detection efforts, as well as of preventive and firefighting resources allocation prior to the following critical wildfire season. Although the increases in predictive capacity enabled by the inclusion of the SSR are modest from a modelling perspective done at the national level, they could be valuable in practical terms when
applied to a territory where areas consumed by summer fires often rank on the hundreds of thousands of hectares (e.g., years 2003 and 2017 in Portugal). This work has showed that the use of spring weather conditions as a predictor of burnt areas in the subsequent summer constitutes a potential alternative path to improve existing spatial models, which should be further explored.

Disclosure statement
There are no relevant financial or non-financial competing interests to report.

Funding
This research was undertaken in the context of the project People&Fire, funded by the Portuguese Foundation for Science and Technology under grant PCIF/AGT/0136/2017, and by the Research Unit (UIDB/00295/2020 and UIDP/00295/2020).

ORCID
José Luis Zézere http://orcid.org/0000-0002-3953-673X

Data availability statement
The data produced are made freely available by the authors. Requests should be enclosed to the corresponding author.

References
Aldersley A, Murray SJ, Cornell SE. 2011. Global and regional analysis of climate and human drivers of wildfire. Sci Total Environ. 409(18):3472–3481.
Barros AMG, Pereira JMC. 2014. Wildfire selectivity for land cover type: does size matter? PLoS ONE. 9(1):e84760.
Beverly JL, Herd EPK, Conner JCR. 2009. Modeling fire susceptibility in west central Alberta, Canada. For Ecol Manag. 258(7):1465–1478.
Calviño-Cancela M, Chas-Amil ML, García-Martínez ED, Touza J. 2017. Interacting effects of topography, vegetation, human activities and wildland-urban interfaces on wildfire ignition risk. For Ecol Manag. 397:10–17.
Carmo M, Moreira F, Casimiro P, Vaz P. 2011. Land use and topography influences on wildfire occurrence in northern Portugal. Landsc Urban Plan. 100(1-2):169–176.
Carvalho A, Flannigan MD, Logan K, Miranda AI, Borrego C. 2008. Fire activity in Portugal and its relationship to weather and the Canadian Fire Weather Index System. Int J Wildland Fire. 17(3):328–338.
Catry FX, Rego FC, Moreira F, Bação F. 2008. Characterizing and modelling the spatial patterns of wildfire ignitions in Portugal: fire initiation and resulting burned area. WIT Trans Ecol Environ. 119:213–221.
Chang Y, Zhu Z, Bu R, Chen H, Feng Y, Li Y, Hu Y, Wang Z. 2013. Predicting fire occurrence patterns with logistic regression in Heilongjiang Province, China. Landscape Ecol. 28(10):1989–2004.
Costa L, Thonicke K, Poulter B, Badeck F. 2011. Sensitivity of Portuguese forest fires to climatic, human, and landscape variables: subnational differences between fire drivers in extreme fire years and decadal averages. Reg Environ Change. 11(3):543–551.

De Vasconcelos MJP, Silva S, Tomé M, Alvim M, Pereira JMC. 2001. Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. Photogramm Eng Remote Sensing. 67:73–81.

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, et al. 2011. The ERA-interim reanalysis: configuration and performance of the data assimilation system. . 137(656):553–597. QJR Meteorol Soc

Fernandes PM, Monteiro-Henriques T, Guiomar N, Loureiro C, Barros AMG. 2016. Bottom-up variables govern large-fire size in Portugal. Ecosystems. 19(8):1362–1375.

Gómez-González S, Ojeda F, Fernandes PM. 2018. Portugal and Chile: longing for sustainable forestry while rising from the ashes. Environ Sci Policy. 81(2017):104–107.

Gudmundsson L, Rego FC, Rocha M, Seneviratne SL. 2014. Predicting above normal wildfire activity in southern Europe as a function of meteorological drought. Environ Res Lett. 9(8):084008.

Hernandez C, Drobinski P, Turquety S. 2015. How much does weather control fire size and intensity in the Mediterranean region? Ann Geophys. 33(7):931–939. https://doi.org/10.5194/angeo-33-931-2015.

Hosmer D, Lemeshow S. 2000. Applied logistic regression. 2nd ed. New York: John Wiley & Sons.

Hosmer D, Lemeshow S. 2000. Applied logistic regression. 2nd ed. New York: John Wiley & Sons.

Jolly WM, Cochrane MA, Freeborn PH, Holden ZA, Brown TJ, Williamson GJ, Bowman DMJSJS. 2015. Climate-induced variations in global wildfire danger from 1979 to 2013. Nat Commun. 6(1):1–11.

Leuenberger M, Parente J, Tonini M, Pereira MG, Kanevski M. 2018. Wildfire susceptibility mapping: deterministic vs. stochastic approaches. Environ Modell Software. 101:194–203.

Maroco J. 2007. Análise Estatística com Utilização do SPSS. 3rd ed. Lisboa: Edições Século.

Meneses BM, Reis E, Reis R. 2018. Assessment of the recurrence interval of wildfires in mainland Portugal and the identification of affected LUC patterns. J Maps. 14(2):282–292.

Moreira F, Ascoli D, Safford H, Adams MA, Moreno JM, Pereira JMC, Catry FX, Armesto J, Bond W, González ME, et al. 2020. Wildfire management in Mediterranean-type regions: paradigm change needed. Environ Res Lett. 15(1):011001.

Moreira F, Viedma O, Arianoutsou M, Curt T, Koutsias N, Rigolot E, Barbati A, Corona P, Vaz P, Xanthopoulos G, et al. 2011. Landscape-wildfire interactions in southern Europe: implications for landscape management. J Environ Manag. 92(10):2389–2402.

Nunes AN, Lourenço L, Meira ACC. 2016. Exploring spatial patterns and drivers of forest fires in Portugal (1980-2014). Sci Total Environ. 573:1190–1202.

Nunes SA, Camara CC, da Turkman K, Ermida SL, Calado T. 2014. Anticipating the severity of the fire season in Northern Portugal using statistical models based on meteorological indices of fire danger. In Viegas DX editor. Advances in forest fire research. Coimbra: Imprensa da Universidade de Coimbra; p. 1634–1645. https://doi.org/10.14195/978-989-26-0884-6_180

Nunes SA, Dacamara CC, Turkman K, Calado T, Trigo RM, Turkman M. 2019. Wildland fire potential outlooks for Portugal using meteorological indices of fire danger. Nat Hazards Earth Syst Sci. 19(7):1459–1470.

Oliveira S, Gonçalves A, Zézere J L. 2021. Reassessing wildfire susceptibility and hazard for mainland Portugal. Sci Total Environ. 762:143121 doi:10.1016/j.scitotenv.2020.143121. PMC: 33129531

Oliveira S, Moreira F, Boca R, San-Miguel-Ayanz J, Pereira JMC. 2014. Assessment of fire selectivity in relation to land cover and topography: A comparison between Southern European countries. Int J Wildland Fire. 23(5):620–630.

Oliveira S, Oehler F, San-Miguel-Ayanz J, Camia A, Pereira JMC. 2012. Modeling spatial patterns of fire occurrence in Mediterranean Europe using multiple regression and random forest. For Ecol Manag. 275:117–129.

Oliveira S, Zézere JL. 2020. Assessing the biophysical and social drivers of burned area distribution at the local scale. J Environ Manag. 264:110449.
Pahl, & IGOT. 2020. Análise da Suscetibilidade de Incêndio Rural em Portugal Continental. Modelação da perigosidade de suporte à produção de cartas de perigosidade de incêndio rural de cariz conjuntural de “outono” e de “primavera.
Pereira JMC, Carreiras JMB, Silva JMN, Vasconcelos MJ. 2006. Alguns Conceitos Básicos sobre os Fogos Rurais em Portugal. In Pereira JS, Pereira JMC, Rego FC, Silva JMN, Silva TP editors. Incêndios Florestais em Portugal: Caracterização, Impactes e Prevenção. Lisboa: ISAPress; p. 133–161.
Pereira JMC, Carreiras JMB, Silva JMN, Vasconcelos MJ. 2006. Alguns Conceitos Básicos sobre os Fogos Rurais em Portugal. In Pereira JS, Pereira JMC, Rego FC, Silva JMN, Silva TP editors. Incêndios Florestais em Portugal: Caracterização, Impactes e Prevenção. Lisboa: ISAPress; p. 133–161.
Pereira MG, Malamud BD, Trigo RM, Alves PI. 2011. The history and characteristics of the 1980-2005 Portuguese rural fire database. Nat Hazards Earth Syst Sci. 11(12):3343–3358.
Pereira MG, Calado TJ, DaCamara CC, Calheiros T. 2013. Effects of regional climate change on rural fires in Portugal. Clim Res. 57(3):187–200.
Pereira MG, Trigo RM, Da Camara CC, Pereira JMC, Leite SM. 2005. Synoptic patterns associated with large summer forest fires in Portugal. Agric For Meteorol. 129(1–2):11–25.
Pinto MM, DaCamara CC, Trigo IF, Trigo RM, Turkman KF. 2018. Fire danger rating over Mediterranean Europe based on fire radiative power derived from Meteosat. Nat Hazards Earth Syst Sci. 18(2):515–529.
Ruffault J, Curt T, Martin-Stpaul NK, Moron V, Trigo RM. 2018. Extreme wildfire events are linked to global-change-type droughts in the northern Mediterranean. Nat Hazards Earth Syst Sci. 18(3):847–856.
San-Miguel-Ayanz J, Durrant T, Boca R, Libertá G, Branco A, de Rigo D, Ferrari D, Maianti P, Vivancos TA, Oom D, et al. 2019. Forest fires in Europe, Middle East and North Africa 2018. EUR 29856 EN. European Union. https://doi.org/10.2760/1128
San-Miguel-Ayanz J, Moreno JM, Camia A. 2013. Analysis of large fires in European Mediterranean landscapes: lessons learned and perspectives. For Ecol Manag. 294:11–22.
Tonini M, D’andrea M, Biondi G, Esposti SD, Trucchia A, Fiorucci P. 2020. A machine learning-based approach for wildfire susceptibility mapping. The case study of the Liguria region in Italy. Geosciences (Switzerland). 10(3):105–118.
Trigo RM, Sousa PM, Pereira MG, Basila D, Gouveia CM. 2016. Modelling wildfire activity in Iberia with different atmospheric circulation weather types. Int J Climatol. 36(7):2761–2778.
Turco M, Ceglar A, Prodhomme C, Soret A, Toreti A, Doblas-Reyes Francisco J. 2017. Summer drought predictability over Europe: empirical versus dynamical forecasts. Environ Res Lett. 12(8):084006.
Turco M, von Hardenberg J, AghaKouchak A, Llasat MC, Provenzale A, Trigo RM. 2017. On the key role of droughts in the dynamics of summer fires in Mediterranean Europe. Sci Rep. 7(81):1–10.
Van Wagner CE. 1987. Development and Structure of the Canadian Forest Fire Weather Index System - Forestry Technical Report no.35.
Garcia CV, Woodard PM, Titus SJ, Adamowicz WL, Lee BS. 1995. A logit model for predicting the daily occurrence of human caused forest-fires. Int J Wildland Fire. 5(2):101–111.
Verde JC, Zézere JL. 2010. Assessment and validation of wildfire susceptibility and hazard in Portugal. Nat Hazards Earth Syst Sci. 10(3):485–497.
Viegas DX, Reis RM, Cruz MG, Viegas MT. 2004. Calibração do sistema Canadiano de Perigo de Incêndio para aplicação em Portugal. Silva Lusitana. 12(1):77–93.
Viegas DX, Viegas MT. 1994. A relationship between rainfall and burned area for Portugal. Int J Wildland Fire. 4(1):11–16.
Vilar del Hoyo L, Isabel MPM, Vega FJM. 2011. Logistic regression models for human-caused wildfire risk estimation: analysing the effect of the spatial accuracy in fire occurrence data. Eur J Forest Res. 130(6):983–996. https://doi.org/10.1007/s10342-011-0488-2.
Wolf R. 1999. LRIT/HRIT global specification. In Coordination Group for Meteorological Satellites (Issue August). Technical Report 2.6. EUMETSAT. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:LRIT+/+HRIT+Global+Specification#0