**Natural Hazards
and Earth
System Sciences**

The history and characteristics of the 1980–2005 Portuguese rural fire database

M. G. Pereira[1,2]**, B. D. Malamud**[3]**, R. M. Trigo**[2]**, and P. I. Alves**[4]

[1]Centre for Research and Technology of Agro-Environment and Biological Sciences (CITAB), University of Trás-os-Montes and Alto Douro, Vila Real, Portugal
[2]CGUL, Centro de Geofísica da Universidade de Lisboa, Lisboa, Portugal
[3]Department of Geography, Kings College London, UK
[4]AFN, Autoridade Florestal Nacional, Lisboa, Portugal

Abstract. We focus here on a mainland Continental Portuguese Rural Fire Database (PRFD) that includes 450 000 fires, the largest such database in Europe in terms of total number of recorded fires in the 1980–2005 period. In this work, we (a) list the most important factors for triggering and controlling the fire regime in mainland Continental Portugal, (b) describe the dataset's production, (c) discuss procedures adopted to identify and correct different fire data inconsistencies, creating a modified PRFD which we use here and make available as Supplement, (d) explore some basic temporal and completeness properties of the data. We find that the dataset's minimum measured burnt areas have changed with time between $A_F = 0.1$ ha (1980–1990), $A_F = 0.01$ ha (1991–1992), and $A_F = 0.001$ ha (1992–2005), with varying degrees of completeness down to these values. These changes in minimum area measured are responsible for greater numbers of fires being recorded. A relatively small number of large fires in the PRFD are responsible for the majority of the burnt area. For example, fires with $A_F > 100$ ha represent about 1 % of all fire records but 75 % of total burnt area. Finally, we consider for each Continental Portugal district and for the 26-yr period, the total number of rural fires and area burnt in forests and shrublands, each normalized by district areas. We find that the highest numbers of fires per unit area are in highly populated districts, and that the largest fraction of burnt area is in forested areas, coinciding with large parcels of continuous forests (predominantly rural and moderately urban areas).

1 Introduction

This paper examines the Portuguese Rural Fire Database (PRFD), provided by the Autoridade Florestal Nacional (AFN, 2011), the current Portuguese Forest Service, which includes information for more than 450 000 fire records that have occurred in Continental Portugal for the 26-yr period, 1980–2005. The PRFD dataset is restricted to fires that have occurred in forests, shrublands, natural grasslands, and agricultural areas. In general, these vegetation areas are not homogenous, but a mixture of different land cover and land use types. We use the term "rural" here for the fire database, as the PRFD does not include urban fires, i.e. fires in buildings and other human-built structures. However, fires that occurred in large forested areas or shrublands within urban areas are included in the dataset.

Continental Portugal has an area of 88 970 km^2 (Statistics Portugal, 2007) with the mainland excluding the areas of Madeira (800 km^2) and Azores (2320 km^2). According to the fire records available in the PRFD, the total of the burnt areas registered 1980–2005 is roughly 3.0×10^6 ha, i.e. 34 % of the total area of Continental Portugal. This value should be considered as a minimum because the PRFD dataset is far from being complete, particularly in the 1980s. Completeness of the dataset is discussed further in Sect. 5.

Portugal holds one of the most comprehensive rural fire databases, not only in Europe but also in comparison to many other countries worldwide. According to Barbosa et al. (2007), the northern Mediterranean countries (Portugal, Spain, France, Italy and Greece) are particularly affected by summer fires and these European countries have the longest consistent time series of fire data during the 1980–2006 period. Using data provided by the European Forest Fire Information System (JRC, 2010) we give (Table 1) the number of fires per year (N_{FT}) and burnt area per year (A_{FT}) for each

Table 1. Number of fires (top) and burnt area (bottom) for five northern Mediterranean countries, for three different periods of time (Table based on yearly values from JRC, 2010)[a].

	Number of Fires						
Period	Quantity	Portugal[b]	Spain	France	Italy	Greece[c]	Total
1980–1989	Average # fires per yr	7381	9515	4910	11 575	1264	34 645
1990–1999	Average # fires per yr	22 250	18 152	5538	11 164	1748	58 851
2000–2009	Average # fires per yr	24 949	18 337	4406	7259	1695	56 645
1980–2009	Average # fires per yr	18 194	15 335	4951	9999	1569	50 047
1980–2009	Total # fires	545 805	460 038	148 531	299 977	47 058	1 501 409
1980–2009	(Total # fires)/(country land area) (# of fires per 100 ha)	5.5	0.8	0.3	0.9	0.3	1.0

	Burnt area (ha)						
Period	Quantity	Portugal[b]	Spain	France	Italy	Greece[c]	Total
1980–1989	Average burnt area per yr	73 484	244 788	39 157	147 150	52 417	556 995
1990–1999	Average burnt area per yr	102 203	161 319	22 735	118 573	44 108	448 938
2000–2009	Average burnt area per yr	150 101	125 239	22 342	83 878	49 238	430 798
1980–2009	Average burnt area per yr	108 596	177 115	28 078	116 534	48 587	478 910
1980–2009	Total burnt area	3 257 886	5 313 457	842 332	3 496 005	1 457 624	14 367 304
1980–2009	(Total burnt area)/(country land area) (ha burnt per 100 ha)	36.6	10.6	1.5	11.9	11.3	9.2

[a] Table does not include information about the completeness of the data (i.e. different countries or administrative areas within each country may not include fires below a given burnt area).
[b] The country land area used here for Portugal is 91 500 km^2, as the statistics here are based on all of Continental Portugal (vs. just mainland Portugal).
[c] Provisional data for 2009.

of the five northern Mediterranean countries and three different periods of time (1980–1989, 1990–1999, 2000–2009). We then normalize (Table 1) the total number of fires and burnt area for each country, by that country's land area (A_C). Then, in Fig. 1, we show the temporal yearly evolution of the density of fire occurrences and percentage of area burnt for 1980–2009. Barbosa et al. (2007) and the JRC (2010) do not consider differing degrees of completeness (e.g. down to "what" burnt area fires are measured, different ways of measuring wildfires, etc.), but some broad conclusions can be drawn from these fire statistics.

The importance of fire activity in Continental Portugal for 1980–2009 is strongly evident in Table 1 (all years considered together) and in Fig. 1 (statistics considered year by year). For the entire 1980–2009 period (Table 1), after normalizing the number of fires and total area burnt for each country (again, within the limitations of completeness of the data gathered), by the correspondent country's total land area (A_C), Portugal has six (three) times more fires (burnt area) per unit area than Italy, the second most fire-affected country. When examining individual years (Fig. 1), differences are even more dramatic. For instance, in 2003 to 2005, the density of fires (percentage of land burnt) in Portugal was over eight (ten) times greater than that of Spain, who had the second largest values of the five countries for those three years. Similar conclusions arise from the comparison with other non-European countries highly affected by fire such

as Canada, the United States and the Russian Federation (ECE/FAO, 2003).

When considering the 30-yr period (1980–2009) broad trends, in Fig. 1a, Portugal has a clear upwards trend for the density of fires, N_{FT}/A_C; whereas, the other four countries have no clear upwards or downwards trend. The upwards trend for Portugal will be discussed in-depth in Sect. 5.1, where we will show that this trend is not due to the actual increasing numbers of fires, but the way statistics have been gathered. In Fig. 1b, for 1980–2009, Portugal has no clear trend for the percentage of land area burnt, A_{FT}/A_C, except for the last few years where there are two higher values (2003, 2005). However, for Greece and Italy, there is a clear downwards trend for the percentage of land area burnt during this 30-yr period, a not as clear downward trend for Spain, and for France there is no trend upwards or downwards.

Versions of the PRFD have been widely used to support a number of fire studies in Portugal (e.g. Pereira et al., 1998; Pereira et al., 2005; Silva and Catry, 2006; Trigo et al., 2006; Catry et al., 2007; Carvalho et al., 2008; Hoinka et al., 2009; Costa et al., 2010) as well as several national and European official reports (DGRF, 2007; Statistics Portugal, 2007; Barbosa et al., 2007; JRC, 2010). Because of the importance of this large dataset in supporting rural fire studies, one should be fully aware of this dataset's characteristics, virtues, limitations, and robustness.
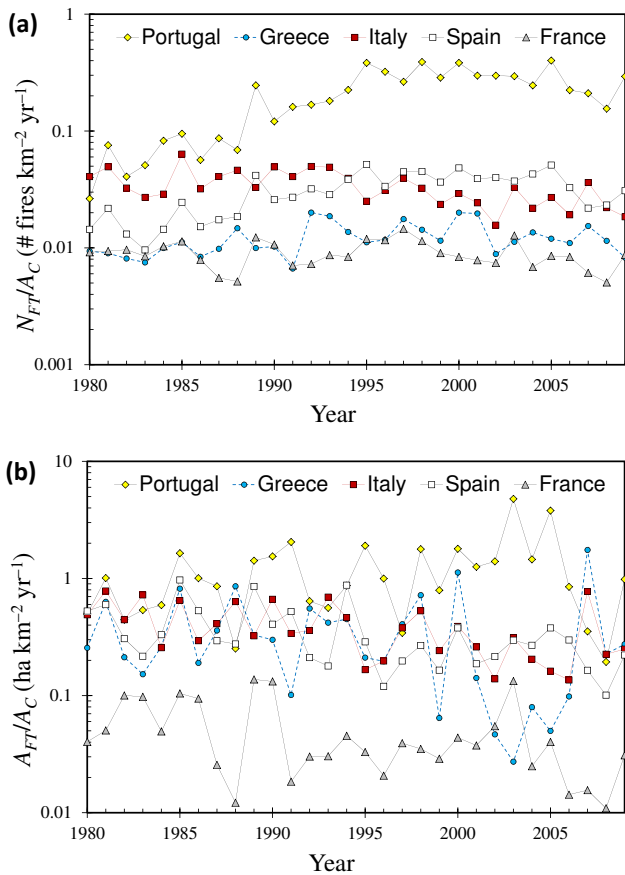
Fig. 1. Rural (wildland) fires in Portugal, Spain, France, Italy and Greece, 1980–2009. For each country, the temporal evolution of **(a)** fire occurrences density (number of fires per year N_{FT} (#) normalized by country area A_C (km^2)) and **(b)** percentage of country land area burnt by rural fires (burnt area per year A_{FT} (ha) normalized by country area (km^2)) is given. Both y-axes are on log scales. Figure based on data provided in the Annual Reports on Forest Fires in Europe of the European Forest Fire Information System (JRC, 2010). Data does not take into account each country's different degrees of completeness in measuring wildfire ("rural" fire) numbers and wildfire areas, both changing over time and depending on administrative areas within each country.

We believe that exposing the characteristics, errors and inconsistencies of the PRFD is of fundamental importance for researchers pretending to use the current or future versions of this database. Moreover, this approach may be useful for the larger wildland fire research community, stressing the need to address similar caveats in the datasets used in their analysis. In this sense, the main objectives of this work are: (i) to describe the assembling and confirmation procedures that have been used to generate the PRFD; (ii) to present the procedures used to identify and correct a number of PRFD errors; (iii) to present and interpret some of the PRFD characteristics, including spatial and temporal distributions of the fires.

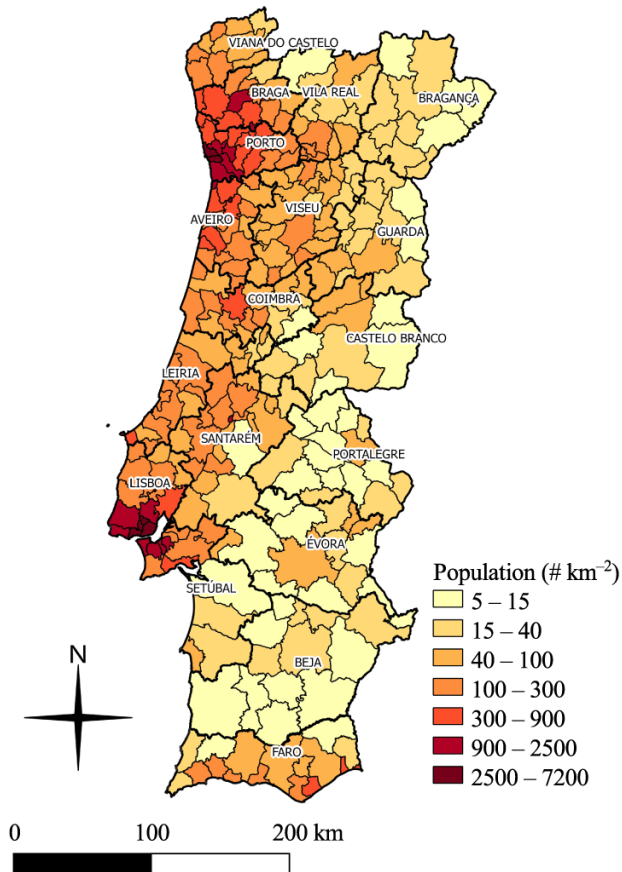In the following section (Sect. 2) we present a description of the most important factors that affect the rural fire regime in Continental Portugal. We then provide a general description of the PRFD (Sect. 3) and the types of errors in the records and corrections applied (Sect. 4). After this, we present temporal and spatial statistics for the rural fires (Sect. 5). Finally, we discuss the results and present some concluding remarks (Sect. 6).

2 Factors influencing rural fires regimes in Portugal

A number of environmental and anthropogenic factors contribute to Portugal being considered the European country most prone (per unit area) to fire occurrence. Environmental factors include climate, topography, and distribution of tree species. Anthropogenic factors include population density, and related, urban vs. non-urban occupation of land. These factors are addressed in this section.

Continental Portugal is located in the southwestern extreme of Europe with a near-rectangular shape (approximately defined by latitudes 37–42° N and longitudes 6–10° W). The country's current division of administrative regions is based on a 2008 classification with 18 districts, 278 counties and 4050 parishes (Statistics Portugal, 2011). The location and names of the Portuguese districts and the county boundaries are shown in Fig. 2. There are significant geomorphologic, climatic and demographic differences between the northern and southern parts of Portugal. The northern half of Portugal is considerably more mountainous, containing 95 % of mainland Portugal areas with elevation >400 m a.s.l. (above sea level). This area is thus wetter with many more watercourses and larger river basins compared to the southern half of Portugal (Trigo and DaCamara, 2000). Precipitation and temperature exhibit a marked seasonal character, with a dry season that has almost no rainfall during the warm and dry summer (June to August) months, when the majority of rural fire activity is registered, and a wetter period is found throughout the rest of the year with maximum values during the cold winter months (November to February) (Trigo and DaCamara, 2000; Meteorological Institute, 2011).

The population density and composition (namely, the urban/rural population ratios, and the percentage of the population that is agricultural) and land structure statistics (both in Table 2) are helpful to understand the context of Portuguese fire statistics within those Mediterranean countries most affected by fires, Portugal, Spain, France, Italy and Greece (Table 1). In Table 2 we see that Portugal, compared to these other countries, has the lowest urban/rural ratio, the highest percentage of population dedicated to agriculture, the highest percentage of forested land area and the lowest percentage of agricultural and arable lands. We also note that the percentage of land that is forested for each of the five northern Mediterranean countries has increased from 1990 to 2005, with Spain and Portugal increasing the most (Spain: from

The 5th National Forestry Inventory (NFI5, 2011) provided by the Autoridade Florestal Nacional (AFN, 2011) was based on a digital aerial-photo coverage obtained during the 2004–2006 period, and on a ground survey performed from December 2005 to June 2006; it is the most recent assessment on land cover and land use in Portugal. According to NFI5 (2011), forest, agriculture and shrubs cover 38.8, 32.9 and 21.6 % of Continental Portugal's area, respectively; inner water bodies and other land cover types cover 1.8 and 4.8 %. This report classified the Portuguese forest between stands of broadleaved (62 %), coniferous (28 %) and mixed forest (10 %). According to this report, the predominant tree species are: *Pinus pinaster* (28 %), *Quercus suber* L. and *Eucalyptus globulus* (23 %) and *Quercus ilex* (13 %). The *pinus pinaster* may be found essentially in the central north part of the country (north of the Tagus river), where high rural fire activity has been registered. The evergreen species *Quercus suber* L. and *Quercus ilex*, which presents a high resistance to unfavourable climate conditions and to fire, are predominant in the southern part of the country where the highest (lowest) values of summer temperature (precipitation) are usually registered. The highly combustible *Eucalyptus globulus* is generally planted in controlled and well managed forests. The large majority of the Portuguese forest is under private control and has been developed for commercial purposes (e.g. Portugal is one of the leading producers of cellulose in Europe).

According to the National Forestry Inventory (DGF, 2001), bushes, pasture (grass-land; pasture ground), spontaneous grazing lands and uncultivated or abandoned fields, also cover a considerable amount (23 %) of land area. Shrublands are common in all districts and due to changes in land cover type and fire regimes, their presence has increased from 16 % of the mainland Portuguese land area in 1993 to 18 % in 1997 (DGF, 1993, 1997; Fernandes, 2001). These fuel types present a high probability of ignition because of their high flammability and are capable of sustaining extreme fire intensities (Fernandes, 2001; Vasconcelos et al., 2001). Open forests of oak and pine with a continuous shrub understorey are also extensive in several regions of the country.

3 Portuguese Rural Fire Database (PRFD)

3.1 Portuguese Rural Fire Database compilation history

Under Portuguese law, firefighters are the source of the information relative to each rural fire occurring in Portugal (DL, 1980; DR, 1981). During the period 1980–2000, the Autoridade Florestal Nacional (AFN) central office in Lisbon received Situation Reports (SITREPs) from firefighters, which included pertinent information about each rural fire. Afterwards, the Portuguese "Forest Guards" further investigated a subset of these reports, particularly with respect to

Fig. 2. Portugal's main administrative regions (18 districts) and their corresponding municipalities (278 counties). The map shows the population density distribution (# km^{-2}) for each county. Population data from Statistics Portugal (2011) and shape file from IGEO (2011).

27.0 % (1990) to 35.9 % (2005), a 33 % increase; Portugal: from 33.9 % (1990) to 41.3 % (2005), a 22 % increase).

The resident population in Portugal, estimated at 10.5 million inhabitants in 2005, is concentrated in the northern and central coastal areas (Statistics Portugal, 2007). From 1950 to present, there has been a massive internal population migration from the rural interior to the coastal urban areas, with more than half the country's population living in towns and cities with more than 2000 inhabitants in 2001 (Statistics Portugal, 2007). According to Catry et al. (2007), the population density is related to the spatial distribution of fire occurrences. They found that for 2001–2005, 70 % of fire ignitions and 14 % of the total burnt area were registered in municipalities with more than 100 persons per km^2 (21 % by area of the mainland Portugal). Other studies (DGRF, 2006) found that the highly populated districts of Lisboa, Porto, Braga and Aveiro have more than half of the fire occurrences in the 1980–2006 period, but represent only a small percentage of burnt area every year.

Table 2. Population and land area statistics for five northern Mediterranean countries for the year of 2005 and percent of land area which was forested for 1990, 1995, 2000, and 2005. All data from FAOSTAT (2011).

Measure	Portugal	Spain	France	Italy	Greece
Population (pop.) $\times 10^3$ (#)	10 547	43 060	61 013	58 645	11 064
Pop. density (# ha^{-1})	1.153	0.863	1.114	1.994	0.858
Rural pop.[a] $\times 10^3$ (# [%])[a]	4470 [42]	10 024 [23]	14 215 [23]	18 994 [32]	4386 [40]
Urban pop.[a] $\times 10^3$ (# [%])[a]	6077 [58]	33 036 [77]	46 798 [77]	39 650 [68]	6678 [60]
Agriculture pop.[b] $\times 10^3$ (# [%])[b]	1293 [12]	2448 [6]	1589 [3]	2440 [4]	1260 [11]
Non-agriculture pop.[b] $\times 10^3$ (# [%])[b]	9254 [88]	40 612 [94]	59 424 [97]	56 205 [96]	9804 [89]
Country area $\times 10^3$ (ha)	9212[d]	50 537	54 919	30 134	13 196
Total land area[c] $\times 10^3$ (ha)	9150[d]	49 909	54 766	29 414	12 890
Forest area $\times 10^3$ (ha [%])[e]	3783 [41]	17 915 [36]	15 554 [28]	9979 [34]	3752 [29]
Agriculture area $\times 10^3$ (ha [%])[e]	3680 [40]	29 164 [58]	29 550 [54]	14 736 [50]	8334 [65]
Other land-type area $\times 10^3$ (ha [%])[e]	1687 [18]	2830 [6]	9662 [18]	4699 [16]	804 [6]
Forest land area percent by year					
1990 [%][e]	[33.9]	[27.0]	[26.5]	[28.5]	[25.6]
1995 [%][e]	[36.7]	[30.0]	[27.3]	[30.3]	[26.8]
2000 [%][e]	[39.2]	[32.9]	[28.0]	[32.1]	[27.9]
2005 [%][e]	[41.3]	[35.9]	[28.4]	[33.9]	[29.1]

[a] Percentages for rural and urban population are in relation to total population.
[b] Percentages for agriculture and non-agriculture population are in relation to total population; agriculture population can include rural, peri-urban and urban populations.
[c] Does not include inland water bodies.
[d] Includes the Azores and the Madeira Islands.
[e] Percentages in relation to total land area for that country.

spatial burnt areas and locations. In order to make the best use of their limited resources, confirmation was extensive for all fires with burnt area $A_F > 100$ ha, the large majority of fires with $A_F > 10$ ha, and for all fires that occurred in "priority" counties. The Portugal Regional Directorates of Agriculture identifies priority municipalities (counties) and appropriate regional prevention policies, with the aim of reducing the number of fires by increasing the number of investigations (DGF, 2002). The number of these "priority" counties was (DGF, 2002) initially during the late 1990s less than 20 counties each year (out of a total of 278 counties in the Portugal mainland) and 71 counties in 2002; these 71 counties covered 1/3 of Portugal's area, and in particular, the forested areas. Many of these "priority" counties have been identified as needing further persuasion and control, which is done by strengthening the dissuasive presence of the Forest Guards and the intensification of research into the cause of forest fires. The criteria for choosing these counties are: the economic or environmental value of the region to protect, a high registered density of ignitions in a given region, and the increased frequency of certain causes of fire (e.g. arson) or of fires conducted by pastors with the aim of the renewal of grassland for grazing. In each year, the priority counties change and, therefore, the percentage of fire records to be further investigated by the Forest Guards varied considerably from year to year, being dependent on the human resources available and the number of fires that oc-

curred in each area during that year. To confirm and further investigate the amount of burnt area recorded in firefighters' initial reports, the Forest Guards registered the area affected on a 1:25 000 map and then in the office, estimating the burnt area (total and by land cover type) with a planimeter. Burnt areas estimates were done in plan view and excluded unburnt areas. This work was performed in the Prevention and Detection Centers (CPD) at a regional level. This information was then sent to the AFN where it was included in the database by the Informatics Division.

After 1997, the database quality increased significantly because of: (i) adoption of a standard fire characteristics classification by firefighters and AFN technicians; and (ii) inclusion of an AFN technician in the National Centre of Operation and Help (Centro Nacional de Operação e Socorro, CNOS) during the fire season. These helped to greatly improve the information exchange between firefighters and AFN regional delegations, contributing to a better perception of the Portugal rural fire situation. This also led, eventually, to more confident information recorded by firefighters in the fire dataset. At present, some parts of the Forest Guard are confirming all the fire sizes (for some districts), or at least checking if the small fire (<2 ha) occurrences were not urban.

In 2001, the AFN established a protocol with the National Center of Geographical Information (Centro Nacional de Informação Geográfica, CNIG) to develop an Information

Management System for Forest Fires (Sistema de Gestão de Informação de Fogos Florestais, SGIF), to link (via intranet) the national dataset on forest fires. The SGIF software allows for the automatic uploading of fire occurrences and details including their spatial GIS information. The SGIF software thus results in a faster transmission of data from the CPD to the AFN central office, where technicians are able to detect some suspicious records (e.g. abnormal high or low A_F values compared to fire duration), quickly update discrepancies (e.g. date and time of ignition) and correct information (e.g. manually adjust the location) (DGRF, 2005). Although the use of SGIF started in 2001, during the early stage of the use of SGIF (2001–2002), fire fighters (in the central region of Portugal) used both the hardcopy and software and the standard hardcopy of manual records of fires as a process of validating the SGIF software application. During the 2001–2002 period, at the district level, the Forest Guards kept their standard method of investigating reports and mapping plan areas.

Since 2001, quality controls on the rural fire database include the central AFN offices checking for burnt areas from the same fire registered in two or more neighbouring administrative regions, and partitioning that fire accordingly. During 2001–2005, this procedure was performed for every fire with $A_F > 100$ ha, with a total of 189 records found (3 in 2001; 3 in 2002; 58 in 2003; 27 in 2004; 98 in 2005). To obtain the total burnt area A_F for each fire, an agglomeration of the multiple records of the same fire was performed. The AFN does not report any other error identification in its database. Satellite remote sensing information was not included in the AFN dataset or used for comparison and/or correction procedures, as the AFN dataset relies on ground measurements made by the fire-fighter teams.

3.2 Portuguese Rural Fire Database general description

The rural PRFD that we analyse here is representative of fires that have occurred in Continental Portugal 1980–2005, and includes the following information for each fire record:

i. Rural fire ignition and extinction date and time.

ii. Rural fire ignition location, in terms of administrative division of Portugal. In 2008 this consisted of 18 districts, subdivided into 278 counties, and further subdivided into 4050 parishes.

iii. The amount of area burnt in the fire, A_F.

iv. The land cover type (forest and shrublands) and total rural fire area. Forests are defined (AFN, 2011) as land occupied by >10 % forest trees, a minimum area of 0.5 ha and width >20 m. Shrublands are defined (AFN, 2011) as land occupied by shrub, natural origin, no agriculture or forestry, a minimum area of 0.5 ha and width >20 m.

v. Cause of rural fire.

vi. Some additional data is available but only for sub-periods of the original 1980–2005 period:

– the spatial coordinates of the rural fire ignition location (2001–2005)

– the burnt area in agricultural lands (2001–2005)

– land ownership (public or private) (1993–2003)

– type of fire (single or reignition) (1984–2005)

– time of the first fire fight attempt (2001–2003)

The amount of burnt area by forest fires (forests and shrublands) in the 1980–2005 period does not include fires in agricultural areas. However, for 2001–2005, individual burnt areas in agricultural lands is also available as an independent field.

A fire's location in terms of urban, rural or urban/rural interface is not available. However, Catry et al. (2007) examined spatial distribution patterns of wildfire ignitions in mainland Portugal, 2001–2005, by analysing 127 492 rural fire ignitions, land cover and population density maps. In this study, they took 42 land cover classes of Corine Land Cover 2000 cartography (IA, 2005) and grouped them into five broader classes:

i. agricultural areas

ii. forests

iii. shrublands and natural grasslands

iv. urban and other "artificial" areas

v. wetlands and water bodies

Catry et al. (2007) concluded that 85 % (98 %) of fire ignitions occurred within 0.5 km (2.0 km) from urban areas, and that 70 % (98 %) of fire ignitions occurred within 0.5 km (2.0 km) from major roads both "urban" and "rural". They also found that 25 % of ignitions were located in social/urban areas, 60 % in agriculture areas, 8.5 % in forested areas, and 6.5 % in other uncultivated areas. We show the spatial distribution for three classes of land use – "predominantly urban", "moderately-urban", and "predominantly rural" – as defined by Statistics Portugal (2011) in Fig. 3, which presents a very similar distribution pattern to the population density map (Fig. 2).

The causes of each fire were determined empirically for 1980–1988 by firemen whenever possible and, only for large fires, confirmed by the Forest Guards. In 1989, one brigade of the Forest Guard started to investigate the cause of fires more in depth, particularly for the central region of Portugal, mainly in the district of Coimbra. The number of brigades attributed to this task increased rapidly: 7 brigades in 1990, 22 in 1991 and 30 in 1992. The AFN (2011) now define 69 distinct fire causes in their database, broadly classified as (i) human (accidental, illegal, prescribed burn), (ii) natural
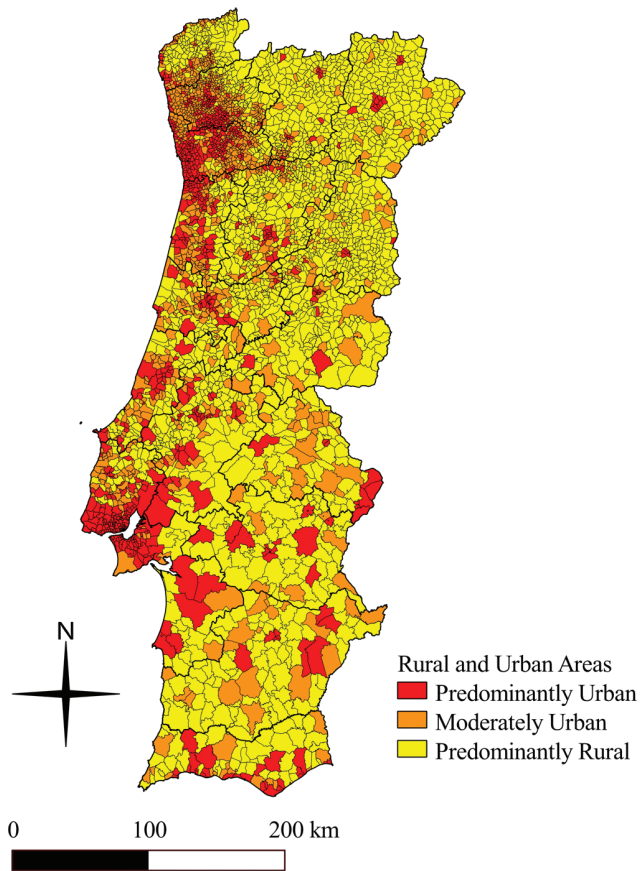
Fig. 3. Urban area types in Portugal (predominately urban, moderately urban, predominately rural) given spatially by principal administrative regions: 18 districts (see Fig. 2 for names), 278 municipalities (counties), and 4050 parishes in 2008. Data from Statistics Portugal (2011), who give detailed definitions for the different urban types and shape file from IGEO (2011).

(lightning), (iii) undetermined. As expected when dealing with tens of thousands of small events, the large majority of fires have an "unknown" cause (92.5 % of all fire records in the entire PRFD dataset, 1980–2005), as a result of the lack of investigation performed by the Forest Guard and unavailability of useful information. The amount of "unknown" causes varies substantially annually, between a minimum of 73.5 % (in 1983) and a maximum of 99.5 % (in 1993).

With 92.5 % of fires in the category "unknown" cause, the second most frequent fire cause (out of 22) in the 1980–2005 PRFD database is "reignition" (3.9 % of total records). Reignition fires are those that start in the same area affected by a recent fire, and are the result of an incomplete extinguishing of a previous fire. However "reignition" (17 790 records) was only recognized as a fire cause for the period 1994–2001. "Intentional" (2.1 %) and "negligence" (1.4 %) fire causes were identified in all years and present similar interannual variability, with higher values in the 1983–1985, 1994–1997 and 2000–2005 periods. The "natural" causes were less fre-

quent (0.1 % average overall) with the two largest extreme values in 1992 (0.9 % of total records in that year) and in 2003 (0.4 % of total records in that year).

Some countries have a relatively large number of "natural" fires caused by lightning, e.g. northwestern USA, Canada, Russia (Pyne, 2001; Wotton and Martell, 2005; Dickson et al., 2006). In contrast, Portugal has only a small percentage of fire records caused by lightning, as the country has a very small density of lightning flashes in the summer, when most of the fire activity occurs (Rivas Soriano and de Pablo, 2002; Tomás et al., 2004; PNDFCI, 2005; Ramos et al., 2011). The two yearly maximum values of fires caused by lightning in 1992 and 2003, noted above, could be due to abnormal high lightning activity in those two years. In fact, Tomás et al. (2004) state that the number of flashes in the Iberian Peninsula during July and August of 1992 (>330 000 flashes) was considerably higher than for the corresponding months of 1993 and 1994 (<240 000 flashes). In the first days of August 2003, Mendes (2003) reported that there was abnormally high lightning activity associated with dry (without rain) thunderstorms; more than 1000 flashes were observed 17:00–21:00 on 1 August 2003, at the same time period when high rural fire activity was observed.

4 Error identification and correction in the PRFD

In order to detect errors and implement correction procedures in the 1980–2005 AFN dataset, we use each rural fire record's date (time) and burnt area (A_F). We also examine a number of additional measures based on the date and A_F values, namely the median (50th percentile) of duration times for rural fires with the same A_F, and the amount of burnt area per unit time.

The original dataset underwent a sequence of procedures to correct data inconsistencies with the increasing order of complexity. Six types of data inconsistencies were found and corrected:

i. Records with zero burnt area ($A_F = 0$ ha).

ii. "Repeat" records, with the same date/time and spatial location, and assumption that only one fire is correct.

iii. Data errors in time (format time errors, negative duration in time errors).

iv. Missing information (date/time, parish names).

v. Multiple records, with similar date/time and spatial location, but assumed to be all pertaining to one larger fire.

vi. Other suspicious records based on area and duration time information.

In Table 3 we give a summary of the types of errors and their frequency. Each single record may be affected by more than

Table 3. Types and numbers of record errors we identified in the 1980–2005 Portuguese Rural Fire Database (PRFD). Total number of values in the original "raw" PRFD is 467 711 records.

Types of record errors	N_F in dataset examined*	number of errors	% of dataset examined
Zero burnt area	467 711	8908	1.90
Repeat records	467 711	5074	1.08
Format errors	453 729	1496	0.33
Negative duration	453 729	3092	0.68
Missing date/time	453 729	2870	0.63
Multiple records	453 729	152	0.03
Missing Parish names	453 577	17 782	3.92

* From the original 467 711 dataset, the zero burnt area and repeats were removed, giving 453 729 records. In the subset remaining (97 % of the original database), six kinds of other inconsistencies were identified and corrected/removed.

one type of error. Therefore, changing a record to correct one of these inconsistencies may solve other discrepancies. For example, after correcting "repeated" record errors, the number of records with "negative duration" decreased from 5786 to 3092. We now discuss each kind of error identified in the AFN rural fire database for Portugal for the entire period analysed, 1980–2005. We give sufficient detail here for each type of error, so that those who work with other large complex fire databases might better identify similar error types in their own datasets.

4.1 Zero burnt area

The original dataset contained a small, but significant percentage (1.9 %) of fire records with zero burnt area ($A_F = 0$ ha). This is due to administrative reporting procedures whereby (i) those fires that were quickly detected and rapidly suppressed and therefore characterized by extremely small values of burnt area, and (ii) false alarms that had to be reported.

4.2 Repeat records

A small (1.1 %) but significant source of error in the 1980–2005 database is related to the existence of two or more fires that occurred at the same location with precisely the same ignition/extinction date and time, but with two distinct values of burnt area, often one much larger than the other. It is assumed that these "repeated" records corresponded to the same fire. We therefore deleted one record and added its burnt area onto the other. There were 5074 records (out of 467 711) identified with this type of error (Table 3). This problem was particularly prominent during 2003, when about 3100 (60 %) of these repetitions were detected and corrected.

Another type of repeat error is the possibility that two or more records do not refer to independent or single fires, but to the same fire. In fact, there are multiple records with the

same ignition date, with similar ignition time and with approximately the same location (e.g. the same parish). Different criteria were applied spatially and temporally to identify these records, with the final number of "repeats" found depending on the magnitude of the criteria.

4.3 Format errors

After removing the zero and repeat values, the database presents a small number of format errors (1496 out of 453 729 records, or 0.3 %), on ignition/extinction time, requiring a relatively simple correction procedure. This error only affects records prior to 1993 and is associated with the fact that the ending and/or starting time were written beyond 23:59:59, for example, 24:30:00 which, in fact is 00:30:00 of the following day. We also detected the use of different formats to write dates or times. These rare situations (on the order of one hundred records) were corrected and saved so that there was internal consistency in the database.

4.4 Negative duration in time

There were 3092 out of 453 729 records (0.7 %) with negative rural fire duration. These errors were corrected and substituted by the median value (50th percentile) of all duration times for other rural fires with the same A_F. We made the assumption that the ignition time was the "correct" one, and added the 50th percentile duration time to this to arrive at a new extinction time (and date).

4.5 Missing date/time

Missing values on ignition and/or extinction time and on extinction date represent 2870 of the 453 729 records (0.6 %) (Table 3) and were corrected and substituted by the median value (50th percentile) of all duration times for other rural fires with the same A_F. Records with missing information were only detected during the years 1992–2005, with 1590 (55 %) of the 2870 records with missing date/time information detected in just two years (1995 and 1996).

4.6 Multiple records

The fire location is determined by the parish of ignition. However, for 2001–2005, whenever a large fire ($A_F > 100$ ha) crossed over a county or district boundary, the total amount of burnt area was divided by the counties/districts affected by the fire and multiple records were generated. Therefore, two or more records referring to the same rural fire were created. These cases of multiple records have one record with a much higher value of total burnt area than the other associated with it. These smaller fires may be secondary outbreaks of fires caused by the projection or transport of incandescent material. Although there is no possibility to ensure which record corresponds to the original fire,

these multiple records were merged into just one value, allocating the total area burnt (in forest and shrublands) to the record that had the highest burnt area value. This procedure minimizes the error on the burnt area on a district basis. There are a small number of multiple records (0.03 %) (Table 3) – 3 in 2001 and in 2002, 49 in 2003, 27 in 2004 and 97 in 2005 – for a total burnt area of 246 193 ha. Correction for 27 of the 179 multiple records was not possible due to missing or invalid information about the original fire.

4.7 Missing and inconsistent parish names

There are three levels of administrative boundaries in Portugal: district, county, and parish. In 17 226 out of 453 577 records (3.80 %), the parish names were missing or assumed as "other" and were changed to "UNKNOWN". In a total of another 556 records (0.12 %), the parish name was also set as "UNKNOWN", namely: (i) 346 records, where the fire ignition location was identified by the name of the locality instead of the parish name, and it was not possible to identify the corresponding parish; (ii) 174 records with the parish name equal to the county name and that a case does not exist; and, (iii) 36 records cases where parish/locality does not belong to the county/district and the correction was not possible. Nevertheless, analyses here were done only at the district level so these changes at the county and parish level (which we have flagged in our records) did not have any consequences on the analyses performed at district level. Naturally, parish names' inconsistencies would need to be taken into account if further analyses are to be done at the parish or county level.

Throughout the 26-yr period studied, districts remained the same; however, several parishes and counties that changed their boundaries also changed their names. Moreover some parishes and counties were annexed (in totality or partially) by neighbours and ceased to exist officially. In other cases, new parishes and counties were created. In this sense, 2471 of other inconsistencies in the parish names were corrected and flagged, specifically: (i) 2210 records where the parish/locality name was inconsistent with the county and district name; and, (ii) 261 of the cases where the name of the locality was provided instead of the parish, since it was possible to replace the locality name by the corresponding name of the parish.

4.8 Missing information on restart

Among all the original records considered as reignitions, only 6294 (5.5 %) contain useful information about the previous/original fire. This information is only available for the 2001–2005 period, but corresponds to the great majority (98.9 %) of the reignition records in this period. Statistics on this type of error were not included in Table 3 because they do not refer to the entire dataset as the information about the previous/original fire only refers to reignition fire records.

4.9 Suspicious records

Finally, we conclude this section with a discussion of a very small number of other "suspicious" records, based on outliers in burnt area or duration time. In particular, records characterized simultaneously (i) by small fire size and a very large fire duration or (ii) exceptionally large burnt area values during an extremely short time period interval. For example, on 15 July 1991 a fire in Arcos de Valdevez (district of Viana do Castelo), was recorded as having occurred over a period of 10 days with $A_F = 0.03$ ha. In another example, on 11 August 1991, a fire in Seia (district of Guarda) that burnt for just 1 h was responsible for $A_F = 1499$ ha.

Naturally, we acknowledge that identifying these "suspicious" records depends on subjective criteria, related to the amount of burnt area per unit time (i.e. the fire propagation). However, a considerable number of factors influence the fire propagation, including (Viegas, 1998): (i) the extension of the fire front; (ii) the local topography (with a uniform fuel bed, the rate of spread of a fire front propagating uphill increases with the slope); (iii) the type, condition and spatial arrangement of the vegetation; (iv) the meteorological conditions, mainly the wind and the vertical stability of the atmosphere. It is certainly acceptable to have rural fires with the same duration but quite different fire size, and fires with similar burnt areas but dissimilar duration. However, we are confident that extreme cases should be considered as suspect. A considerable proportion of these suspicious records is probably due to data entry errors, and is not real. For these reasons, these records have not undergone any treatment/correction. Table 3 does not include any estimates of suspicious records (and correspondent statistics) because such estimates (and correspondent statistics) will strongly depend on the different criteria used to classify a record as suspicious. In studies where the fire duration information plays an important role, these records have to be carefully analysed.

Finally, the database includes a heterogeneous description of the administrative boundaries in Portugal. The district, county, and parish names were homogenized by removing hyphens, replacing abbreviations by the full name, spelling, replacing characters and accent characteristic of the Portuguese language (e.g. ç, õ, à, é, etc.) and turning their names in uppercase. This correction resulted in changes in 54 378 district, 129 706 county and 161 235 parish names.

4.10 Modified PRFD (Supplement)

From the original rural fire dataset (467 711 values), and after all types of errors were corrected following the procedures described in Sect. 4, we have retained 97 % (453 577) fire events (Table 3), creating a modified PRFD, which we use in the rest of this paper. We have included the modified PRFD as a Supplement, which is freely available for others to use in their research.
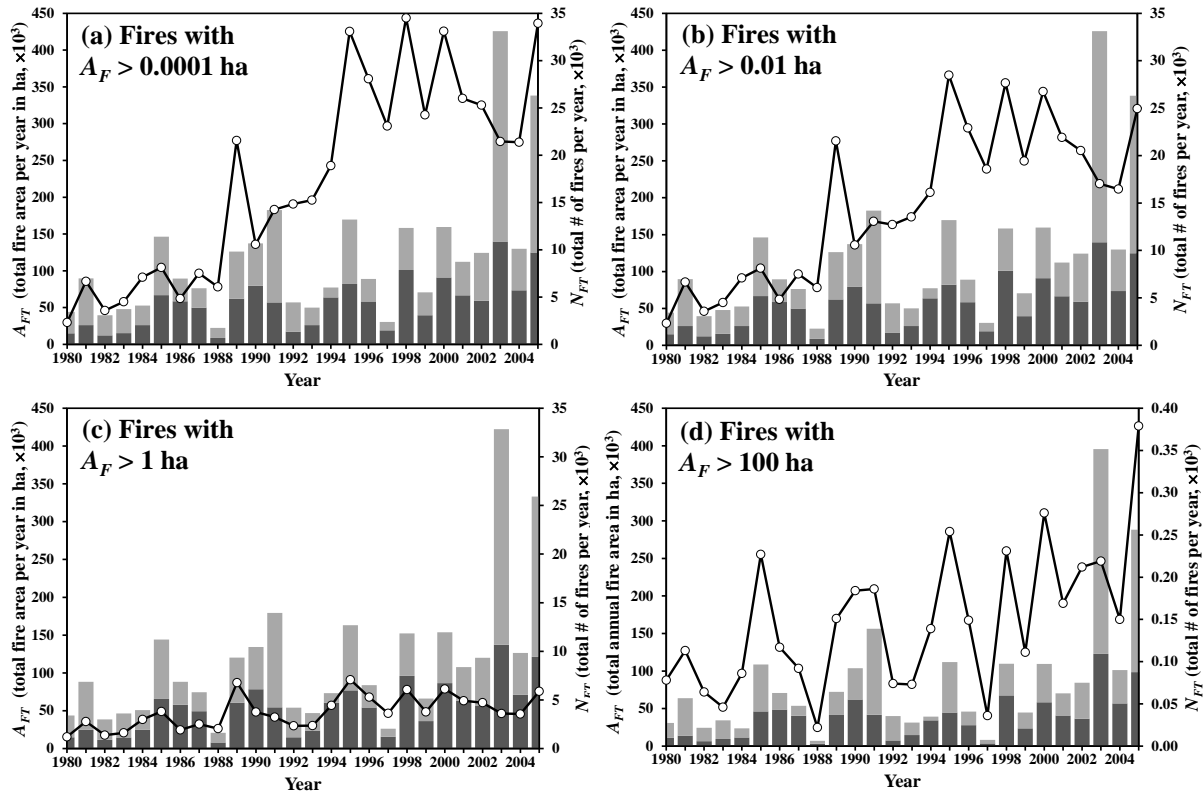
Fig. 4. Temporal evolution of the annual total number (N_{FT}, solid line) and area (A_{FT}, vertical bars) of Continental Portugal fire records. The vertical bars are subdivided into burnt area of forests (light grey bars) and shrublands (dark grey bars). Four lower thresholds are included: $A_F >$ **(a)** 0.0001 ha, **(b)** 0.01 ha, **(c)** 1 ha and **(d)** 100 ha. The A_{FT} (principal vertical) axis remains the same for all four figure parts, but the secondary vertical axes maximum value changes from $N_{FT} = 35\,000$ for **(a)**, **(b)** and **(c)** to $N_{FT} = 400$ for **(d)**, thus changing by a factor of about 100 the vertical exaggeration.

5 Temporal and spatial Portugal rural fire statistics

We now restrict our analysis to this modified PRFD and evaluate (i) the impact of changing characteristics of the dataset with time, such as changes in the minimum area burnt systematically detected and reported in the database, and (ii) the density of fires and fraction of burnt area as a function of administrative district area in Portugal. We are aware that changes in reporting practices for the minimum area burnt have taken place since 1980 and are bound to result in inconsistencies such as biased reporting of the total number of fires per year. Inconsistencies, both over time and spatially, have also been shown to be true for other countries. For example, Brown et al. (2002) highlighted some of these inconsistencies in a course assessment of 658 000 fire records from USA federal wildlands for 1970–2000.

5.1 Temporal statistics results

In Fig. 4 we show the total number of fires per year (N_{FT}) in mainland Continental Portugal, for 1980–2005, based on the corrected PRFD using fires with burnt area "lower"

thresholds: $A_F > 0.0001$, 0.001, 1 and 100 ha. The number of fire records in the database with $A_F > 0.0001$ ha (Fig. 4a, solid line with circle symbols) presents a roughly linear increase for the 1980–1994 period (from about 2500 to 20 000 records), but with an "anomalous high" of 22 000 records for 1989. This is followed by a roughly horizontal trend in N_{FT} for the 1995–2005 period, but with values going back and forth between 21 000/23 000 fires per year (lows) and 33 000/35 000 fires per year (highs). Similar behaviour in N_{FT} is seen for $A_F > 0.01$ ha (Fig. 4b). For $A_F > 1$ ha (Fig. 4c), when considered on its own, a distinct linear trend in N_{FT} as a function of year, would be difficult to visually pull out. When only those burnt areas with $A_F > 100$ ha are considered (Fig. 4d), a potential linear trend in N_{FT} over the entire period, 1980–2005, can be seen, compared to $A_F > 0.0001$, 0.01, 1 ha (Fig. 4a, b and c), which shows a trend only for 1980 to mid-1990s. This is probably because most fires with $A_F > 100$ ha were included in the PRFD from 1980 (i.e. the trend is "real" for 1980–2005); whereas, fires increasingly smaller than this were not included in the database until the mid-1990s (i.e. the trend
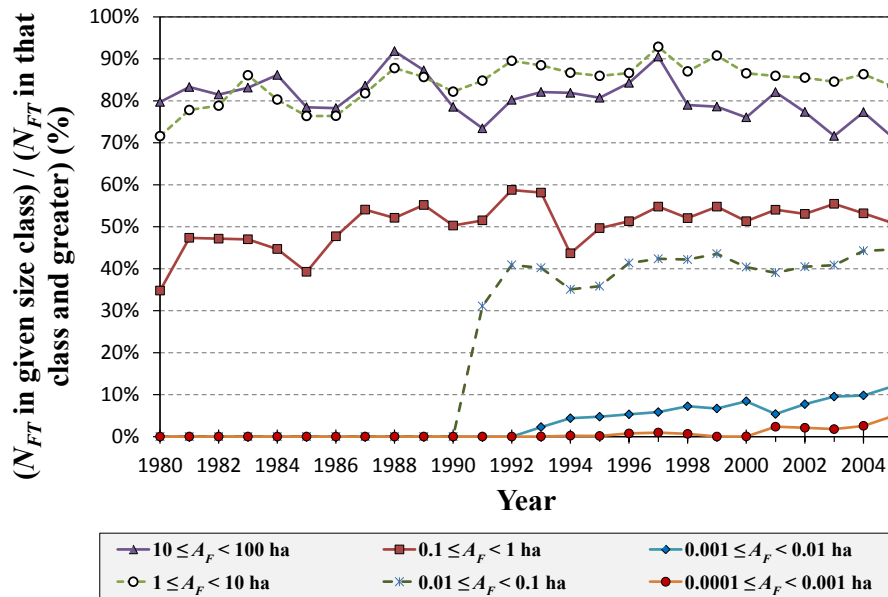
Fig. 5. Percentage of Continental Portuguese fires N_{FT}, in specified burnt area size class (see legend), divided by all fires in that burnt area class and greater. The N_{FT} (size class)/N_{FT} (size class and greater) is given as a function of year, 1980–2005.

in number of fires shown in Fig. 4a and b are not "real" for 1980–2005, but an artefact of underreporting of smaller burnt areas).

In Fig. 4 we also give the total rural fire burnt area A_{FT} per year, for $A_F > 0.0001$, 0.01, 1, 100 ha. Increasing the A_F lower limit (minimum threshold) from 0.0001 to 0.01 to 1 ha does not have considerable impact on the burnt area statistics A_{FT} for any given year. Only for $A_F > 100$ ha, do the values for A_{FT} noticeably diminish for most years. We also differentiate between area burnt in shrublands (dark grey bars), and forests (light grey bars). For $A_F > 0.0001$ ha (Fig. 4a) the proportion of burnt area per year amongst the two land-use types (forest and shrublands) was not constant. In the sub-periods 1980–1983, 1991–1992 and in 1988, 2003 and 2005, there was more burnt area in forested land areas than in shrublands, with a ratio higher than 60/40. On the other hand, during the sub-periods 1986–1987, 1996–1998 and in 1994, the proportion was inverted. Similar ratio changes can be seen for the other burn area reference values ($A_F > 0.01$, 1, 100 ha; Fig. 4b to d). For $A_F > 100$ ha (Fig. 4d) the inter-annual variability of N_{FT} tends to follow quite well the A_{FT} inter-annual variability with a few exceptions (e.g. 2003).

Although the total burn area per year, A_{FT}, does not decrease significantly as the lower bound increases, the number of fires per year, N_{FT}, decreases dramatically. This is because a relatively small number of large and very large fires are responsible for the majority of the burnt areas.

To examine the changing pattern of fire reporting over the period 1980–2005 in more depth, one method is to graph the number of fires in a given burnt area fire class (e.g. 1 ha

to 10 ha). Therefore, we examine in Fig. 5 the number of fires per year (N_{FT}) as a function of six burnt area classes: $0.0001 \leq A_F < 0.001$ ha, $0.001 \leq A_F < 0.01$ ha, $0.01 \leq A_F < 0.1$ ha, $0.1 \leq A_F < 1$ ha, $1 \leq A_F < 10$ ha, $10 \leq A_F < 100$ ha. So that we can better observe broad trends, potentially due to reporting practice, we divide, in Fig. 5, N_{FT} in given burnt area size class by N_{FT} in given burnt area size class or greater. We do not normalize by "all" fires for that year in all classes, because the number of smaller fires reported is increasing greatly year on year. We believe this measure will reduce some of the interannual variability observed in Fig. 4, but acknowledge that this percentage is also subject to natural variability.

Examining Fig. 5 over the 26 yr of records to see how percentages have changed with time for the number of fires in each fire size classes (as a percentage of fires in that class or bigger), particularly with respect to reporting practices, we see that there are no records in the PRFD with $A_F < 0.1$ ha ($A_F < 0.01$ ha) before 1990 (1992); further examination of the data show that less than 4 % of fires in the 1993–2000 period have $A_F < 0.001$ ha. For the three size classes, $0.0001 \leq A_F < 0.001$ ha, $0.001 \leq A_F < 0.01$ ha, $0.01 \leq A_F < 0.1$ ha, from 1992–2005, a gentle trend upwards is observable in Fig. 5, indicating that more fires in these size classes are being reported. For the three burnt area size classes, $0.1 \leq A_F < 1$ ha, $1 \leq A_F < 10$ ha, $10 \leq A_F < 100$ ha, we see the percentages are approximately constant, although there are slightly positive and negative trends evident. However, for the size class $0.1 \leq A_F < 1$ ha, the period 1980–1985 is on average about 10 % lower than 1986–2005, indicating
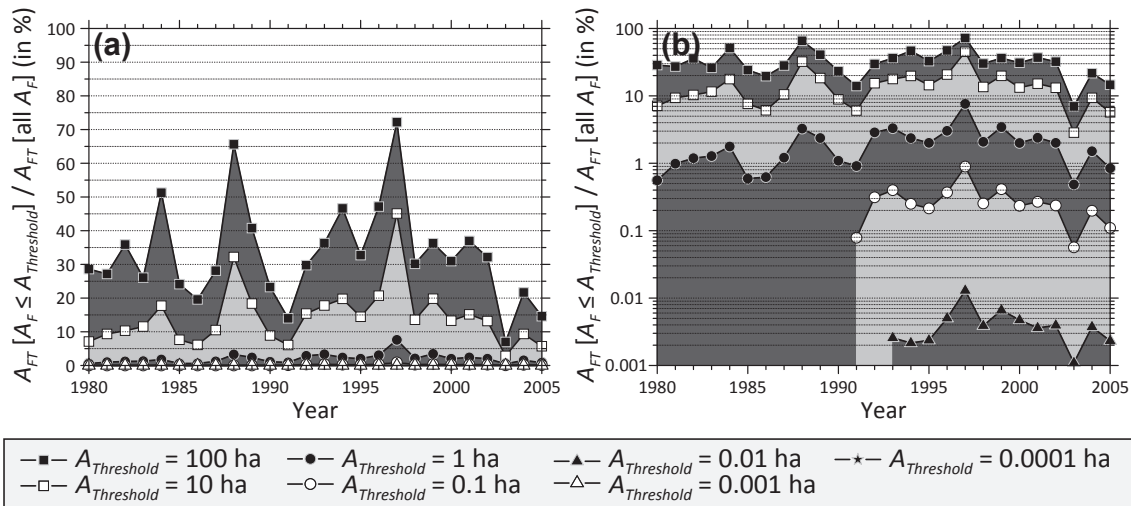
Fig. 6. Percentage of Continental Portugal yearly burnt area (A_{FT}) using just those burnt area fire records $A_F \leq A_{Threshold} = 0.0001, 0.001, 0.01, 0.1, 1, 10$ and 100 ha, each given as a percentage of the corresponding total values in each year, 1980–2005: $A_{FT} [A_F \leq A_{Threshold}] / A_{FT}$ [all A_F], is given for **(a)** linear axis and **(b)** log axis.

that the early to mid-1980s might have had reporting practices which underestimated fires $0.1 \leq A_F < 1$ ha to a small degree.

The results shown in Fig. 5 are further evidence that fire sizes were not measured down to the same value throughout the years and that dataset completeness was substantially different in the 1980s compared to the 1990s and the present. Therefore, extreme care must be taken when using data from earlier years in the PRFD 1980–2005 data, and only records "above" a minimum value should be included when comparing earlier years (e.g. the 1980s) to later years (e.g. the 2000s).

In Fig. 6a (linear y-axis) and Fig. 6b (log y-axis), we present the annual percentage of total burnt fire area (A_{FT}) for increasing values of a burnt area threshold, $A_{Threshold}$. This is graphing the same data as that shown in Fig. 4, but now as a percentage and as a function of year, so that we can better observe any changes over the 26-yr period being examined. Log axes are used as the y-axis data ranges over 16 orders of magnitude (of which only 5 orders are shown in Fig. 6b). For example, in Fig. 6a and b if we take $A_{Threshold} = 10$ ha (white squares), then we see that in 1980, 8 % of the reported total burnt area for that year (A_{FT} for all A_F) was recorded as due to fires $\leq A_{Threshold} = 10$ ha, and (100–8) % = 92 % of the total area due to fires > $A_{Threshold} = 10$ ha. As expected, the relative maximum and minimum values for different values of the $A_{Threshold}$ correspond well to relative extreme values of total area per year, A_{FT}, shown in Fig. 4. Overall, as smaller fires contribute only a small amount to the total area burnt per year, the systematic underreporting of smaller fires in the 1980s and 1990s that we found in Fig. 6a and b does not overly effect the total burnt area per

year. Examining the 26 yr overall, we find that fires with $A_F \leq 100$ ha, 10 ha, 1 ha, 0.1 ha, contribute (respectively) on average, 33 %, 14 %, 2 % and 0.3 % of the area burnt by all fire records in the dataset.

Based on Figs. 4 to 6, we conclude that (a) the minimum burnt area for fires recorded 1980–2005 has in later years changed toward smaller values, (b) this has influenced dramatically the number of fires recorded per year, such that one must be careful to only consider those fires above a given value when comparing early years with later years, (c) this has not influenced very much the total area reported per year, as smaller fires do not contribute greatly to the overall burnt area. Finally, in addition to taking care with comparing the number of fires per year for early (e.g. 1980s) vs. later (e.g. 2000s) years in the period under consideration, likewise, A_{FT}/N_{FT}, or "average" burnt area is likely to be strongly biased and incorrect. Based on the results of Figs. 4 to 6 to broadly determine completeness of the data for different decades, we have restricted the remaining analysis to fires with $A_F \geq 0.1$ ha (a total of 286 751 records) to ensure an appropriate comparison of statistics for the whole period of record, 1980 to 2005. In comparison, Malamud et al. (2005), when examining wildfire statistics in US Forest Service lands for the conterminous United States over the time period 1970–2000, found similar behaviour in that database, such that smaller fires were under-reported in the earlier years of the record. They therefore took $A_F \geq 0.4$ ha as a lower threshold for completeness.

5.2 Fire statistics as a function of Continental Portugal district

In Fig. 7, we consider for the Continental Portugal district, the entire period 1980–2005 and only for records with $A_F \geq 0.1$ ha, the total number of rural fires (N_{FT}), and area burnt (A_{FT}) in forests, shrublands, forests + shrublands. In each case the variables are normalized by the district area (A_D), resulting in "densities" of rural fires burnt (per km^2) and percentage (%) of the total geographic area burnt (i.e. ha per km^2).

The highest density of fire records are found in the northwestern part of the country, in the highly populated districts of Porto, Braga, and Viana do Castelo, but also the districts of Lisboa, Vila Real, Viseu and Guarda present relatively high values (Fig. 7a). The burnt area in forested areas (Fig. 7b) presents relatively high values in the central and northern parts of the territory, following the spatial distribution of the larger parcels of continuous coniferous forests (Fig. 7b; Pereira and Santos, 2003). These continuous forest areas coincide roughly (Fig. 2) with predominately rural and moderately urban areas. Shrublands are most affected by rural fires in the mountainous districts of Guarda, Viseu and Vila Real, where this type of vegetation is more abundant (Fig. 7c). Finally, the spatial distribution of the total area affected by rural fires (Fig. 7d) resembles the result of the combination of the burnt areas in forests (Fig. 7b) and in shrublands (Fig. 7c), with higher values in the districts of Guarda and Coimbra, precisely the districts with higher burnt areas in shrublands and in forests, respectively.

6 Summary and discussion

The datasets used by geophysical and environmental research communities may be classified into two broad classes: (i) datasets that have undergone several levels of data treatment procedures to correct the different types of errors and evaluate a dataset's potential to be used in specific types of analyses; (ii) datasets in mostly "raw" format. Examples of a dataset that falls into the first classification (various levels of data treatment) include the meteorological re-analysis databases of NCEP/NCAR (Kalnay et al., 1996), where there is a quality control implemented in the model and data assimilation system, and also the European Climate Assessment database (Klein Tank et al., 2002), where statistical procedures are applied in order to test the homogeneity of the database. The Portuguese Rural Fire Database (PRFD) that we have examined is an example of a dataset that falls into the second class, i.e. that of essentially "raw" data.

With this work we have made an effort to provide a comprehensive description of the AFN Portuguese rural fire dataset. This is one of the largest rural fire databases in Europe and is based exclusively on local ground measurements (no satellite data). Our first aim was to describe the errors and inconsistencies in this dataset and to assess its limitations and potential. Missing values and different formats were reported while the procedures to detect and correct the records affected by errors were described. All the procedures followed to detect and correct errors in the dataset were based on information available in the dataset itself. Three categories of data inconsistencies were identified: assorted formats, data errors and suspicious records. Records affected by date/time errors or with parish name inconsistencies were not immediately excluded from the dataset; instead, a quality flag was set for all records with identified errors and set to zero for all remaining reports. Nevertheless, only a relatively small percentage of records were flagged. It should be recognized that without additional information, namely satellite data, there are few additional robust criteria to employ in order to detect and correct the errors in this kind of dataset. The Portuguese Rural Fire Database (PRFD) covering the period 1980–2005, provided by the Portuguese Forest Service (AFN, 2011) and modified in accordance with the procedures described in this paper, is available to interested researchers as a Supplement. The package includes two text files: (i) readme, (ii) the 453 577 fire records tab delimited for the 29 different variables.

The total number of fires per year (N_{FT}) and of burnt area per year (A_{FT}) for fires with burnt area (A_F) above a given threshold presents a general linear increase with time due to higher values in the last years of the 1980–2005 period (Fig. 4). As the burnt area threshold increases (from 0.0001 ha to 100 ha), the inter-annual variability of the total burnt area per year (A_{FT}) and the ratio between shrublands and forest burnt areas, in each year, remains essentially constant while the yearly values of N_{FT} tend to resemble the total burnt area per year (A_{FT}) time series (Fig. 4).

Similarly to what can be observed over many other ecosystems of the world, smaller rural fires are significantly more frequent than larger fires; however, the vast majority of total burnt area is due to the few larger fires. In fact, the fire records with $A_F < 1$ ha ($A_F < 10$ ha) represent about 70% (95%) of all fire records in the PRFD but only accounts for 1.5% (10%) of total burnt area. In addition, the PRFD presents an asymmetry in the temporal distribution of smaller fires, mainly because the minimum value of A_F reported in the database decreases during the 26 yr of data: fires down to sizes of $A_F = 0.1$ ha are systematically reported before 1991, compared to down to $A_F = 0.01$ ha in 1991–1992 and $A_F = 0.0001$ ha in the last 1992–2005 period (Fig. 5). This decrease is most likely evidence that smaller fires were reported less consistently from 1980–1990 compared to 1991–2005 due to better recording procedures. As a consequence, the proportion of the total number of fires N_{FT} in a given burnt area size class, with respect to fires burnt in all size classes (not just that class and above), will decrease in time, as smaller fires are reported.

When considering Continental Portugal yearly burnt areas at different thresholds, the inter-annual variability of the
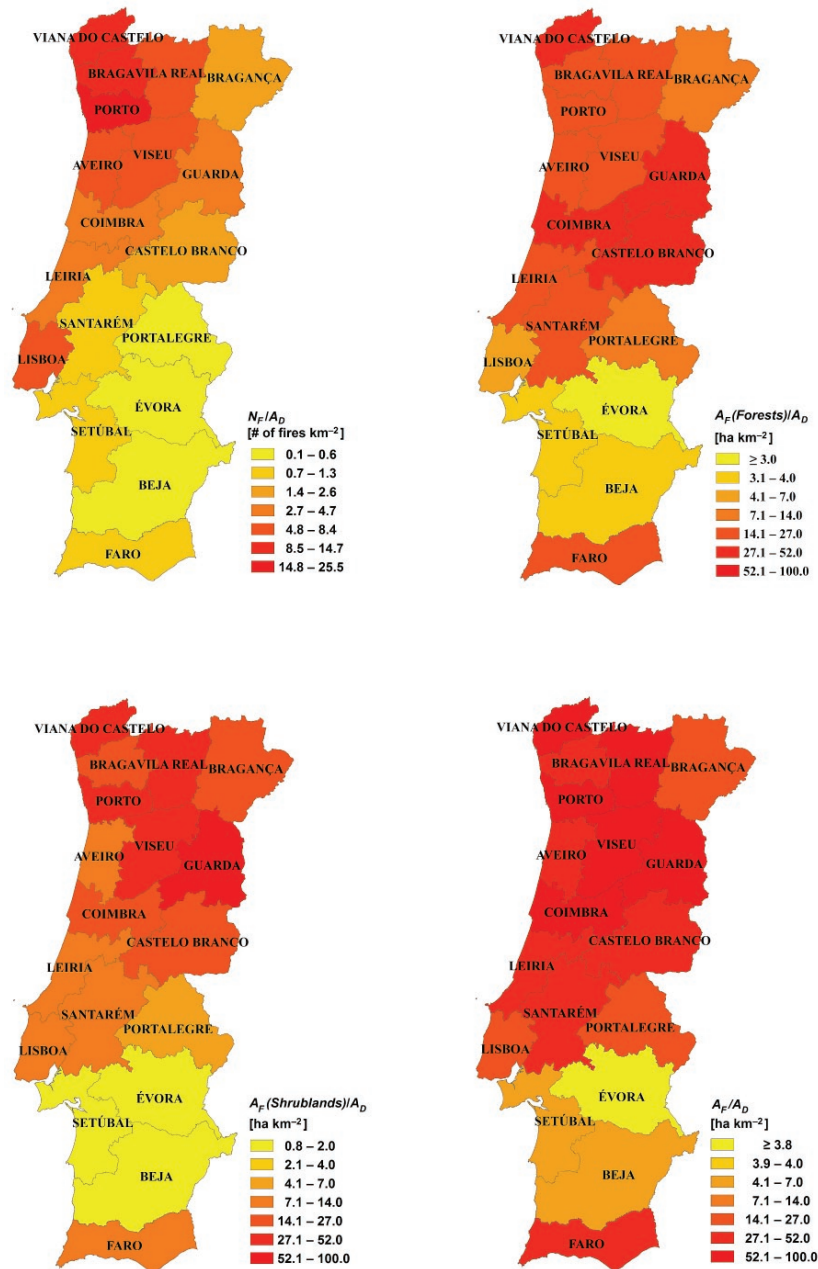
Fig. 7. Fire numbers and burnt areas in Continental Portugal districts normalized by each district's area, A_D (km^2). Given in **(a)** are the normalized number of fire records, N_{FT}/A_D (in # of fires per km^2). In **(b)** to **(d)** are given the normalized total burnt areas, A_{FT}/A_D (in ha per km^2), with a logarithmically increasing legend, for **(b)** forests, **(c)** shrublands, **(d)** forest plus shrublands. In these analyses, fires included have burnt area $A_F \geq 0.1$ ha and are for the 1980–2005 period.

proportion of A_{FT} by fires with A_F less than or equal to a given threshold seems to resemble the variability of A_{FT} (Figs. 4 and 6). All these results about the dataset completeness allow us to question the meaning of the fire statistics based on the number of fires and supports the decision to restrict the remaining exploratory analysis to fires with $A_F \geq 0.1$ ha (286 751 records).

In Fig. 7, we consider for each Continental Portugal district and for the 26-yr period, the total number of rural fires and area burnt in forests and shrublands, each normalized by district areas. We find that the highest numbers of fires per unit area are in highly populated districts, but also in the less-populated districts of Vila Real, Viseu and Guarda (Figs. 2 and 7). We also find that the burnt area in forested areas

follows the spatial distribution of larger parcels of continuous coniferous forests (predominantly rural and moderately urban areas). Evaluating the spatial distribution of the density of fires and burnt area based on fires with size $A_F \geq 0.1$ ha provides us with much more confidence in the results obtained, due to the completeness of the data used.

Besides revealing the most general characteristics of this unique and large dataset of Portuguese fires, this work intends to alert the fire research community to several structural problems with fire datasets that could compromise the results of their analysis. Knowing the history and the specificities of the entire data compilation procedure – from the detection procedures to the insertion of the fire records into the dataset – helps to become acquainted with the completeness and individual inaccuracy of each fire record. We may also conclude that some artificial misleading characteristics can be introduced in the datasets as a consequence of changing compilation procedures (e.g. minimum area detected). In this sense, we believe that this work could be a valuable contribution for those who are responsible for rural (wildland) fire data acquisition and assemblage in order to avoid leaving their signature in the database.

Supplementary material related to this article is available online at: http://www.nat-hazards-earth-syst-sci.net/11/3343/2011/ nhess-11-3343-2011-supplement.zip.

References

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

Barbosa, P., Amatulli, G., Boca, R., Camia, A., Kucera, J., Libertà, G., San-Miguel Ayanz, J., Schmuck, G., Schulte, E., and Dierks, H.-H.: 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, 79 pp., 2007.

Brown, T. J., Hall, B. L., Mohrle, C. R., and Reinbold, H. J.: Coarse Assessment of Federal Wildland Fire Occurrence Data, Report for the National Wildfire Coordinating Group, Program CEFA (Climate, Ecosystem and Fire Applications), Report 02-04, 30 pp., 2002.

Carvalho, A., Flannigan, M. D., Logan, K., Miranda, A. I., and Borrego, C.: Fire activity in Portugal and its relationship to weather and the Canadian Fire Weather Index System, Int. J. Wildland Fire, 17, 328–338, 2008.

Catry, F. X., Damasceno, P., Silva, J. S., Galante, M., and Moreira, F.: Spatial distribution patterns of wildfire ignitions in Portugal, in: Proceedings of the 4th International Wildland Fire Conference, 13–17 May 2007, UNISDR; FAO; EU, CD Rom (10 pp.), Seville, Spain, 2007.

Costa, L., Thonicke, K., Poulter, B., and Badek, F.-W.: 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, 543–551, doi:10.1007/s10113-010-0169-6, 2010.

DGF: Distribuição da floresta em Portugal Continental, Áreas florestais por distritos: informação disponível em 1992, Estudos e informação no. 303, Direcção Geral das Florestas, Lisbon, 1993.

DGF: Distribuição da floresta em Portugal Continental, Direcção Geral das Florestas, Lisbon, 1997.

DGF: Inventário florestal nacional, Portugal continental, 3a Revisão, 1995–1998, Direcção-Geral das Florestas, Lisbon, 2001.

DGF: Defesa da floresta contra incêndios, Direcção-Geral das Florestas, Lisbon, 2002.

DGRF: Sistema de Gestão de Fogos Florestais, Divisão de Defesa da Floresta Contra Incêndios, Direcção-Geral dos Recursos Florestais, Lisbon, 2005.

DGRF: Incêndios florestais – Relatório de 2005, Divisão de Defesa da Floresta Contra Incêndios, Direcção-Geral dos Recursos Florestais, Lisbon, 2006.

DGRF: Incêndios florestais – Relatório de 2006, Divisão de Defesa da Floresta Contra Incêndios, Direcção-Geral dos Recursos Florestais, Lisbon, 2007.

Dickson, B. G., Prather, J. W., Xu, Y., Hampton, H. M., Aumack, E. N., and Sisk, T. D.: Mapping the probability of large fire occurrence in northern Arizona, USA, Landscape Ecol., 21, 747–761, 2006.

DL: Decreto-Lei n° 327/80 (Law-decree n° 327/80). Diário da República, N° 196, Serie I, 1980-08-26, available at: http://www.dre.pt (last access: 1 June 2009), 1980.

DR: Decreto Regulamentar n° 55/81 (Regulation-decree n° 55/81). Diário da República, N° 290, Serie I, 1981-12-18, available at: http://www.dre.pt (last access: 1 June 2009), 1981.

ECE/FAO Agriculture and Timber Division: Forest fire statistics, 1999–2001, Timber Bulletin (ECE/FAO), Bulletin du Bois (CEE/FAO), 0259-4323, (United Nations, Economic Commission for Europe, Geneva), 55, 21 pp., 2003.

FAOSTAT: Food and Agriculture Organization of the United Nations statistical database, available at: http://faosat.fao.org/, last access: 1 April 2011.

Fernandes, P. A. M.: Fire spread prediction in shrub fuels in Portugal, Forest Ecol. Manage., 144, 67–74, 2001.

Hoinka, K. P., Carvalho, A., and Miranda, A. I.: Regional-scale

weather patterns and wildland fires in central Portugal, Int. J. Wildland Fire, 18, 36–49, 2009.

IA (Institute for the Environment): CORINE Land Cover 2000 Portugal, 2005.

IGEO (Instituto Geográfico Português): available at: http://www.igeo.pt/, last access: 15 April 2011.

JRC: Forest Fires in Europe 2009, Report EUR 24502 EN – 2010, JRC-IES / Land Management & Natural Hazards Unit, ISSN 1018-5593, available at: http://effis.jrc.ec.europa.eu/reports/fire-reports/ (last access: 1 April 2011), 2010.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Wollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, Jenne, R. and Joseph, D.: The NCEP/NCAR 40-years reanalyses project, B. Am. Meteorol. Soc., 77, 437–471, 1996.

Klein Tank, A. M. G., Wijngaard, J. B., Können, G. P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., Szalai, S., Pálsdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., van Engelen, A. F. V., Forland, E., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Antonio López, J., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L. V., and Petrovic, P.: Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment, Int. J. Climatol., 22, 1441–1453, 2002.

Malamud, B. D., Millington, J. D. A., and Perry, G. L. W.: Characterizing wildfire regimes in the U.S.A., Proc. Nat. Ac. Sci., 102, 4694–4699, 2005.

Mendes, C.: Incêndios florestais e onda de calor – Análise do período entre 27 de Julho e 15 de Agosto de 2003, Serviço Nacional de Bombeiros e Protecção Civil, Divisão de Riscos Naturais e Tecnológicos, unpublished, Lisbon, Portugal, 2003.

Meteorological Institute: Climate Profile, Mainland Portugal, available at: http://www.meteo.pt/pt/areaeducativa/otempo.eoclima/clima.pt/index.html, last access: 1 May 2011.

NFI5 (The 5th National Forestry Inventory): Attributes of the Portuguese forest, available at: http://www.afn.min-agricultura.pt/portal/ifn, last access: 20 July 2011.

Pereira, J. M. C. and Santos, T. N.: Fire Risk and Burned Area Mapping in Portugal, Direcção Geral das Florestas, Lisboa, 2003.

Pereira, J. M. C., Carreiras, J. M. B., and Vasconcelos, M. J. P.: Exploratory data analysis of the spatial distribution of wildland fires in Portugal, 1980–1989, Geogr. Syst., 5, 355–390, 1998.

Pereira, M. G., Trigo, R. M., DaCamara, C. C., Pereira, J. M. C., and Leite, S. M.: Synoptic patterns associated with large summer forest fires in Portugal, Agr. Forest Meteorol. 129, 11–25, doi:10.1016/j.agrformet.2004.12.007, 2005.

PNDFCI: Plano Nacional Defesa Floresta Contra Incêndios – Relatório intercalar, Estudo Técnico I: Diagnóstico, Visão e Objectivos Estratégicos, available at: http://www.isa.utl.pt/pndfci/ (last access: 1 March 2011), 2005.

Pyne, S. J.: Fire: A Brief History, University of Washington Press, 204 pp., 2001.

Ramos, A. R., Ramos, R., Sousa, P., Trigo, R. M., Janeira, M., and Prior, V: Cloud to ground lightning activity over Portugal and its association with circulation weather types, Atmos. Res., 101, 84-101, 2011.

Rivas Soriano, L. and de Pablo, F.: Study of lightning event duration and flash rate in the Iberian Península using cloud-to-ground lightning data, Atmos. Res., 61, 189–201, 2002.

Silva, J. S. and Catry, F.: Forest fires in cork oak (Quercus suber L.) stands in Portugal, Int. J. Environ. Stud., 63, 235–257, 2006.

Statistics Portugal: Statistical year book of Portugal, (Portugal National Institute of Statistics: Lisbon), available at: http://www.ine.pt (last access: 7 March 2008), 2007.

Statistics Portugal: Territorial structure website, (Portugal National Institute of Statistics: Lisbon), available at: http://sig.ine.pt/ (last access: 7 March 2011), 2011.

Tomás, C., de Pablo, F., and Soriano L.R.: Circulation weather types and cloud-to-ground flash density over the Iberian Peninsula. International Journal of Climatology 24, 109-123, 2004.

Trigo, R. M. and DaCamara, C. C.: Circulation Weather Types and their impact on the precipitation regime in Portugal, Int. J. Climatol., 20, 1559–1581, 2000.

Trigo, R. M., Pereira, J. M. C., Pereira, M. G., Mota, B., Calado, T. J., DaCamara, C. C., and Santo F. E.: Atmospheric conditions associated with the exceptional fire season of 2003 in Portugal, Int. J. Climatol., 26, 1741–1757, doi:10.1002/joc.1333, 2006.

Vasconcelos, M. J. P., Silva, S., Tomé, M., Alvim, M., and Pereira, J. M. C.: Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks, Photogr. Eng. Remote Sens., 67, 73–81, 2001.

Viegas, D. X.: Forest fire propagation, Phil. Trans. R. Soc. Lond. A, 356, 2907–2928, 1998.

Wotton, B. M. and Martell, D. L.: A lightning fire occurrence model for Ontario, Can. J. For. Res., 35, 1389–1401, 2005.

Nat. Hazards Earth Syst. Sci., 11, 3343–3358, 2011

www.nat-hazards-earth-syst-sci.net/11/3343/2011/