INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2013)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.3749

RMetS
Royal Meteorological Society

Modelling wildfire activity in Iberia with different atmospheric circulation weather types

Ricardo M. Trigo,[a,b]* Pedro M. Sousa,[a] Mário G. Pereira,[c,a] Domingo Rasilla[d] and Célia M. Gouveia[a]

[a] *IDL, University of Lisbon, Lisbon, Portugal*
[b] *Departamento de Engenharias, Universidade Lusófona, Lisboa, Portugal*
[c] *Centro de Investigação e de Tecnologias Agro-Ambientais e Biológicas (CITAB), Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal*
[d] *GIMENA, Departamento de Geografía, Urbanismo y OT, Universidad de Cantabria, Santander, Spain*

ABSTRACT: This work focuses on the spatial and temporal variability of burnt area (BA) in the entire Iberian Peninsula (IP) and on the construction of statistical models to reproduce the inter-annual variability. A novel common dataset was assembled for the whole IP by merging the registered BA from 66 administrative regions of both Portugal and Spain. We applied a cluster analysis to identify larger regions with similar fire regimes and results point to the existence of four clusters (Northwestern, Northern, Southwestern and Eastern) whose spatial patterns and seasonal fire regimes are shown to be related with constraining factors such as topography, vegetation cover and climate conditions. The relationship between BA at monthly time scale with both long-term climatic pre-conditions and short-term synoptic forcing was assessed using correlation and regression analysis based on: (1) temperature and precipitation from 2 to 7 months in advance to fire peak season, (2) synoptic weather patterns derived from 11 distinct Weather Types Classifications (WTC). Different relations were obtained for each IP region with a relevant link being identified between BA and short-term synoptic forcing for all clusters, while the relation with long-term climatic preconditioning was relevant for all but one cluster. Stepwise regression models based on the best climatic and synoptic circulation predictors were developed with cross-validation to avoid over fitting. The performance of the models varies within IP regions, though models exclusively based on WTC tend to better reproduce the annual BA time series than those merely based on pre-conditioning climatic information. Nevertheless, the use of both synoptic and climatic predictors provides the best results, particularly for the two western clusters, with Pearson correlation coefficient values higher than 0.7. Finally, it is shown that typical synoptic configurations that favour high values of BA correspond to dry and warm wind flows associated with anti-cyclonic regimes.

KEY WORDS wildfires; Iberian Peninsula; weather types; cluster analysis; regression models

Received 14 November 2011; Revised 29 April 2013; Accepted 3 May 2013

1. Introduction

It is widely known that, as a consequence of the Mediterranean type of climate, the southern European countries (Portugal, Spain, France, Italy and Greece) are particularly affected by summer fires (Barbosa *et al*., 2007; JRC, 2010). The frequent warm and dry meteorological conditions found in summer play an obvious role in the triggering and spreading of these fires. Moreover, recent and future trends towards a dryer (Mariotti *et al*., 2008; Mariotti, 2010; Sousa *et al*., 2011) and warmer (Giorgi, 2006; Fischer and Schär, 2010) Mediterranean climate will tend to exacerbate the problem. However, while the prevailing weather conditions play an important role it must be acknowledged that human contribution is also significant either in Spain (Valbuena-Carabaña *et al*., 2010) or in Portugal (Costa *et al*., 2010). For example, the widespread abandonment and depopulation of rural areas in Iberia in the last decades has contributed to an increase in fuel availability for the ignition and spreading of fires (Pausas and Vallejo, 1999; Lloret *et al*., 2002). This pattern of events is in line with the rest of the Mediterranean basin, where areas of scarcely relevance for agriculture were either converted to forest plantations or abandoned to the natural process of ecological succession, often converted to shrublands and woodlands (Moreno *et al*., 1998; Pausas and Vallejo, 1999). In fact, it should be noticed that the use of fire as a tool to modify the landscape has been used on much longer temporal scales throughout the Holocene in both southern Iberian Peninsula (IP) (Gil-Romera *et al*., 2010) as well as in northern Iberia (Rubiales *et al*., 2008).

Among all northern Mediterranean regions, it must be stressed that wildfires constitute a major and recurrent hazard in the IP. Their occurrence is responsible for a very large amount of burnt area (BA) every year, as well as for significant human and socio-economical impacts.

* Correspondence to: R. M. Trigo, Centro de Geofísica da Universidade de Lisboa, Faculdade de Ciências, Univ. de Lisboa, Campo Grande, Ed C8, Piso 6, 1749-016 LISBOA, Portugal. E-mail: rmtrigo@fc.ul.pt

For example, during the period 1980–2009, the average annual BA by wildfires in Portugal (Spain) was about 109 000 ha (177 000 ha) (Pereira *et al*., 2011a). However, if we compute the ratio of total BA by the country total land area, Portugal leads the rank of European countries most affected by fires, with a ratio up to three times higher than Italy, the second most fire-affected country (Pereira *et al*., 2011a). Nevertheless spatial distribution patterns of fire occurrences and BA are not homogeneous within southern Europe, neither within each country (JRC, 2010). On the basis of data compiled between 1985 and 1997 (European Commission, 1998), northwestern Iberia (Portugal and the Spanish Provinces of Galicia) is the European region with the largest percentage of fire occurrences (roughly 50% of total number of fires), followed by southern Italy (10%) and the triangle Provence-Tuscany-Corsica (7%) (EC, 1998; Pereira *et al*., 2005). However, one should be careful because these annual average values hide a remarkable level of inter-annual variability. For example, focusing on the burned surface in Portugal between 2003 and 2007, the exceptional summers of 2003 (450 000 ha) and 2005 (380 000 ha), present the highest values of BA since robust statistics started to be compiled in 1980 (Trigo *et al*., 2006; Pereira *et al*., 2011a), contrast with average-summers like 2004 or 2006 (120 000 and 80 000 ha of BA, respectively), or even 2007, which represents the second lowest record with just 30 000 ha of BA (Pereira *et al*., 2011a).

The number of fires and their extent in mid-latitude regions are controlled primarily by natural factors (e.g. topography, weather conditions and vegetation cover), by land management practices responsible for anthropogenic ignitions, and also by fire prevention, management and suppression efforts. Changes in the frequency of occurrence of specific atmospheric conditions, which are favourable to the outbreak and consequent spreading of large wildfires, can help explaining the magnitude of this inter-annual variability (Viegas and Viegas, 1994; Pyne *et al*., 1996, Kunkel, 2001). In the context of large fire occurrences in the IP, a number of works has focused objectively on the role played by meteorological factors, namely in Portugal (Trigo *et al*., 2006; Hoinka *et al*., 2009; Carvalho *et al*., 2008) and in Spain (Vázquez and Moreno, 1993; Piñol *et al*., 1998; Pausas, 2004; Padilla and Vega-García, 2009; Rasilla *et al*., 2010; Pausas and Fernández-Muñoz, 2011). Moreover, several works have proved that Iberian fires are also associated with the occurrence of anomalous climate in the preceding seasons (Viegas and Viegas, 1994; Pereira *et al*., 2005). However, a growing number of studies focused on the anthropogenic contribution towards the spatial distribution of Iberian fires (Vélez, 1993, Badía *et al*., 2002; Lloret *et al*., 2002; Vázquez *et al*., 2002; Cueva *et al*., 2006; Costa *et al*., 2010). Among these, one of the most comprehensive approaches corresponds to the recent work of Costa *et al*. (2010), where the authors focus on the sensitivity of Portuguese forest fires to climatic, human activity and landscape. Similarly, in Cueva *et al*. (2006), the authors showed the dependence of regional patterns of forest fires in Spain on numerous human, landscape and climatic factors that change frequently in time and space.

Despite the amount of studies, there are still issues to be further addressed in relation to the interactions between meteorological conditions, vegetation dynamics and fires. In the Mediterranean region, drought periods can have strong impacts in vegetation activity (Gouveia *et al*., 2009), with significant losses of crop yields (Austin *et al*., 1998), decreasing of terrestrial net primary production (Zhao and Running, 2010) and forest growing (Martínez-Villalta *et al*., 2008) and increasing the risk of forest fires (Pausas, 2004). It is now widely accepted that changes in large-scale atmospheric circulation is partially responsible for decreasing precipitation trends and higher frequency of drought episodes over Iberia (Paredes *et al*., 2006; García-Herrera *et al*., 2007). Moreover, such a trend towards a drier Mediterranean climate is in very good agreement with state-of-the-art climate change projections for different future scenarios (Giorgi, 2006). Finally, other studies have proved that the occurrence of major droughts in southern Europe during the preceding winter and spring seasons can enhance the amplitude of heat waves on the following summer (Seneviratne *et al*., 2006; Fischer *et al*., 2007), implying that these two phenomena (droughts and heat waves) are closely related.

A number of studies have used remote sensing imagery-derived datasets over the Mediterranean region to analyse changes in vegetation activity (Vicente-Serrano and Heredia-Lacaustra, 2004; Gouveia *et al*., 2008), have determined the impact of droughts (Vicente-Serrano, 2007; Gouveia *et al*., 2009) and have monitored plant recovery after fire (Viedma *et al*., 2006, Gouveia *et al*, 2010). Several vegetation indices can be used to assess vegetation dynamics; however, the Normalized Difference Vegetation Index (NDVI) is the most widely employed (Myneni *et al*., 1995; Gouveia *et al*., 2008).

Pereira *et al*. (2005) found that the daily synoptic variability in Portugal is the most important driver of favourable local fire-prone weather conditions. Bearing this in mind, atmospheric circulation classifications become an important potential tool to study the role of weather in wildfire occurrence in the IP. Furthermore, impacts on fire activity of large scale circulation patterns in the atmosphere and in the ocean have been used with the aim of developing short- and long-range forecast models for wildfire occurrences by several authors. In this respect, Flannigan *et al*. (2001) explored the relationship between Pacific sea surface temperature (SST) and BA in Canada to build a forecasting model of monthly and seasonal fire activity. Here, we intend to apply a similar approach on the specific context of the IP, and to introduce Weather Types Classification (WTC) on a deeper assessment of fire-season predictability for this region. In this work, we will be mostly focused on the relationship between meteorological and BA inter-annual variability, through the use of summer synoptic circulation and previous climatological conditions in the entire IP.

WTCs are a recurrent and fairly simple method to describe atmospheric circulation variability at daily and sub-daily scale, as they provide discrete characterizations of the atmospheric conditions. They have been widely used in several contexts and, as a consequence, a large variety of classification methods (automated or not) have been developed with multiple purposes. Here we are considering the WTCs developed within the framework of the COST Action 733—Harmonization and applications of WTC for European Regions (COST733) catalogue.

The COST733 catalogue comprehends the largest set of classifications ever assembled for different regions of Europe. These are based on numerous methods and meteorological variables and have been applied to several spatial domains centred on different European sub-regions. We intend to make the best use of this catalogue and perform an objective comparison on the ability of each classification to explain wildfire variability in the IP domain. Numerous objective comparisons for the catalogue have been performed on several areas of study, such as heavy precipitation occurrence (Lupikasza, 2010; Twardosz, 2010), extreme values of air temperature (Ustrnul *et al*., 2010), lightning activity (Pineda *et al*., 2010), air pollution (Lésniok *et al*., 2010). Preliminary works relating WTC with wildfire occurrences have been developed for Iberia (Rasilla *et al*., 2010) and Greece (Kassomenos, 2010). However, Rasilla *et al*. (2010) did not use WTC from the COST733 catalogue, but rather his own classifications, based on Principal Components and Cluster analyses.

Therefore, the main objectives of this work are the following:

(1) To obtain a novel comprehensive wildfire dataset for the entire IP and to distinguish fairly independent regions in terms of BA characteristics and fire regime for each region;
(2) To provide an interpretation on the spatial distribution of BA in terms of the topography, climate and vegetation dynamics in the IP;
(3) To analyse the relationships and dependence of BA inter-annual variability with previous seasonal climatological conditions, namely precipitation and air temperature;
(4) To examine the performance of the different WTC available from COST733 catalogue as indicators for the BA inter-annual variability;
(5) To test the ability of simple statistical regression models to reproduce the inter-annual variability of BA, using climatological data and WTC as predictors.

The rest of this manuscript has the following structure: the datasets used in this study and the climate of IP are described in Section 2, while the identification of the most important spatial regions with identical fire behaviour was obtained with cluster analysis in Section 3. The links between BA for each cluster and previous meteorological fields and/or the frequency of each weather type (WT) are evaluated in Section 4, while Section 5 provides a description of the most important WT patterns associated with fire occurrences at each cluster previously identified. Finally, section 6 is devoted to the discussion of obtained results and to present the conclusions of this study.

2. Datasets and main characteristics of the study area

This work relies on three types of data series, namely: (1) WT catalogues and meteorological variables; (2) BA in both Iberian countries (Portugal and Spain) and (3) vegetation index.

2.1. WT catalogues

WTC data was obtained from the COST733 project (Philipp *et al*., 2010) and although available since 1958 it ends in 2001, which constrains its use to the 22-year long period (1980–2001) to be coincident with fire data. We only considered classifications with nine categories (in a total of 11 classifications), and exclusively worked in the domain D09 that is centred over the IP. This domain (amongst all available for the catalogue) is represented in Figure 1, together with all the remaining regional and pan-European domains. The acronym of the considered 11 classifications is presented in Table I, together with the climatic elements and main methodological methods used in their construction.

2.2. A climate characterization of the IP

The climate of the study area is characterized by using the mean monthly precipitation and temperature series obtained from long Climatic Research Unit (CRU) gridded dataset—version TS3.0 at the resolution of $0.5^\circ \times 0.5^\circ$ latitude (Mitchell and Jones, 2005). To assess the IP climatology within the appropriate context, we must acknowledge the main physical attributes of the region, namely the topography and mountain ranges. These geographical characteristics and the political borders of the Iberian countries are depicted in Figure 2.

Maps of annual averages of surface temperature and precipitation in the IP computed for the 1980–2005 period using the CRU gridded dataset are shown in Figure 3. The overall patterns are quite similar to higher resolution Iberian climate atlas maps of annual average mean air temperature and precipitation computed for the 1971–2000 period (AEMET-IM, 2011) obtained with observed data from a large number of meteorological and udometric stations. The annual precipitation in the IP (Figure 3, top panel) presents a SE–NW gradient, with higher values in the coastal regions extending from Lisbon, in central Portugal to Galicia (northwestern Spain) and also to the northern mountainous regions (including the Cantabria and Pyrenees) and the lowest values found in the southern-eastern corner of the IP. Annual mean air temperature (Figure 3, bottom panel)

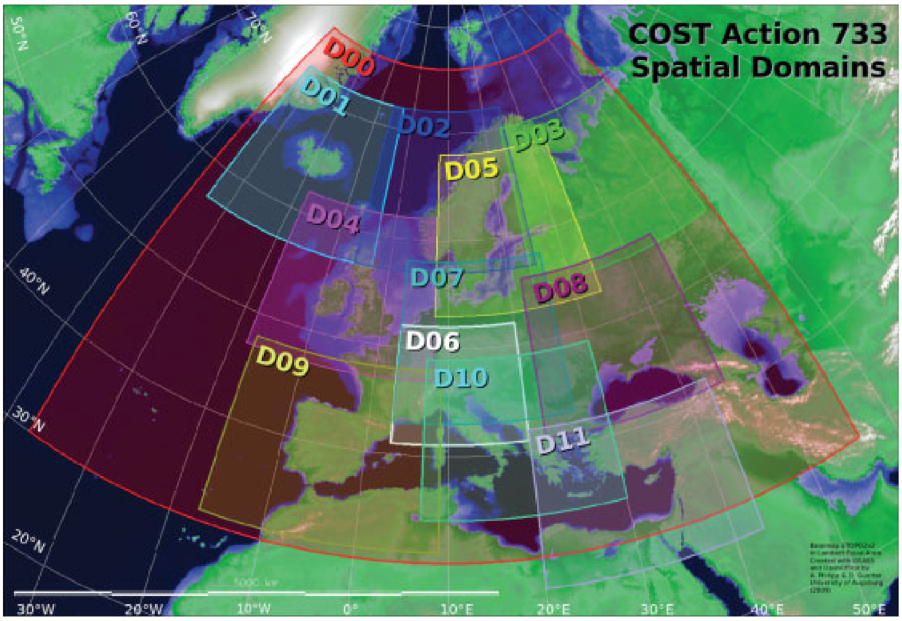

Figure 1. European Spatial domains defined by COST Action 733—Harmonization and Applications of Weather Type Classifications for European regions (COST733Wiki, 2011). Catalogues/classifications used in this work were produced for domain D09 (Iberian Peninsula, Western Mediterranean), 31°N—48°N (18 points), 17°W—9°E (27 points).

Table I. Summary of the 11 Weather Type Classifications (WTC) characteristics used in this study obtained from COST733 (COST733Wiki, 2011), namely the acronym of the WTC, main methodology and meteorological variables used for the weather typing classification.

WTC	Method	Meteorological variables
CKMEANSC09	k-means	SLP
ESPLC09	spatial similarity index of daily maps	SLP
KHC09	spatial correlation of daily maps	SLP
LITADVE	direction of advection, cyclonicity	SLP
LUNDC09	correlation-based	SLP
NNWC09	artificial neural networks	SLP
PCACAC09	S-mode PCA, k-means	SLP
PETISCOC09	correlation-based	SLP
SANDRAC09	simulated annealing	SLP
SANDRASC09	simulated annealing of sequences	SLP
WLKC09	direction of advection, cyclonicity	U700,V700, Z925 and Z500

seems to be dominated by the latitudinal change and the topography with values ranging between about 20 °C in southern Spain to about 2.5 °C in the high regions in the north of Spain and Pyrenees.

2.3. A fire database for the IP (1980–2005)

We have considered BA data from fire occurrences over Portugal and Spain in order to produce a joint database for Iberia (and Baleares islands) as these European countries have long and homogeneous time series of fire data since 1980 (Barbosa *et al*., 2007; JRC, 2010). Portuguese fire data was obtained from the Autoridade Florestal Nacional (AFN, 2011), while Spanish data was obtained from the Dirección General de Biodiversidad. Pereira *et al*. (2011a) provides a comprehensive description of the history and characteristics of the Portuguese fire database for the 1980–2005 period. It was found that during the 1980s the minimum amount of BA was 0.1 ha (i.e. all fires smaller than 0.1 ha where not considered) and this threshold has been reduced subsequently, being currently in the order of 0.0001 ha. These changes are important as they introduce artificial trends on the number of fires time series and, consequently, on the average BA per fire. However, the contribution of all these very small fires does not change significantly the annual (or summer) BA (Pereira *et al*., 2011a). The Spanish database has also been analysed in the recent past (Moreno *et al*., 1998) and has been shown to be increasingly reliable (Carracedo Martín *et al*., 2009). Although available since 1976, two northern Spanish regions present significant amounts of missing data: Álava (1980–1984) and Navarra (1980–1984 and 1994–2002). These missing values were replaced by the long-term monthly mean BA in each of these regions, a procedure with very limited impact as these regions present mean annual BA values below 0.5 and 0.1% of their total areas, respectively.

This new Iberian database refers to the common period of 1980–2005 for both countries BA time series. Monthly BA time series are spatially disaggregated on similar size administrative regions (AR): in the case of Portugal by *Distritos*, and in the case of Spain by *Provincias* (Figure 4). For both countries, we have

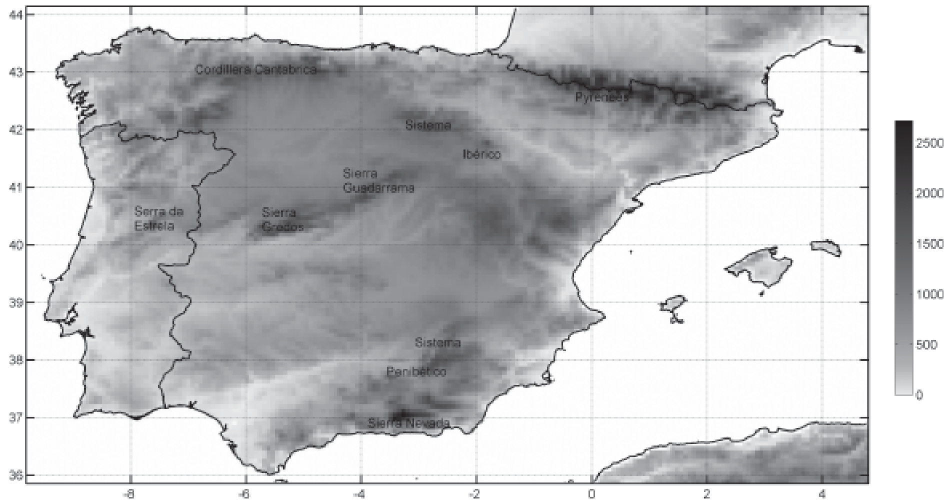

Figure 2. Main topographical features of Iberian Peninsula, including mountain ranges and the political borders of both Iberian countries.

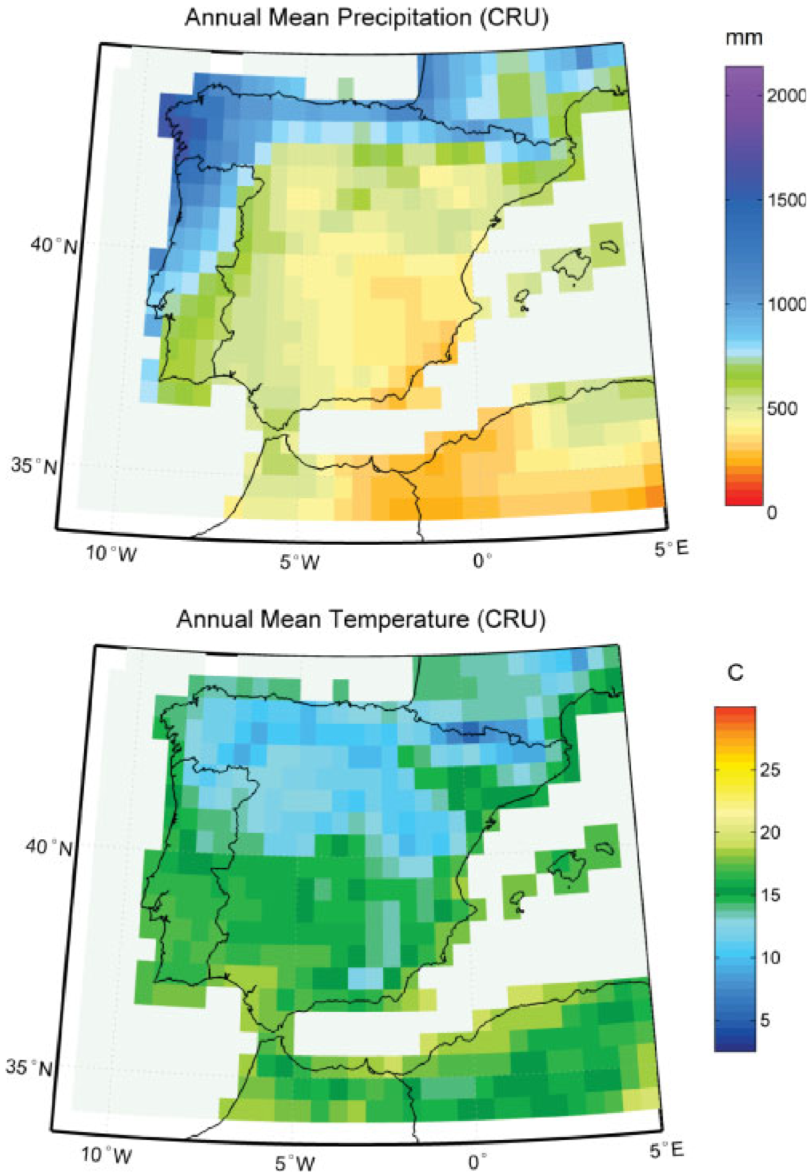

Figure 3. Annual mean precipitation (top panel) and temperature (bottom panel) over Iberia for the 1980–2005 period, based on the high resolution dataset from the Climatic Research Unit (CRU) gridded dataset version TS3.0 (Mitchell and Jones, 2005).

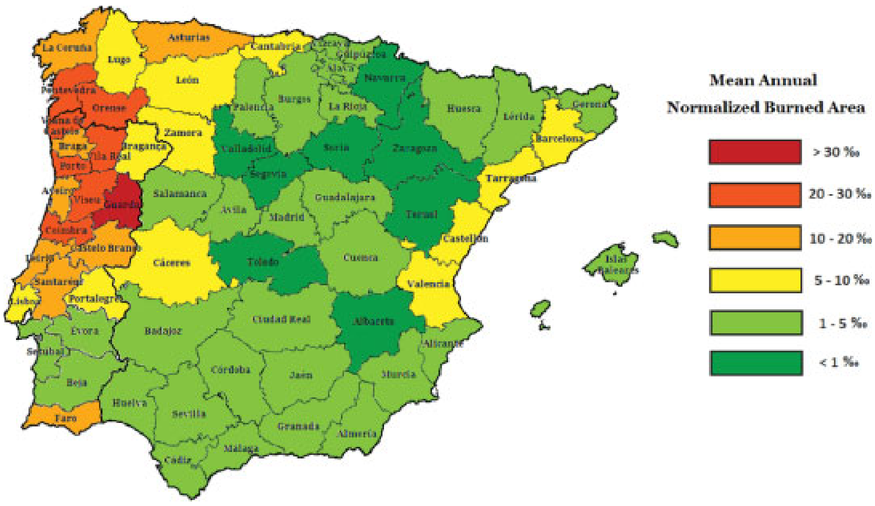

Figure 4. Annual mean normalized burnt area (NBA) in the Iberian Peninsula for the 1980–2005 period (in permillage) on a regional basis, composed by the Portuguese (*Distritos*) and Spanish (*Provincias*) administrative regions (AR). The NBA is defined as the quotient between the amount of burned area (BA) in each of the AR and the area of that AR.

not used the BA observed in the Portuguese (Azores and Madeira for Portugal) and Spanish (Canary Islands) Atlantic archipelagos and, in the end, the database spatial domain covers a total of 66 AR: 18 in Portugal and 48 in Spain. As these AR have considerably different surface areas, it is not appropriate (for some purposes) to compare BA directly among them. To circumvent this problem, besides the absolute amount of BA per region we have also computed the monthly Normalized Burnt Area (NBA) for each AR, defined as the quotient between the amount of BA in each of the AR and the area of the corresponding AR. Since both these quantities have the same units, NBA will be presented in permillage. After this normalization procedure, the monthly and annual averages for the NBA dataset for each of the 66 AR were computed. Annual mean NBA values for all the 66 Iberian AR (Figure 4) reveal that the northwestern sector of Iberia, which includes central and northern Portugal and northwestern Spain (mostly Galicia and Asturias), is the most affected by wildfires, presenting averaged values higher than 20‰. The eastern and northeastern coastal areas of Spain (Catalunya and Valencia) are also moderately prone to wildfires (with NBA values between 10 and 20‰ of their total area). Most central and southern regions of Iberia (except the Algarve region) do not present significant amounts of annual NBA (below 5‰). However, it must be acknowledged that this analysis based on NBA values still presents a caveat, since the normalization with the total districts/province's surface area can induce a bias in the final results. In fact, ARs have different percentage of vegetated area and some ARs (e.g. Zaragoza in Spain or Algarve in Portugal) have a relatively small fraction of forestry surface, albeit highly concentrated in smaller mountainous sub-domains.

2.4. Vegetation dynamics in IP

NDVI monthly anomalies, with 8 km of spatial resolution, were obtained from the Global Inventory Modelling and Mapping Studies (GIMMS) dataset and correspond to the most complete and longest remote sensing dataset, covering the entire Mediterranean region, for the period 1982–2006 (Tucker *et al*., 2005).

Vegetation over the Mediterranean basin and specifically over the IP presents high diversity, resulting from climate, landscape, topography and other factors. To identify the main types of vegetation cover over Iberia, a cluster analysis of the monthly means of NDVI for the entire period was performed, using four clusters depicted in Figure 5 (top panel). These four clusters correspond approximately to the four main vegetation types over Iberia. The cluster represented in blue coincides almost exactly with the extension of the Eurosiberian phytoclimatic region across the IP (Rivas Martínez, 1987), characterized by low temperature and abundant and regular water supply. The spatial domain of this cluster broadly corresponds to regions included within the 800 mm isohyets, a threshold traditionally used to delimitate the 'green' (Csa, Csb and Cfb on Köppen Classification) Iberia from the drier Mediterranean regions. Its natural vegetation is alpine or deciduous forest, currently composed by pastures and secondary forest plantations (eucalyptus, pinus) that presents high activity throughout the entire year, with a maximum in spring (Figure 5, bottom panel), but deeply dependent on water resources. The cluster represented in yellow groups two different landscapes: (1) an ecotone transition (semideciduous forest) through the southern piedmont of the mountains of northern Spain (Cordillera Cantábrica and Pyrenees) and (2) the typical Mediterranean sclerophyllous forests (sometimes grazed—*dehesas*) across most of the mountains and plateaus of central and southern IP. This cluster comprehends large areas of annual (mainly not irrigated) crops characterized by the overall highest NDVI values during the entire seasonal cycle and with a maximum in summer (Figure 5, bottom panel). The cluster represented in green corresponds to the central areas of both Mesetas

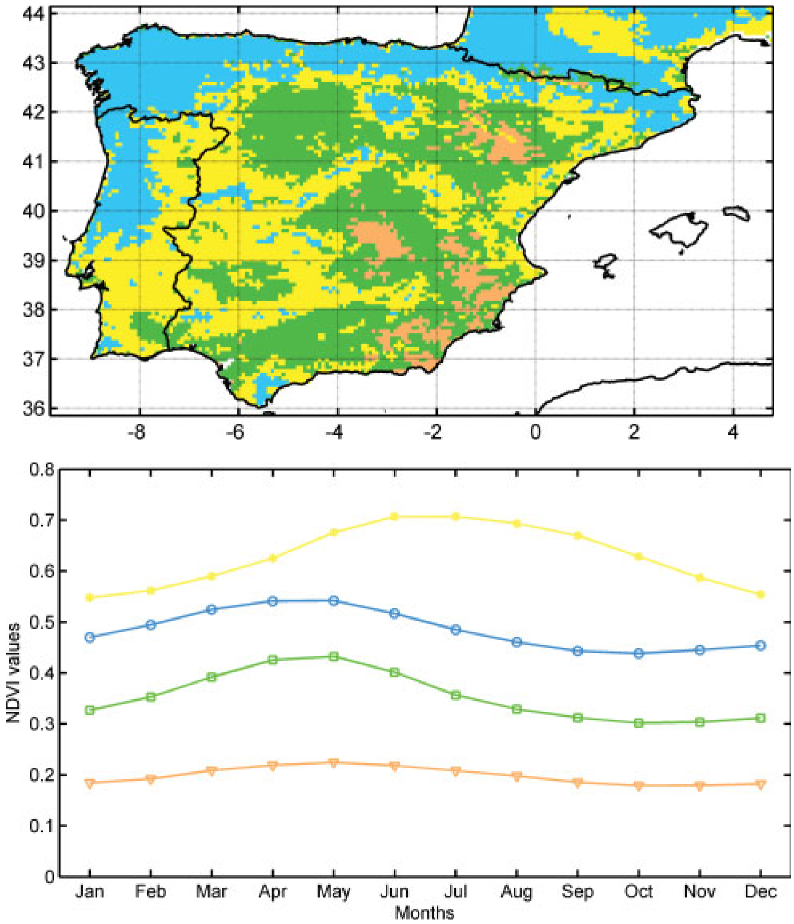

Figure 5. Spatial distribution of the four clusters of NDVI obtained with GIMMS dataset for the period 1982–2006 (top panel) and the annual cycles of monthly NDVI that characterize the centroids of the four identified clusters (bottom panel).

and the Ebro and Guadalquivir basins, besides some spots along the Mediterranean coast of the IP. These regions seem to share in common a dominant (in some cases irrigated) agricultural activity. This cluster presents lower vegetation activity (lower NDVI values) than the previous two, with a maximum during spring, due to the availability of water during this season—a feature usually associated to scrublands. The last cluster, plotted in orange presents very low values of NDVI over the entire year (Figure 5, bottom panel), corresponding to the typical regions characterized with very sparse vegetation or even bare soil. It resembles the most arid environments of the IP, (Ebro valley, La Mancha, SE Spain), characterized by a stepped-like vegetation (when/where there is vegetation) as consequence of high summer temperatures and scarce annual precipitation (below 400 mm).

3. Regionalizing fire regimes in the IP

In the previous section, BA datasets from both Iberian countries were merged and a preliminary analysis was performed to obtain the spatially NBA dataset. These procedures were required before applying the methodology that aims to objectively identify the most important fire regimes within the IP. For this purpose, we performed a Cluster Analysis, using the K-means algorithm, on the NBA in the 66 considered AR disposed in a T-mode matrix. We decided to retain four statistically significant clusters (Figure 6), since the obtained pattern presents regions with higher spatial homogeneity in comparison with spatial configurations attained with a higher number of clusters, as it will be shown later. Furthermore, this result is concordant with the defined regionalization defined by Rasilla *et al*. (2010), also with four sub-regions, and presents a strong resemblance with the operative regionalization used by the Spanish authorities (MARN, 2011).

We must stress that the configuration presented in Figure 6 does not correspond exactly to the cluster analysis output. A small number (six) of neighbouring Spanish AR were swapped in order to define a more spatially homogeneous set of four clusters. These AR are identified with a small circle in Figure 6 and these changes were considered valid to perform, as they refer to ARs with low mean annual values of NBA (usually below 5‰, as shown in Figure 2). Taking into account the geographical location of the resulting four clusters, these were named respectively: northwest (NW_CLU); southwest (SW_CLU); north (N_CLU) and east (E_CLU).

Figure 6. Spatial extension of the four clusters obtained for the normalized burnt area in the 66 administrative regions (AR) of Iberia. Black dots identify the AR that originally belonged to different clusters but were reassembled to increase spatial coherence. Legend: blue—northwest cluster (NW_CLU); yellow—southwest cluster (SW_CLU); green—east cluster (E_CLU); and, magenta—north cluster (N_CLU).

The NW_CLU aggregates the northern half of Portugal and the extreme northwest of Spain, including most of the ARs with higher values of mean annual BA. The SW_CLU represents the southern and interior areas of Portugal (including Guarda, the AR with the highest mean annual BA), as well as many ARs of central and southwestern of Spain. The coastal and pre-coastal ARs east of Gibraltar correspond to the E_CLU (including the Baleares islands). Finally, the N_CLU corresponds to the regions located in the mountainous sectors of northern Spain (including Asturias, Cantabria and the Basque Country).

The main purpose of applying this Cluster Analysis was to identify the Iberian regions with similar temporal variability. In fact, all four regions shown in Figure 6 have a clear annual cycle of the mean monthly NBA, with the main differences being found in the amplitude and timing of the maximum values (Figure 7, top panel). The NW_CLU clearly represents the AR more affected by wildfires, with nearly 10‰ of the total area being burned, on average, every August. The mean peaks of the other clusters are all below 3‰ of their area. Interestingly, two of the identified clusters present two maxima, with the NW_CLU revealing a much larger peak in August and a smaller one in March. The northernmost areas of Iberia are concentrated in the N_CLU that presents two peaks, both comparable in magnitude (relatively small although), in spring (March) and late summer (September). Finally, the two remaining clusters are characterized by a single summer maximum in July and in August in the case of E_CLU and SW_CLU cluster, respectively.

The boxplots of monthly NBA (Figure 7, bottom panel) also help to characterize the intra-annual evolution of location (median), dispersion (IQR), range (maximum minus minimum) and asymmetry characteristics of the NBA distribution in each cluster. In general, the annual evolution of location, range and dispersion statistics present a similar annual cycle of the mean monthly NBA, with similar ratios between the amplitude and timing of the maximum values. Only a few number of discrepancies of that annual cycle are worth to mention, namely: (1) the median in July is similar than the one in August for the SW_CLU; (2) secondary maximum values of the median (the first peak, in late winter and beginning of spring) occurs in February, and not in March as for the mean monthly NBA, and is about 75% higher for N_CLU than for the NW_CLU. The temporal evolution of the quartile skewness reveals that monthly NBA distribution is essentially positive skewed, (with a few exceptions), presenting high intra-annual variability and smaller values for the western clusters (NW_CLU and SW_CLU) during summer/fire season and the opposite behaviour for the other two cluster (except for N_CLU in June and July).

The spatial patterns of the four NBA clusters and the correspondent fire regimes are also in good agreement with topography (Figure 2), climate conditions (Figure 3) and vegetation dynamics (Figure 5). Actually, the NW_-CLU corresponds to the relatively low land coastal areas in the north western Atlantic coast of the IP while the SW_CLU includes the lowlands in the SW region of the IP and the large Meseta Central (Figure 2). This central

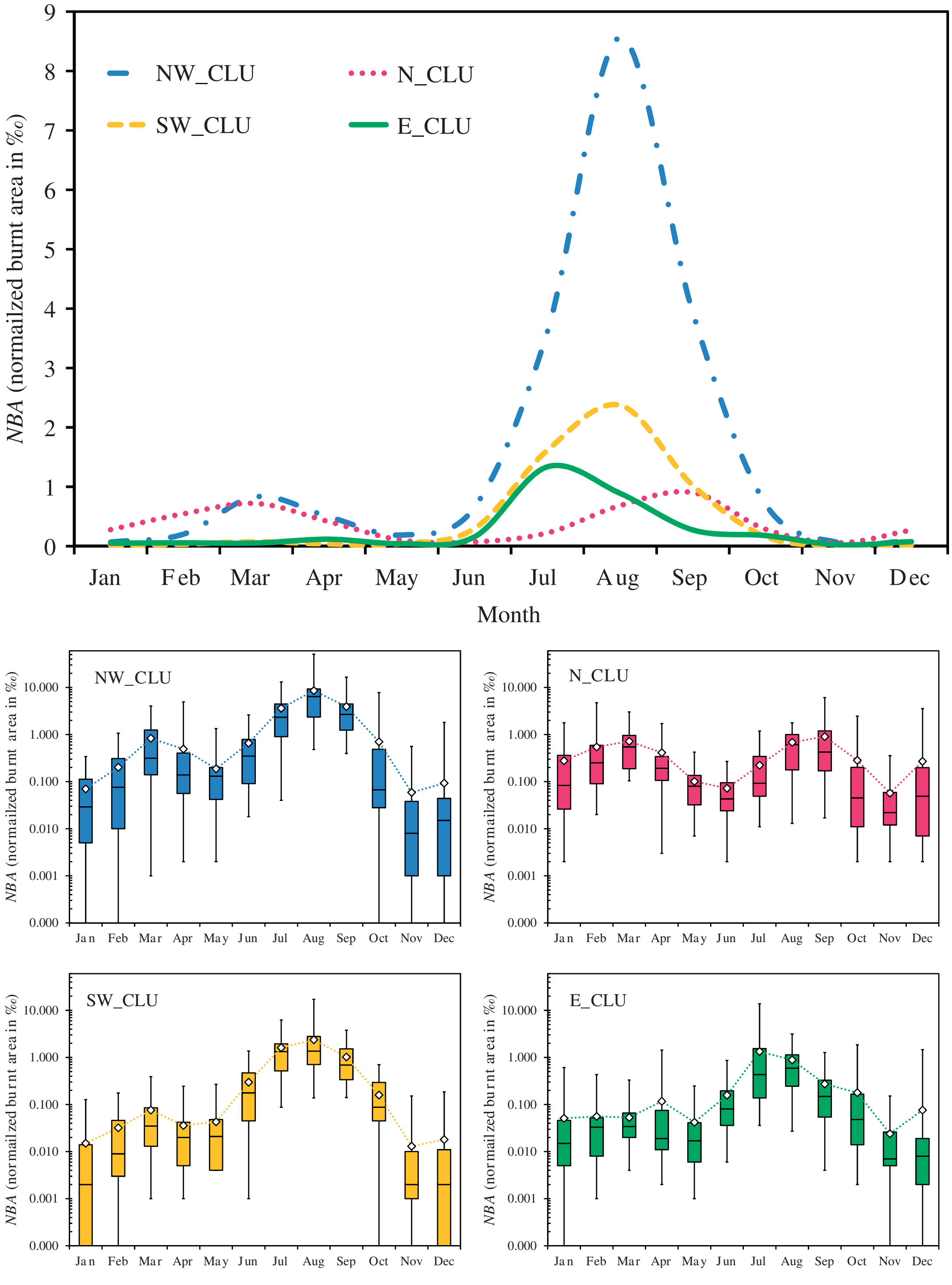

Figure 7. Mean monthly time series (above) and boxplots (below) of normalized burned area (NBA—defined as the quotient between the amount of BA in each of the AR and the area of the AR, both in hectares), in the 1980–2005 period for the four NBA clusters shown in Figure 6 and using the same colour palette as in Figure 6.

and high plateau is surrounded by the Cantabrian Mountains (Cordillera Cantábrica) region that corresponds to the N_CLU and by the Ibérico and Penibético systems which delimits the E_CLU along with the Mediterranean coasts of Iberia (Figure 2). The climate of these four regions is also very different and helps to put into perspective some characteristics of their vegetation cover (Figure 5) and fire regime (Figures 6 and 7), namely the yearly amount of BA, the fire season duration and peak timing. Besides mean annual precipitation and temperature spatial patterns (Figure 3), the climate characterization of IP requires some knowledge on the intra-annual temporal and spatial variability, e.g. the typical conditions during the cold (winter) and warm (summer) seasons. According to the Iberian Climate Atlas (AEMET-IM, 2011), the winter and summer mean precipitation and temperature conditions in these four regions are as follows. The N_CLU and NW_CLU are clearly the wettest clusters (Figure 3) and that implies a compatible vegetation cover (blue colour in Figure 5). However, these two fire regime clusters diverge in terms of their annual temperature. Thus the N_CLU is characterized by wet and cold winters and by humid and fresh summers (AEMET-IM, 2011) which supports the fact that this is one of the regions less affected by wildfires (Figure 7). On the contrary, the climate (and associated vegetation cover) of the NW_CLU is the most appropriate for the occurrence and development of fires during the dry and hot summer season because these conditions follow the humid and mild winters (AEMET-IM, 2011). Figure 4 confirms (through the NBA spatial pattern) that these are the most favourable conditions for the occurrence of wildfires, allowing the existence and development of live fuels during the wet and mild season (winter and early spring) and the vegetation stress conditions during hot and dry seasons (late spring and summer). The two remaining clusters are associated with drier climates and mostly non-forested vegetation cover (Figure 5). Climate in the E_CLU is characterized by dry winters (with small exceptions in the Pyrenees) and by even dryer and hot summers (AEMET-IM, 2011). The existence of meteorological favourable dry and hot conditions that promote the occurrence of wildfires during the summer season is not supported by the existence of appropriate vegetation cover, because this arid environment seriously limits the fuels availability (Pausas and Fernández-Muñoz, 2011). In fact, the cluster represented in light green in Figure 5 corresponds to areas with extremely low NDVI values (Figure 5 bottom panel), which is a typical of bare soil or very sparse vegetation. In this situation, and even with meteorological conditions that favour fire ignition, there is a very low probability of having large fire events/activity. The SW_CLU region is characterized by extremely hot and dry summers but has a less homogeneous climate conditions during the mild and humid winter, with a gradient of both temperature and precipitation in the NE–SW direction. Vegetation cover in this area is dominated by sparse vegetation, with a maximum of vegetation activity in spring (green in Figure 5) and rain feed crops that mature in late spring and early summer (yellow in Figure 5).

4. Modelling the inter-annual variability of BA

4.1. The influence of previous climatic conditions

Although fire ignition is most of the times dependent on human activities (either by accident, negligence or intentionally) the probability of a fire ignition to grow into a large and severe event is highly dependent on the existence of the appropriate meteorological conditions (Pyne *et al*., 1996). Nevertheless, besides the meteorological conditions at the time of the fire ignition, previous climatic conditions play a very important role, as they are directly related to soil dryness and vegetation hydric stress during the fire season. Some relations have been identified in specific regions, such as over Portugal (Pereira *et al*., 2005; Trigo *et al*., 2006; Hoinka *et al*., 2009), where besides the usual link between summer fires and previous months dryness and warmness, wet months in mid-spring also reveal to be related to more BA in the fire season, as they promote the growth of vegetation which may act as fuel when fires are triggered. However, the exact nature of these relations between appropriate previous climatic conditions and BA varies from area to area and from season to season.

Taking into account the similar fire regime within each of the four clusters defined in Section 3, and in order to assess the role of previous climatic conditions on the inter-annual variability of BA, values of the Pearson correlation coefficient (R) between BA time series (not NBA) in each cluster and relevant antecedent climatic variables (precipitation and temperature) over the region were computed as a procedure to reduce the dimensionality and pre select predictors with greater potential. It should be noted that BA and NBA present similar seasonal cycles, but the need of using NBA in the previous section was related with the necessity of performing the cluster analysis using a more appropriate fire statistic that facilitates the comparison among clusters (Figure 6) and the corresponding seasonal evolution (Figure 7). The CRU-TS3.0 mean monthly precipitation and temperature dataset were used to compute mean spatial average series for each cluster domain, i.e. selecting all grid-points coincident with the area of the spatial configuration associated to each cluster depicted in Figure 6. Instead of using absolute values of climate variables to test their usefulness as predictors, the developed models will rely on monthly anomalies time series, computed by removing the long-term mean from the spatially averaged monthly time series in each cluster. This procedure does not change the Pearson correlation coefficient but effectively removes the annual cycle in order to retain time series of temperature and precipitation anomalies in each of the four Iberian fire cluster regions. As it was seen in Section 2.3, each cluster presents a peak of BA at different timings of the year, or even two peaks during different seasons. Bearing this in mind, eventual links between BA in those peak months and the time series of anomalous meteorological fields were checked from 2 to 7 months in advance. For example, in order to study potential previous climatic influences for the August BA observed in the NW_CLU, we tested the existence and statistical significance of Pearson correlation coefficients between this variable and the meteorological monthly time series from January to June. The highest values of R between each variable and each BA cluster for the considered months are presented in Table II. No results are presented for N_CLU, as none of the tested R values were statistically significant at the 5% level.

These previous climatological variables can integrate the pool of potential predictors and be used as predictors in statistical hindcast models of BA (hereafter CLI_models). The purpose of such relative crude models is not to

Table II. Summary of the highest correlation coefficients obtained between observed BA series for NW_CLU, SW_CLU and E_CLU clusters, in July and August (the months of highest values of BA, for these clusters) and the monthly mean temperature (T) and precipitation (P) time series in the seven previous months.

Month	Cluster		
	NW_CLU	SW_CLU	E_CLU
July	May T (+0.55)	May T (+0.41)	May T (+0.44)
	May P (−0.48)	May P (−0.49)	
August	May T (+0.45)	May T (+0.36)	Jan T (−0.51)
	May P (−0.47)	Jun P (−0.58)	

Correlations presented here are statistically significant at the 5% level.

develop quasi-operational summer BA models for each region but simply to assess the contribution potential to BA predictability based on pre-disposing climatic conditions. For this purpose, a forward stepwise regression was applied to the temperature and precipitation monthly anomaly time series, in order to reproduce BA monthly time series for each cluster. This method tests all possible predictors and selects the one that explain more variance of the predictand (BA), and then looks for the most relevant remaining predictors that surpass a pre-defined level of statistical significance (here defined as the 5% level). For the three clusters that present the main peaks of BA during summer months (e.g. NW_CLU, SW_CLU and E_CLU), three models were developed and tested to simulate the BA in (1) July, (2) August and (3) July + August. In the case of the N_CLU, which presents two BA peaks (in March and September) models for those two specific periods were tested. Eventual problems related to over fitting were mitigated by employing the always stringent cross-validation scheme, usually known as the leave-one-out scheme (Wilks, 2006). In this approach, a single observation from the original sample is retained for validation, and the remaining observations are used as training data. Besides, all models were forced to retain four predictors at most. All these precautionary steps were taken because the time series are only 22 years long (1980–2001, since WTC data was only available until 2001), and this implies that using a large number of predictors could be misleading.

Values of the Pearson correlation coefficient between observed and modelled (with CLI_models) BA time series in each cluster for all these cases are summarized in Figure 8. These results reveal that the role played by previous climatic conditions diverges for the different clusters (Figure 8). Winter and spring climatic conditions play an important role for the very active NW areas. Although we find modest R values between observed and modelled BA time series for both July and August BA, R increases (to about 0.5) and surpasses the statistical significance level for the July + August BA. Results obtained for the SW_CLU and E_CLU show that pre-conditioning climatic predictors appear to be useful for estimating the BA in August, but not as clearly as for July. For the N_CLU (not presented in Figure 8) we did not obtain statistical models for any of the considered months, as none of the possible previous climatic predictors were retained by selection method. This result means that BA in March and September are not clearly associated to remote previous anomalous climatic conditions in this cluster.

4.2. Links with WT frequencies

The objective of this section (and the main motivation of this work) is to assess the performance of the WTC from the COST733 catalogue over the Iberia region in what respects to their capacity of discriminating weather conditions that favour the occurrence of large values of BA. The focus will be on the conditions in the months of occurrence of fires, or at the most, the previous one.

Every single day in the study period (1980–2001) was classified in one of the nine specific categories of each WTC. To make this information operational, monthly frequencies of each category were computed for all WTC used in this study. Then, following the same approach used for the CLI_models, correlation analysis was used as a pre-selection procedure to identify the WTC most related with BA time series in each cluster.

The highest values of the Pearson correlation coefficients between time series of BA in each cluster and monthly frequencies of each WT are presented in Table III. Taking into account that several classifications present few significant values of R with BA time series, the remaining of the analysis was restricted to the five most promising WTC, namely: LUNDC09, NNWC09, PCACAC09, PETISCOC09 and WLKC09. This restriction is objective as it is solely based on the magnitude of the statistical relationships observed between the BA in each cluster and the WT frequencies for certain classifications.

The BA statistical models based on time series of monthly frequencies of each WT (WTC_models) for each of the four clusters were developed following the same approach used to developed the CLI_models (i.e. a forward stepwise predictor selection procedure, limited to retain at most four predictors and using cross-validation). This approach will enable us to distinguish which of the considered five classifications tend to better reproduce the appropriate classes of weather conditions that are related to months with high values of BA in Iberia. At this stage, the climatic information from previous seasons was neglected and we are only focused on the WTC predictors. Aiming to develop the best BA models and to assess the possible added value of incorporating both types of information (rather than using only WTC or only previous meteorological data alone) a third type of hindcast models (TOT_models) was developed. In this case, both classes of variables are integrated in the potential pool of predictors used to construct the models and the remaining exact same procedures as in the previous models are carried out.

Despite the simplifications assumed previously (number of clusters retained, months to analyse and WTC kept

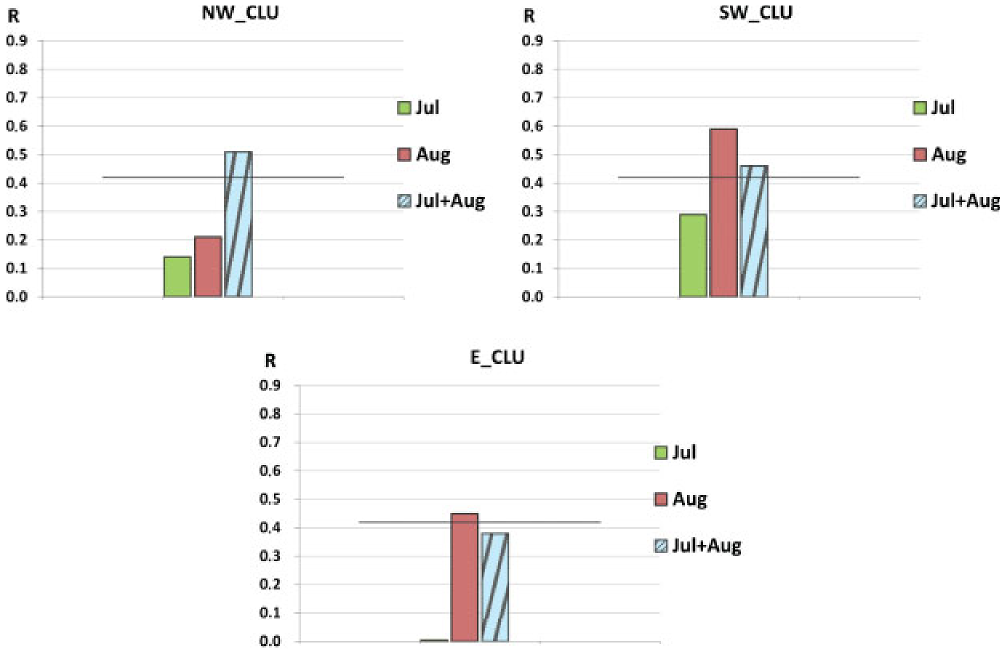

Figure 8. Values of the Pearson correlation coefficient (*R*) between observed and modelled time series of monthly burnt areas in the northwest (NW_CLU), southwest (SW_CLU) and east (E_CLU) clusters of normalized burned area (NBA). The models use climatic variables as predictors (CLI_models) in hindcast mode and the grey line represents statistical significance at the 5% level.

Table III. As Table II, but for the best correlation coefficients between observed BA series (for each of the four clusters) in the considered months and the COST733 catalogue WTCs.

	Cluster			
Month	NW_CLU	SW_CLU	E_CLU	N_CLU
March	–	–	–	Feb WLKC09-5 (+0.64)
	–	–	–	Mar CKMEANS09-5 (+0.58)
	–	–	–	Mar SANDRASC09-6 (+0.56)
July	Jul PCACAC09-8 (+0.66)	Jun SANDRAC09-4 (−0.55)	Jul WLKC09-4 (+0.94)	–
	Jul PETISCOC09-5 (+0.61)	Jul KHC09-2 (+0.54)	Jul PETISCOC09-7 (+0.64)	–
	Jun NNWC09-4 (−0.58)	Jun PETISCOC09-3 (−0.53)	Jun LUNDC09-2 +0.57	–
August	Aug LUNDC09-4 (+0.70)	Aug PCACAC09-4 (+0.72)	Aug PCACAC09-4 (+0.62)	–
	Jul NNWC09-3 (+0.65)	Aug PETISCOC09-8 +(0.59)	Jul LUNDC09-6 (+0.56)	–
	Aug PETISCOC09-2 (+0.68)	Aug LUNDC09-2 (+0.56)	Jul WLKC09-4 (+0.53)	–
September	–	–	–	Sep KHC09-4 (+0.62)
	–	–	–	Sep SANDRASC09-1 (−0.52)
	–	–	–	Sep LUNDC09-2 (+0.51)

Only WTCs catalogues whose weather types frequencies show significant correlation (at the 5% level) with Burned Area series are displayed. WTCs names and characteristics are described in Table I.

in the analysis) a large number of models were developed as a result of retaining five different WTC, four clusters and at least 2 months to predict. Thus, for the sake of simplicity, we have summarized results in a simpler display (Figure 9) of the Pearson correlation coefficient value between the observed and simulated BA time series for all the tested models. For each cluster, classification and month, the plot includes two bars, the lighter one (at front) representing the result achieved with the WTC_models and the darker one (at back) representing the TOT_models, that incorporates both WTC and previous climatic conditions as possible predictors. For this reason, the darker bar is only visible in the cases where TOT_models produce better models than WTC_models. This representation simplifies the comparison between the performances of all models, the assessment of the ability of each WTC to describe the BA variability, and the understanding to what extent (and for each cluster) previous climatic data improves the results obtained from WTC predictors-based models. Several major results emerge from the analysis of Figure 9, namely:

(1) Higher values of *R* were obtained for both western sectors of Iberia (NW_CLU and SW_CLU), fact which suggest that it will be more susceptible to reconstruct BA series using this type of statistical models and predictors in these regions;
(2) In general, TOT_models present higher performance than the corresponding WTC_models and CLI_-models, which means that the use of both WTC

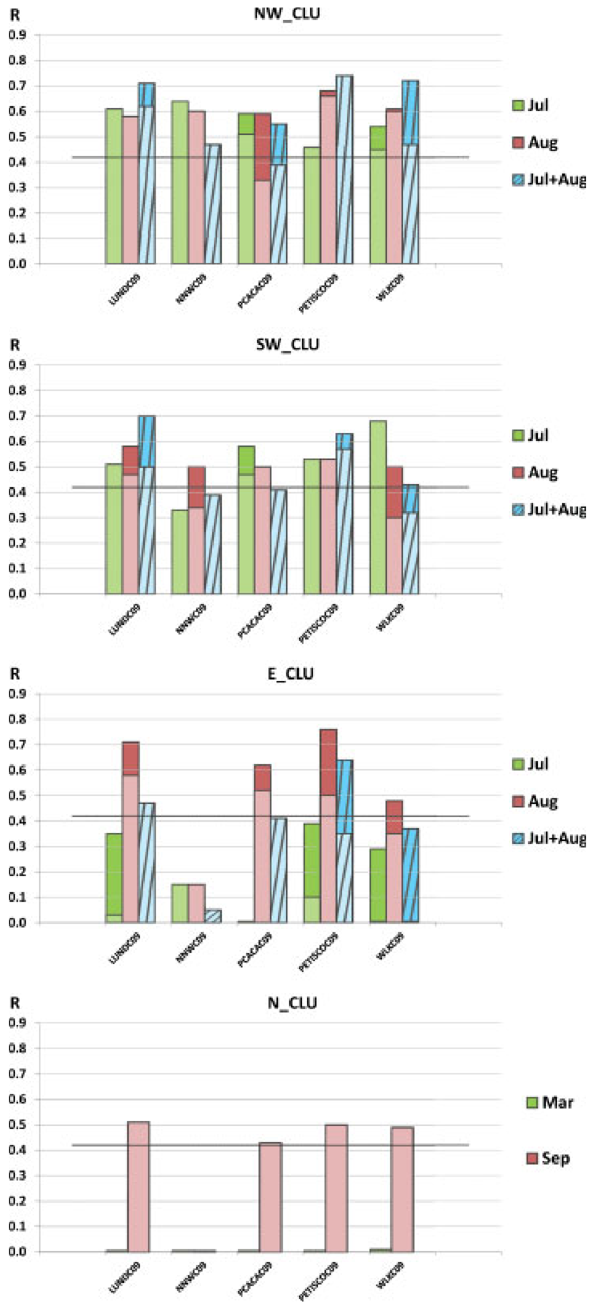

Figure 9. Values of the Pearson correlation coefficient (R) between observed and modelled time series of monthly burnt areas in the northwest (NW_CLU), southwest (SW_CLU), east (E_CLU) and north (N_CLU) clusters of normalized burned area (NBA), using as predictors in light shaded bars just the WTC (WTC_models) and in dark shaded bars using both WTC and climatic variables (TOT_models). Dark shaded bars are only visible whenever TOT_models present better performance than WTC_models.

and previous meteorological conditions as predictors provides an added value in the vast majority of cases.

For example, in the NW_CLU the value of the Pearson correlation coefficient for most of the models surpasses the 5% significance line. In this region, good results are found for both July and August BA time series. Nevertheless, the July + August models present the best performance in NW_CLU (in three of the five tested WTC). In the SW_CLU, the month which models are more able to reproduce BA variance varies according to the considered classifications. TOT_models present better results than CLI_models but, the main feature is that the use of WTC as predictors improves substantially the reconstruction of BA series in July, which was quite modest when using climatological information only. Results for August are slightly better but not statistically significant, showing that the wildfire climatology in SW_CLU and this month is essentially related to years with propitious conditions cumulated in previous months. In the E_CLU August BA models tend to present better results (in four of the five classifications), while July models fall below the 5% level of statistical significance. As no satisfying CLI_models were obtained for the N_CLU, we find the best TOT_models for this region to be equivalent to the best WTC_models. In this cluster, statistically significant BA models were only obtained for September.

The procedure for statistical validation of the models is not based solely on the Pearson correlation coefficient. The Mean Absolute Error (MAE) has also been computed for all models presented here. We found a good agreement between the results obtained with the two measures of goodness of fit, for example, the higher the correlation coefficients, the lower the MAE between observed and modelled series (not shown).

Finally, it is not an easy task to rank the best WTC classifications for any cluster as it depends on the target month. However, PETISCOC09, WLKC09 and LUNDC09 can be considered the most satisfying for NW_CLU. It is also rather difficult to select the best performing classifications in the SW_CLU, but LUNDC09, PCACAC09 and PETISCOC09 may be considered to be those that provide the best results on the overall. The PETISCOC09, LUNDC09 and PCACAC09 classifications present the best results for the E_CLU. For the N_CLU, the region with the poorest model performance, statistically significant models were only found for September and with the WLKC09, LUNDC09 and PETISCOC09 WTC.

5. Synoptic analysis

It was shown in the previous section that the statistical models (TOT_models) which combine prior climatic information (temperature and precipitation anomalies) and non-lagged WTC series usually present better performance, when reconstructing the inter-annual variability of BA series of Iberia, particularly in the most affected western sectors. To have a further understanding on the role that the latter predictors have on these models, the mean sea level pressure anomaly composite fields of the retained classes for each model in each cluster were analysed in more detail. Once again, taking into account the considerably large number of models and predictors, we opted to summarize results in a concise way (Figure

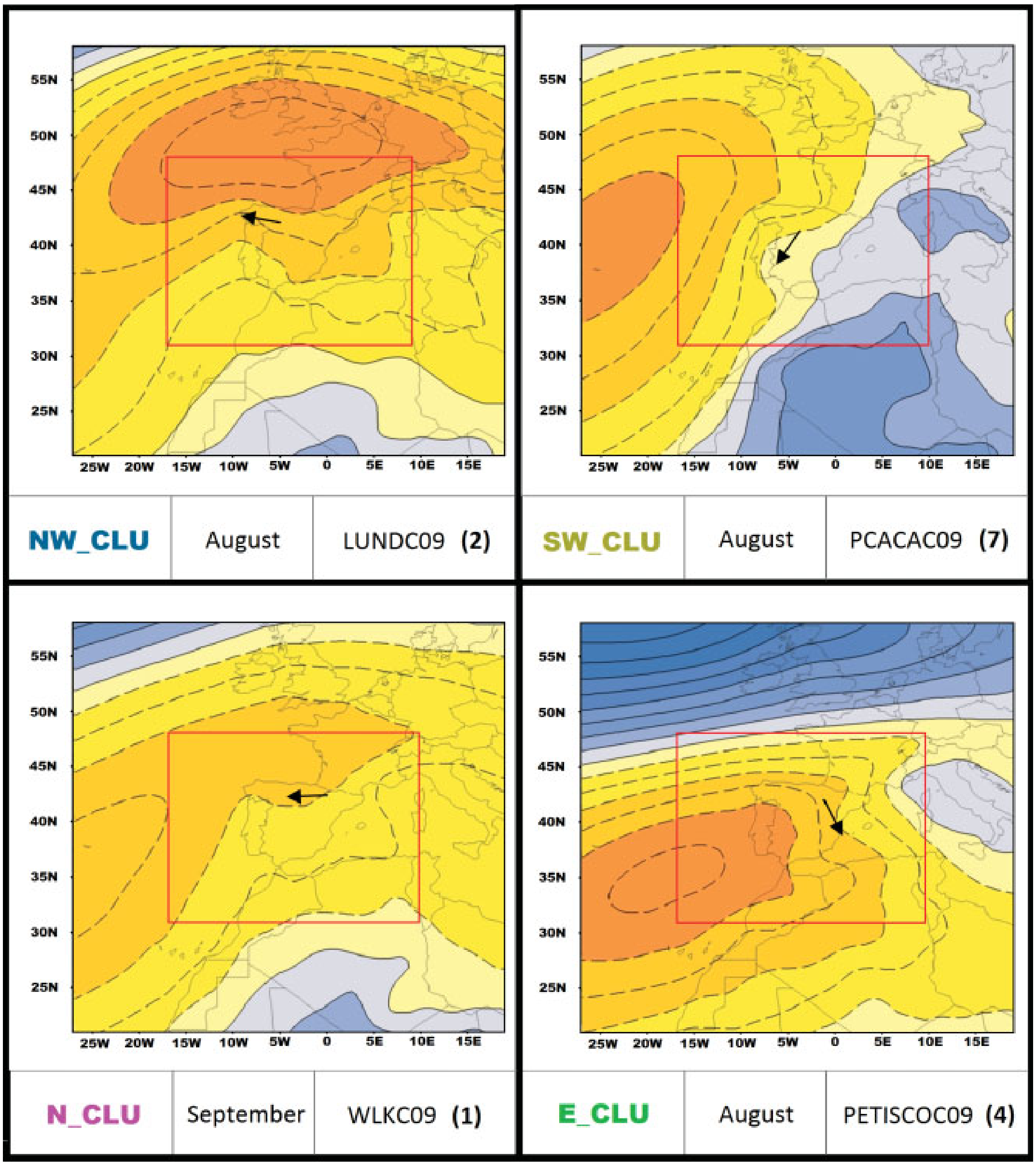

Figure 10. Composites of surface pressure anomaly (2 hPa interval isolines, with dashed lines corresponding to positive anomalies) in months (lower central box) with higher values of BA in each cluster (lower left box) and for specific weather type classifications (lower right box) that are highly correlated with BA series. Black arrows represent the anomalous wind direction in the respective cluster, revealing the origin and characteristics of the advected air masses.

10). The shown composites characterize the typical synoptic configurations driving the main surface wind flow and heat advection that predominate during months with very high values of BA in each cluster.

The majority of the retained classes correspond to very similar synoptic configurations, independently of the WTC. In all cases, and as expected, the composites reveal atmospheric anti-cyclonic circulation patterns usually associated with stable and dry conditions. The obtained spatial configurations foster prevailing warm and dry winds from inland, rather than moist and cool air advection from the sea, either from the Atlantic Ocean or the Mediterranean. The main difference between the results for each cluster can be found in the location of the high-pressure region, which induces different directions of the surface wind flow. This type of circulation is clearly evident in the composites identified when analysing the most coastal confined clusters (NW, N and E). In the NW_CLU of Iberia, we find that hazardous fire seasons are related to WTs with prevailing winds from the east or southeast components. The composite of surface pressure anomaly obtained for the NW_CLU (Figure 10) is very similar to the spatial pattern obtained for the sea level air pressure composite for the 10% highest BA days in Portugal and the corresponding 10% anomaly for the 1980–2000 period (Pereira *et al*., 2005). The Mediterranean areas that compose the E_CLU present important fire episodes when the circulation is essentially from the northwest (Figure 10). Regions from the N_CLU have fire incidence with dominating southeast or east winds. In this latter case, and in particular for coastal areas, it is also quite evident that besides being an inland wind, it is also a downslope wind from the Cantabria mountainous region towards the coast, promoting additional drying and warming of air masses (Foehn effect). This is also the process that helps to understand the obtained results for the SW_CLU, where the anomalous composites point to prevailing winds from the north or northwest components. It should be reminded that in this cluster the highest values of the NBA are found for some ARs of Portugal (namely Guarda, Castelo Branco and Santarém), which are located downstream of the main orographic barriers in that domain. It should also be kept in mind that this cluster covers a much wider area of the peninsula (Figure 6), comprehending not only the referred Portuguese AR but also a large part of central and southern Spain, where NBA values are generally low. The wide structure of this cluster may partially explain the absence of a clearly inland originated wind patterns in the presented composite.

6. Discussion and conclusions

In this work, a comprehensive and novel wildfire database for the entire IP was established by merging information from both Iberian countries which is essential for studying spatial and temporal patterns of fire occurrences and BAs in Iberia. Moreover, such a database can be useful for further developments on a wider European context. Using this Iberian dataset, we were able to perform a cluster analysis to distinguish between different subregions with similar monthly fire regimes within the IP. This analysis proved to be efficient in discriminating four independent regions, as they differ both on the timing and intensity of the seasonal peaks of NBA. The analysis of the mean annual cycle of NBA clearly highlights the western half of Iberia as more prone to large wildfire activity, particularly in central and northern areas of Portugal and northwestern areas of Spain. Paradoxically, being neither the driest nor the most irregular region in terms of rainfall in the Iberian context, the exceptional high frequency of forest fires in NW_CLU denotes the relevant role of a combination of weather-climate, vegetation, geographical and human factors upon fire dynamics. Some northeastern areas of Spain also present relevant mean annual values of NBA, but generally at a smaller magnitude than the previous. This fact leads us to pay particular attention to the results and statistical models of westernmost clusters. In fact the relatively low values of NBA found in the other clusters reduce the possibility of a deeper analysis, and increase the danger of misinterpretation of the results and of the efficiency of the tested statistical models. As expected, the highest mean monthly values of NBA area were found in summer months, as higher temperatures and lower humidity foster appropriate conditions for fire ignitions. Westernmost regions tend to present higher mean values of NBA in the month of August, while areas along the Mediterranean basin present a peak in July. Northern areas (both N_CLU and NW_CLU) have two annual peaks, one in summer and one in late winter or early spring. While in the NW_CLU this cold season peak is almost insignificant when compared to the summer peak, in the regions around Cantabria and Asturias both peaks are comparable, although of relatively small amplitude.

The purpose of this study was to undertake an objective comparison between the ability of different WTC available from the COST733 project, in terms of discriminating the synoptic conditions that favour large wildfire activity in Iberia. This was done for classifications with nine classes, and the procedure includes: (1) correlation analysis between monthly BA and the monthly frequencies of each WT, used as a predictor pre-selection tool; followed by a (2) regression analysis to derive statistical hindcast models of monthly BA (WTC_models) for all the clusters, and testing their ability to reproduce the inter-annual variability of the observed series. Similar approaches were adopted in previous studies in many different areas (Pereira *et al*, 2011b). To avoid multicolinearity, in ecological modelling studies it is common to adopt two criteria: (i) pairwise correlation using Spearman's or Pearson's R correlation coefficient between predictors lower than given thresholds, e.g. 0.7 (Elith *et al*., 2006; Wisz and Guisan, 2009); and (ii) Generalized Variance Inflation Factor (GVIF) lower than 5 (Neter *et al*., 1996). In this study, R between every pair of predictors was less or equal 0.5, while GVIF computed with SAM (Rangel *et al*., 2010) was always less than 2.

To contextualize this modelling approach, models based solely on climatic information from the pre-fire season (CLI_models) were initially developed in order to compare the results obtained with both models. This procedure enabled us to distinguish between regions of the IP where pre-fire season temperature and precipitation play an important role on the magnitude of the fire season, and areas where this information has relatively low predictability value. Summer large wildfire activity is highly related to previous climatic conditions in the northwestern sector of Iberia. This result is in agreement with the previous work by Pereira *et al*. (2005), which states that spring conditions are determinant on the propensity for large values of BA during the warmest months, mainly by (1) the level of dryness imposed by previous climatic conditions and (2) the rate of vegetation growth in spring, which leads to an additional fuel accumulation during summer extreme warm episodes. We have attempted to use other variables related with the pre-fire season climate, such as the Palmer Drought Standardized Index (PDSI), however, this has not improved the quality of the models, being rejected during the stepwise regression procedure.

A third set of statistical models was computed using both referred types of information as potential predictors (TOT_models). This was done in order to see the added value of merging WTC and previous climatic conditions information. It should be stressed that all models were cross-validated to avoid over fitting. As expected, in almost all cases, best results were obtained for the TOT_models. The larger improvements were found for months and clusters where CLI_models results were poorer. In northwestern areas, previous climatic conditions information was already sufficient to produce quite robust hindcast models. The difference in the predictability of BA amid different areas is quite evident in the correlation coefficient and MAE between hindcast and observed time series. Overall, we found the most robust statistical hindcast models for the NW_CLU (the one with the highest mean annual values of NBA), and the worst performing models in the N_CLU (the one with the lowest mean annual values of NBA and with two similar sized peaks in March and September).

There is no clear or consensual result pointing to one particular WTC that performs better on the purpose of modelling BA time series in all Iberian regions. Depending on the considered region and month, the best performing classification change. However, some of the tested classifications tend to present the best results more frequently: PETISCOC09, WLKC09 and LUNDC09. Each of these WTC presents some of the most satisfying results in at least two clusters, while other classifications do not present good performances in none of the considered regions and months that were analysed and were therefore discarded from the remaining of the analysis. In the western sectors of Iberia (NW_CLU and SW_CLU), several WTC used alone are sufficient to developed statistical significance models at the 5% level. In these clusters, the variance of July BA time series is less explained by previous climatic conditions alone than in August or July + August BA time series, but there was a significant gain by introducing current WTC information. This difference between July and August models might result from the fact that the ignition of a fire in the middle of August does not always require an extreme 'synoptic forcing', due to the accumulation of previous hot and dry spells, while earlier in July a large fire might require a more intense 'synoptic forcing' to be initiated. In the Mediterranean coastal areas, BA in July was also difficult to model based purely on previous climatological conditions, and in this particular case, the WTC information does not improve significantly the models—only the BA in August and in July + August time series can be well reproduced by these statistical models. A study by Millán *et al*. (1998) suggests that in this region, early summer fire activity is highly related to the penetration of sea breezes inland, during well established warm and dry conditions in the interior, while in late summer westerly dry and warm inland flows (*Poniente*) play a more important role. This fact may provide some context for the poor results obtained with the models of BA in July, as sea breezes occur frequently at the mesoscale (i.e. at a finer resolution than the considered datasets in this work), while the *Poniente* conditions can be well represented at the synoptic scale. Models for the northernmost regions present the worst results even when both types of information are used as predictors, although some few WTC can present interesting results. Three possible facts might help to explain such regional discrepancies: (1) the low performance of the N_CLU and E_CLU models might result from an insufficient level of spatial aggregation to solve the climatological and topographical complexity of both regions; (2) the reliability of the seasonal models is reduced, as the BA in the NW_CLU reveals to be a consequence of a large number of 'small' fires, while when we move to the southeast most of the BA rests on a few large forest fires; (3) as a response of such level of complexity, atmospheric mechanisms might combine themselves in a nonlinear way, sometimes reinforcing the total forcing, sometimes weakening it. In any case, the analysis of those facts is beyond the scope of this paper.

It was also important to assure that WT predictors which were retained by the selection method corresponded to synoptic patterns associated with large wildfire activity episodes. The composites for the majority of WTs that were retained as predictors for each cluster correspond to stable conditions, where the specific region is influenced by warm and dry air masses. The atmospheric processes that could enhance the susceptibility for large BA are associated to either the advection of hot and dry inland air to coastal regions, or diabatic heating and drying on the downslope side of the mountainous systems.

On the overall, and despite being almost impossible to identify the WTC more suitable to discriminate BA in Iberia, it is quite clear that their information is very useful and valuable enough to develop statistical models that can explain the variability of BA in this region. In fact, instead of picking and focusing on one or two particular classifications *a priori*, we find a more efficient type of approach to take them all initially into account, and use objective criteria and automatic methods to select the classifications with better performance in each particular case.

Regionalized scenarios of climate change for Spain (MARN, 2005; AEMET-IM, 2011) and Portugal show a sustained warming (Ramos *et al*., 2011) and a reduction of the precipitation (Gouveia *et al*., 2011), more remarkable during spring, outcome of the reduction of Atlantic storm systems (Lorenzo *et al*., 2011). Consequently, the overall increase of the fire risk associated to a warmer world should be enhanced in western Iberia due to the long-term reduction of the springtime precipitation. The good performance of models in this region offers a promising tool to simulate the BA amount from the outputs of the regional climate models. Finally, we must acknowledge the fact that some of the Iberian regions characterized in this work might be affected by the existence of asymmetric trends in BA annual values. In particular, it has been shown that Spain reveals a negative trend of BA in the last three decades while Portugal has not revealed such pattern, being more appropriately characterized by an increasing tendency between the 1980s and 2000s (Pereira *et al*., 2011a).

Acknowledgements

This work was supported by European Union Funds (FEDER/COMPETE—Operational Competitiveness Programme) and by national funds (FCT—Portuguese Foundation for Science and Technology) under the project FCOMP-01-0124-FEDER-022692, the project FLAIR (PTDC/AAC-AMB/104702/2008) and the EU 7th Framework Program through FUME (contract number 243888). We are also in debt to the Autoridade Florestal Nacional, the Portuguese Forest service for providing the Portuguese fire data and to Cátia Teixeira for the final text revision of the manuscript.

References

Agencia Estatal de Meteorología (AEMET-IM). 2011. Atlas climático de España y Portugal http://www.aemet.es/es/serviciosclimaticos/datosclimatologicos/atlas_climatico

AFN (Autoridade Florestal Nacional). 2011. Statistics: Wildfire Data [Online] Available at: http://www.afn.min-agricultura.pt/portal/dudf/estatisticas/estatistica-2013-dados-sobre-incendios-florestais [Accessed 1 March 2011].

Austin RB, Cantero-Martínez C, Arrúe JL, Playán E, Cano-Marcellán P. 1998. Yield–rainfall relationships in cereal cropping systems in the Ebro river valley of Spain. *European Journal of Agronomy* **8**: 239–248.

Badía A, Saurí D, Cerdan R, Llurdés JC. 2002. Causality and management of forest fires in Mediterranean environments: an example from Catalonia. *Environmental Hazards* **4**: 23–32.

Barbosa P, Amatulli G, Boca R, Camia A, Kucera J, Libertà G, San-Miguel Ayanz J, Schmuck G, Schulte E, Dierks HH. 2007. Forest Fires in Europe 2006, EUR 22931 EN—Joint Research Centre—*Institute for Environment and Sustainability, Luxembourg: Office for Official Publications of the European Communities, EUR—Scientific and Technical Research series*: ISSN 1018–5593.

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. *International Journal of Wildland Fire* **17**(3): 328–338.

Carracedo Martín V, Diego Liaño C, García Codrón JC, Rasilla Álvarez D. 2009. *Incendios forestales*. Davinci Continental, *S.L.*

COST 733 Wiki. 2011. Harmonisation and Applications of Weather Type Classifications for European regions or COST733 spatial domains for Europe. Available at http://geo21.geo.uni-augsburg.de/cost733wiki/Cost733_Wiki_Main [Accessed 30 May 2013].

Costa L, Thonicke K, Poulter B, Badek FW. 2010. Sensitivity of Portuguese forest fires to climatic, human, and landscape variables: subnational differences between fire drivers in extreme fire years and decadal averages. *Regional Environmental Change* **11**(3): 543–551. DOI: 10.1007/s10113-010-0169-6

Cueva AV, Barrio JMG, Quero MO, Palomares OS. 2006. Recent fire regime in peninsular Spain in relation to forest potential productivity and population density. *International Journal of Wildland Fire* **15**(3): 397–405. DOI: 10.1071/WF05071

Elith J, Graham CH, Anderson RP, Ferrier M, Dudík S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, McC Overton J, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberón J, Williams S, Wisz MS, Zimmermann NE. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**: 129–151.

European Commission. 1998. Common Information System on Forest Fires, 1985–1997 study. *European Commission, Directorate-General Agriculture*: DG VI FII-2.

Fischer EM, Schär C. 2010. Consistent geographical patterns of changes in high-impact European heatwaves. *Nature Geoscience* **3**: 398–403. DOI: 10.1038/NGEO866

Fischer EM, Seneviratne SI, Lüthi D, Schär C. 2007. Contribution of land-atmosphere coupling to recent European heatwaves. *Geophysical Research Letters* **34**: L06707.

Flannigan MD, Wotton B. 2001. Climate, weather and area burnt. In *Forest Fires—Behavior and Ecological Effects*, Johnson EA, Miyanishi K (eds). Academic Press: San Diego; 351–373.

García-Herrera R, Paredes D, Trigo RM, Trigo IF, Hernández H, Barriopedro D, Mendes MT. 2007. The outstanding 2004/05 drought in the Iberian Peninsula: associated atmospheric circulation. *Journal of Hydrometeorology* **8**: 483–498.

Gil-Romera G, Carrión JS, Pausas JG, Fernández S, Burjachs F. 2010. Holocene fire activity and vegetation response in South-Eastern Iberia. *Quaternary Science Reviews* **29**: 1082–1092.

Giorgi F. 2006. Climate change hot-spots. *Geophysical Research Letters* **33**: L08707.

Gouveia C, Trigo RM, DaCamara CC, Libonati R, Pereira JMC. 2008. The North Atlantic Oscillation and European Vegetation Dynamics. *International Journal of Climatology* **28**: 1835–1847.

Gouveia C, Trigo RM, DaCamara CC. 2009. Drought and vegetation stress monitoring in Portugal using Satellite Data. *Natural Hazards and Earth System Sciences* **9**: 185–195.

Gouveia C, DaCamara CC, Trigo RM. 2010. Post fire vegetation recovery in Portugal based on SPOT-VEGETATION data. *Natural Hazards and Earth System Sciences* **10**: 673–684.

Gouveia C, Liberato MLR, DaCamara CC, Trigo RM. 2011. Modelling past and future wine production in the Portuguese Douro Valley. *Climate Research* **48**: 349–362. DOI: 10.3354/cr01006

Hoinka KP, Carvalho A, Miranda AI. 2009. Regional-scale weather patterns and wildland fires in central Portugal. *International Journal of Wildland Fire* **18**(1): 36–49.

JRC: Forest Fires in Europe 2009. Report EUR 24502 EN. 2010. JRC-IES/Land Management & Natural Hazards Unit. ISSN 1018–5593. [Online] Available at: http://effis.jrc.ec.europa.eu/reports/fire-reports/ [Accessed 1 April 2011].

Kassomenos P. 2010. Synoptic circulation control on wild fire occurrence. *Physics and Chemistry of the Earth* **35**: 544–552. DOI: 10.1016/j.pce.2009.11.008

Kunkel KK. 2001. Surface energy budget and fuel moisture. In *Forest Fires—Behaviour and Ecological Effects*, Johnson EA, Miyanishi K (eds). Academic Press: San Diego; 303–350.

Leśniok M, Małarzewski L, Niedźwiedź T. 2010. Classification of circulation types for Southern Poland with an application to air pollution concentration in Upper Silesia. *Physics and Chemistry of the Earth, Parts A/B/C* **35**(9–12): 516–522.

Lloret F, Calvo E, Pons X, Díaz-Delgado R. 2002. Wildfires and landscape patterns in the Eastern Iberian Peninsula. *Landscape Ecology* **17**: 745–759.

Lorenzo MN, Ramos AM, Taboada JJ, Gimeno L. 2011. Changes in Present and Future Circulation Types Frequency in Northwest Iberian Peninsula. *PLoS One* **6**(9–12): e16201. DOI: 10.1371/journal.pone.0016201.

Lupikasza E. 2010. Relationships between occurrence of high precipitation and atmospheric circulation in Poland using different classifications of circulation types. *Physics and Chemistry of the Earth, Parts A/B/C* **35**(9–12): 448–455.

Mariotti A. 2010. Recent changes in Mediterranean water cycle: a pathway toward long-term regional hydroclimatic change? *Journal of Climate* **23**: 1513–1525.

Mariotti A, Zeng N, Yoon J, Artale V, Navarra A, Alpert P, Li LZX. 2008. Mediterranean water cycle changes: transition to drier 21st century conditions in observations and CMIP3 simulations. *Environmental Research Letters* **3**: 044001. DOI: 10.1088/1748-9326/3/4/044001

Ministerio de Medio Ambiente, Marino y Rural (MARN). 2011. Los incendios forestales en españa decenio 1996–2005 (http://www.marm.es/es/biodiversidad/temas/defensa-contra-incendios-forestales/decenio_1996_2005_tcm7-19437.pdf) [Accessed 21 September 2011].

Martínez-Villalta J, López BC, Adel N, Badiella L, Ninyerola M. 2008. Twentieth century increase of Scots pine radial growth in NE Spain shows strong climate interactions. *Global Change Biology* **14**: 2868–2881.

Millán M, Estrela MJ, Badenas C. 1998. Meteorological Processes Relevant to Forest Fire Dynamics on the Spanish Mediterranean Coast. *Journal of Applied Meteorology* **37**: 83–100. DOI: 10.1175/1520-0450(1998)037<0083:MPRTFF>2.0.CO;2

Ministerio de Medio Ambiente, Marino y Rural (MARN). 2005. Evaluación preliminar de los impactos en España del cambio climático (http://www.marm.es/es/cambio-climatico/temas/impactos-vulnerabilidad-y-adaptacion/iniciativas-en-el-ambito-nacional/evaluacion-preliminar-de-los-impactos-en-espana-del-cambio-climatico/) [Accessed 29 October 2011].

Myneni RB, Hall FG, Sellers PJ, Marshak AL. 1995. The interpretation of spectral vegetation indexes. *IEEE Transactions on Geoscience and Remote Sensing* **33**: 481–496.

Mitchell TD, Jones PD. 2005. An improved 258 method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology* **25**: 693–712.

Moreno JM, Vázquez A, Vélez R. 1998. Recent history of forest fires in Spain. In *Large Forest Fires*, Moreno JM (ed). Backhuys: Leiden; 159–186.

Neter J, Kutner M, Nachtshei C, Wasserma W. 1996. *Applied Linear Regression Models*, 3 edn. Boston: McGraw-Hill.

Padilla M, Vega-García C. 2009. On the comparative importance of fire danger rating indices and their integration with spatial and temporal variables for predicting daily human-caused fire occurrences in Spain. *International Journal of Wildland Fire* **20**(1): 46–58. DOI: 10.1071/WF09139

Paredes D, Trigo RM, García-Herrera R, Trigo IF. 2006. Understanding precipitation changes in Iberia in early spring: weather typing and storm-tracking approaches. *Journal of Hydrometeorology* **7**: 101–113.

Pausas JG. 2004. Changes in fire and climate in the Eastern Iberian Peninsula (Mediterranean Basin). *Climatic Change* **63**: 337–350.

Pausas JG, Vallejo R. 1999. The role of fire in European Mediterranean ecosystems. In *Remote Sensing of Large Wildfires in the European Mediterranean Basin*, Chuvieco E (ed). Springer-Verlag: Berlin; 3–16.

Pausas JG, Fernández-Muñoz S. 2011. Fire regime changes in the Western Mediterranean Basin: from fuel-limited to drought-driven fire regime. *Climatic Change* **110**(1-2): 215–226. DOI: 10.1007/s10584-011-0060-6

Pereira MG, Trigo RM, DaCamara CC, Pereira JMC, Solange ML. 2005. Synoptic patterns associated with large summer forest fires in Portugal. *Agricultural and Forest Meteorology* **129**: 11–25. DOI: 10.1016/j.agrformet.2004.12.007

Pereira MG, Malamude BD, Trigo RM, Alves PI. 2011a. The History and Characteristics of the 1980–2005 Portuguese Rural Fire Database. *Natural Hazards and Earth System Sciences* **11**: 1–16. DOI: 10.5194/nhess-11-1-2011

Pereira MG, Caramelo L, Gouveia C, Gomes-Laranjo J, Magalhães M. 2011b. Assessment of weather-related risk on chestnut productivity. *Natural Hazards and Earth System Sciences* **11**: 1–12. DOI:10.5194/nhess-11-12-011

Philipp A, Bartholy J, Beck C, Erpicum M, Magalhães M, Esteban P, Huth R, James P, Jourdain S, Krennert T, Lykoudis S, Michalides S, Pianko K, Post P, Rassilla Álvarez D, Schiemann R, Spekat A, Tymvios FS. 2010. COST733CAT – a database of weather and circulation type classifications. *Physics and Chemistry of the Earth* **35**: 360–373.

Pineda N, Esteban P, Trapero L, Soler X, Beck C. 2010. Circulation types related to lightningactivity over Catalonia and the Principality of Andorra. *Physics and Chemistry of the Earth, Parts A/B/C* **35**(9–12): 469–476.

Piñol J, Terradas J, Lloret F. 1998. Climate warming, wildfire hazard and wildfire occurrence in coastal Eastern Spain. *Climatic Change* **38**: 345–357.

Pyne SJ, Andrews PL, Laven RD. 1996. *Introduction to Wildland Fire*, 2 edn. New York: John Wiley & Sons.

Ramos A, Trigo RM, Santo FE. 2011. Evolution of extreme temperatures in Portugal: reporting on recent changes and future scenarios. *Climate Research* **48**: 177–192. DOI: 10.3354/cr00934

Rangel TFLVB, Diniz-Filho JAF, Bini LM. 2010. SAM: a comprehensive application for Spatial Analysis in Macroecology. *Ecography* **33**: 46–50.

Rasilla DF, García-Codron JC, Carracedo V, Diego C. 2010. Circulation patterns, wildfire risk and wildfire occurrence at continental Spain. *Physics and Chemistry of the Earth* **35**: 553–560. DOI: 10.1016/j.pce.2009.09.003

Rivas Martínez S. 1987. Memoria del mapa de series de vegetación de España 1: 400.000. *ICONA*. Ministerio de Agricultura, Pesca y Alimentación, Madrid. ISBN 84-85496-25-6.

Rubiales JM, García Amorena I, García Álvarez S, Gómez Manzaneque F. 2008. The Late Holocene extinction of Pinus sylvestris in the West of the Cantabrian Range. *Journal of Biogeography* **35**(10): 1840–1850. DOI: 10.1111/j.1365-2699.2008.01925.x

Seneviratne SI, Lüthi D, Litschi M, Schär C. 2006. Land-atmosphere coupling and climate change in Europe. *Nature* **443**: 205–209. DOI: 10.1038/nature05095

Sousa P, Trigo RM, Aizpurua P, Nieto R, Gimeno L, García-Herrera R. 2011. Trends and extremes of drought indices throughout the 20th century in the Mediterranean. *Natural Hazards and Earth System Sciences* **11**: 33–51.

Trigo RM, Pereira JMC, Pereira MG, Mota B, Calado MT, DaCamara CC, Santo FE. 2006. Atmospheric conditions associated with the exceptional fire season of 2003 in Portugal. *International Journal of Climatology* **26**(13): 1741–1757.

Tucker CJ, Pinzon JE, Brown ME, Slayback D, Pak EW, Mahoney R, Vermote E, El Saleous N. 2005. An extended AVHRR 8-km NDVI data set compatible with MODIS and SPOT vegetation NDVI data. *International Journal of Remote Sensing* **26**: 4485–4498.

Twardosz R. 2010. An analysis of diurnal variations of heavy hourly precipitation in Kraków using a classification of circulation types over southern Poland. *Physics and Chemistry of the Earth, Parts A/B/C* **35**(9–12): 456–461.

Ustrnul Z, Czekierda D, Wypych A. 2010. Extreme values of air temperature in Poland according to different atmospheric circulation classification. *Physics and Chemistry of the Earth, Parts A/B/C* **35**(9–12): 429–436.

Valbuena-Carabaña M, López de Heredia U, Fuentes-Utrilla P, González-Doncel I, Gil L. 2010. Historical and recent changes in the Spanish forests: A socio-economic process. *Review of Palaeobotany and Palynology* **162**(3): 492–506.

Vázquez A, Moreno MJ. 1993. Sensitivity of fire occurrence to meteorological variables in Mediterranean and Atlantic areas of Spain. *Landscape and Urban Planning* **24**: 129–142.

Vázquez A, Pérez B, Fernández-González F, Moreno JM. 2002. Recent fire regime characteristics and potential natural vegetation relationships in Spain. *Journal of Vegetation Science* **13**: 663–676.

Vélez R. 1993. High intensity forest fires in the Mediterranean Basin: natural and socioeconomic causes. *Disaster Manage* **5**: 16–20.

Vicente-Serrano SM. 2007. Evaluating the Impact of drought using Remote Sensing in a Mediterranean, semi-arid region. *Natural Hazards* **40**: 173–208.

Vicente-Serrano SM, Heredia-Laclaustra A. 2004. NAO influence on NDVI trends in the Iberian Peninsula (1982–2000). *International Journal of Remote Sensing* **25**: 2871–2879.

Viedma O, Moreno JM, Rieiro I. 2006. Interactions between land use/land cover change, forest fires and landscape structure in Sierra de Gredos (central Spain). *Environmental Conservation* **33**: 212–222.

Viegas DX, Viegas MT. 1994. A relationship between rainfall and burned area for Portugal. *International Journal of Wildland Fire* **4**(1): 11–16.

Wilks DS. 2006. *Statistical Methods in the Atmospheric Sciences*, 2 edn. Amsterdam: Academic Press.

Wisz MS, Guisan A. 2009. Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology* **9**: 8.

Zhao M, Running SW. 2010. Drought-Induced Reduction in Global Terrestrial Net Primary Production from 2000 through 2009. *Science* **20**: 940–943.

