



1 The Portuguese Large Wildfire Spread Database (PT-FireSprd)

2 Akli Benali¹, Nuno Guiomar², Hugo Gonçalves^{3,4}, Bernardo Mota⁵, Fábio Silva^{3,4}, Paulo M. Fernandes⁶,
3 Carlos Mota^{3,4}, Alexandre Penha⁴, João Santos^{3,4}, José M.C. Pereira^{1,7}, Ana C.L. Sá¹

4 ¹Centro de Estudos Florestais, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa,
5 Portugal

6 ²MED - Mediterranean Institute for Agriculture, Environment and Development; CHANGE - Global Change and
7 Sustainability; EaRSLab - Earth Remote Sensing Laboratory Institute; IIFA - Institute for Advanced Studies and Research;
8 University of Évora, 7006-554 Évora, Portugal

9 ³Força Especial de Proteção Civil, 2080-221 Almeirim, Portugal

10 ⁴Autoridade Nacional de Emergência e Proteção Civil, 2799-51 Carnaxide, Portugal

11 ⁵National Physical Laboratory (NPL), Climate Earth Observation (CEO), Hampton Rd. Teddington, TW11 0LW, UK

12 ⁶CITAB - Centro de Investigação e de Tecnologias Agro-Ambientais e Biológicas, Universidade de Trás-os-Montes e Alto
13 Douro, 5001-801 Vila Real, Portugal;

14 ⁷Laboratório Associado TERRA, Tapada da Ajuda, 1349-017 Lisboa, Portugal

15
16 *Correspondence to:* Akli Benali (aklibenali@gmail.com)

17 **Abstract.** Wildfire behaviour depends on complex interactions between fuels, topography and weather, over a wide range of
18 scales, being important for fire research and management applications. To allow for a significant progress towards better fire
19 management, the operational and research communities require detailed open data on observed wildfire behaviour. Here, we
20 present the Portuguese Large Wildfire Spread Database (PT-FireSprd) that includes the reconstruction of the spread of 80 large
21 wildfires that occurred in Portugal between 2015 and 2021. It includes a detailed set of fire behaviour descriptors, such as rate-
22 of-spread (ROS), fire growth rate (FGR), and fire radiative energy (FRE). The wildfires were reconstructed by converging
23 evidence from complementary data sources, such as satellite imagery/products, airborne and ground data collected by fire
24 personnel, official fire data and information in external reports. We then implemented a digraph-based algorithm to estimate
25 the fire behaviour descriptors and combined it with MSG-SEVIRI fire radiative power estimates. A total of 1197 observations
26 of ROS and FGR were estimated along with 609 FRE estimates. The extreme fires of 2017 were responsible for the maximum
27 observed values of ROS (8956 m/h) and FGR (4436 ha/h). Combining both descriptors, we defined 6 fire behaviour classes
28 that can be easily communicated to both research and management communities and support a wide number of applications.
29 Analysis also showed that the area burned by a wildfire is mostly determined by its FGR rather than by its forward speed.
30 Finally, we explored a practical example to show the PT-FireSprd database can be used to study the dynamics of individual
31 wildfires and build robust case studies for training and capacity building.

32 The PT-FireSprd is the first open access fire progression and behaviour database in Mediterranean Europe, dramatically
33 expanding the extant information. Updating the PT-FireSprd database will require a continuous joint effort by researchers and
34 fire personnel. PT-FireSprd data are publicly available through <https://doi.org/10.5281/zenodo.7495506> (last access: 30th



35 December 2022) and have a large potential to improve current knowledge on wildfire behaviour and support better decision-
36 making (Benali et al. 2022).

37

38 **Keywords:** fire behaviour; satellite; airborne; ground; rate of spread; fire radiative energy; graphs; progression

39



40 **1 Introduction**

41 Wildfire behaviour is broadly defined as the way a free-burning fire ignites, develops and spreads through the landscape (Albini
42 1984; Rothermel 1972). It depends on complex interactions between fuels, topography and weather, over a wide range of
43 temporal and spatial scales (Santoni et al., 2011; Countryman, 1972). Wildfire behaviour can be described using common
44 metrics such as the spread rate, propagation mode, area growth rate, perimeter, rate of energy release and flame size (Albini
45 1984). Fire behaviour information is important for fire research and management applications (Finney et al., 2021).

46
47 To allow for a significant progress towards better fire management, the operational and research communities require detailed
48 open data on observed wildfire behaviour (Gollner et al., 2015). In this context, systematic mapping of the fire front progression
49 through space and time is critical to address existing needs, for wildfires burning under a wide range of environmental
50 conditions, including extreme ones (Storey et al., 2021; Gollner et al., 2015). Compiling quality fire behaviour information is
51 paramount to develop reliable and well-suited fire spread models and for a much-needed extensive evaluation of fire behaviour
52 predictions, which is crucial for its ultimate aim: support the decision-making process (Alexander and Cruz, 2013a; Scott and
53 Reinhardt, 2001). This includes planning pre-suppression activities and defining resources dispatch to wildfires, delineating
54 safe and effective fire suppression strategies and tactics during a wildfire, and for early alert and evacuation purposes (Finney
55 et al., 2021). Comprehensive fire progression and behaviour information is also useful to develop burned area/fire perimeter
56 mapping algorithms (Valero et al., 2018), understand fire effects (Collins et al., 2009), fire danger rating (Parisien et al., 2011),
57 fire hazard mapping and risk analysis (Alcasena et al., 2021, Palaiologou et al., 2020), planning and implementation of
58 preventive fuel treatments (Salis et al., 2018), and also to foster robust training of operative personnel and researchers
59 improving their learnings from past wildfires (Alexander and Thomas, 2003). Unfortunately, reliable quality information on
60 the progression and behaviour of wildfires, especially those burning under extreme conditions, is difficult to collect (Gollner
61 et al., 2015).

62
63 Fire behaviour data can be collected from laboratory experiments, experimental fires, prescribed fires or wildfires. A large
64 number of laboratory-scale experiments have been made for the development of semi-empirical rate-of-spread (ROS) models
65 (Rothermel 1972; Catchpole et al., 1998). Experimental fires have been set up to collect fireline data, estimate fire behaviour
66 descriptors and develop empirical fire spread models (Forestry Canada Fire Danger Group 1992; Fernandes et al., 2009; Cruz
67 et al., 2015; Gollner et al., 2015), requiring significant time and resources. Neither laboratory-scale nor experimental fires
68 represent the spatial and temporal variability of environmental conditions under which uncontrolled wildfires most often burn
69 (e.g. Gollner et al., 2015).

70
71 Due to the unpredictability of their timing and location, conventional measurements on wildfires are difficult to perform and
72 lead to slow accumulation of data (Alexander & Cruz 2013b). Generally, they are of poor quality or incomplete (Duff et al.,



73 2013), although outstanding reconstruction examples exist (e.g. Wade & Ward 1973; Alexander & Lanoville 1987; Cheney
74 2010). Dedicated efforts do exist (Vaillant et al., 2014), but wildfire behaviour estimates often result from opportunistic
75 observations (e.g. Santoni et al., 2011) or post-fire interviews (e.g. Butler and Reynolds, 1997). Some authors have made
76 relevant efforts in compiling a large amount of direct field observations on wildfire behaviour (Alexander and Cruz, 2006;
77 Cheney et al., 2012), some combined with experimental fire data (Cruz and Alexander, 2013, 2019; Anderson et al., 2015;
78 Cruz et al., 2018, 2021, 2022; Khanmohammadi et al., 2022). An additional limitation lies on the fact that some of the existing
79 fire behaviour datasets are not freely available for the operational and research communities (Gollner et al., 2015).

80
81 Remote sensing technology, either through airborne or satellite platforms, can provide relevant data to document wildfires.
82 Manned or unmanned airborne visible and infrared (IR) images have been collected to document fire progression, and in some
83 cases to retrieve fire radiative power estimates (Schag et al., 2021; Storey et al., 2020, 2021; Coen & Riggan 2014; Sharples
84 et al., 2012). Satellite data provide easy-to-use, autonomous, synoptic observations of fire activity throughout the entire globe.
85 Recent advances in satellite technology have made available a panoply of imagery and products that range from moderate to
86 high spatial resolution, and from every 5 days to sub-daily frequency. Several authors have used satellite data to map daily fire
87 progression at country-level (Parks et al., 2014; Veraverbeke et al., 2014, Briones-Herrera et al., 2020; Sá et al., 2017) and at
88 the global scale (Artés et al., 2019; Oom et al., 2016). Some have estimated fire behaviour metrics, such as ROS (Humber et
89 al., 2022; Frantz et al., 2017; Andela et al., 2019). Recently, Chen et al., (2022) improved this line of research by using Visible
90 Infrared Imaging Radiometer Suite (VIIRS) data to automatically reconstruct sub-daily fire progression at a higher resolution.
91 Other authors exploited the capabilities of geostationary satellites to monitor wildfires and estimate fire behaviour descriptors
92 (Sifakis et al., 2011; Storey et al., 2021).

93
94 The different data sources used to characterise wildfire progression and behaviour have inherent limitations and potentialities.
95 Ground-collected data can be characterised by large uncertainties, particularly when taken by fire personnel whose focus is on
96 suppression and not on data collection (Alexander and Thomas, 2003). In addition, ground-collected data have poor synoptic
97 capability and provide a limited representation of fire behaviour variability. For example, distribution of ROS values for single
98 fire runs are seldom available (Cruz, 2010). Airborne data can provide wider coverage of the fire progression, however, have
99 limited temporal acquisition windows (e.g. USFS National Infrared Operations - NIROPS - provides data once per night) and
100 in some cases require manual digitization of fire perimeters (Veraverbeke et al., 2014).

101
102 The tradeoff between spatial and temporal resolution of satellite data, as well as the presence of clouds and thick smoke can
103 significantly limit their fire monitoring capability. In addition, the correct location of a wildfire cannot be determined inside a
104 burning pixel whose size varies with viewing geometry and sensor properties (Wolfe et al., 1998). Daily or sub-daily satellite-
105 derived fire progressions can also fail to reflect the influence of extreme conditions in fire behaviour due to the effect of
106 averaging over relatively long periods (Collins et al., 2009).



107
108 Considering that all data sources have limitations and provide information for very limited time frames, combining different
109 sources is key to capture the spread and behaviour variability of wildfires. The example provided in Figure 1 highlights the
110 potential of combining different data sources to overcome inherent acquisition gaps, particularly in the afternoon, when both
111 field and airborne data overcome the satellite gap, and during dawn, when ground-collected and satellite data complement each
112 other. Note that observation frequencies of ground and airborne data strongly depend on daily fire activity patterns.

113
114 **(Figure 1 near here)**

115
116 Systematic multi-source acquisition of wildfire data collection was recently done by Kilinc et al., (2012) and Storey et al.,
117 (2020, 2021) for Australia, by Crowley et al., (2019) for Canada (only satellite data) and by Fernandes et al., (2020) at the
118 global scale. The pursuit of this goal requires a monitoring framework and a concerted joint effort between research and
119 operational communities (Stocks et al., 2004; McCaw et al., 2012, Storey et al., 2020, 2021). Additional data on constantly
120 evolving wildfires, accompanied by robust replicable methods, is needed, namely in southern Europe where a substantial data
121 gap is manifest (Fernandes et al., 2018).

122
123 Here, we present the Portuguese Large Wildfire Spread Database (PT-FireSprd) that combines data from multiple sources,
124 using a “convergence of evidence” approach to characterise in detail the progression and behaviour of large wildfires in
125 Portugal. Fire behaviour is described in *sensu stricto*, thus analysis of its drivers and effects is beyond the scope of the current
126 work. The work results from a joint co-creation effort between researchers and fire personnel, integrating data collected from
127 airborne and ground operational resources.

128 **2 Data and Methods**

129 **2.1 Overview**

130 We first collected data for all the large wildfires (>100 ha) that occurred in mainland Portugal between 2015 and 2021. These
131 large wildfires were responsible for almost 1 million hectares burned during this period, of which half in the extreme fire
132 season of 2017. About 90% of the total burned area resulted from the 760 larger wildfires.

133 Multi-source input data (L0, section 2.2) were collected and only wildfires with good quality and representative data were
134 kept. Fire progressions were reconstructed from the input data and fire behaviour metrics were estimated. The PT-FireSprd
135 database was then organised in three levels:

- 136 • L1: Wildfire Progression (section 2.3), representing the spatial and temporal evolution of the wildfire spread (i.e.
137 where and when).



- 138 • L2: Wildfire behaviour (section 2.4), including quantitative behaviour descriptors of how a wildfire burned, such as
139 the rate-of-spread (ROS), fire growth rate (FGR), fire radiative energy (FRE), and FRE flux;
140 • L3: Simplified Wildfire behaviour (section 2.5), averaging fire behaviour over longer periods that represent relatively
141 homogenous fire runs.

142 The data from the different levels is composed by a large set of maps that can be useful for several applications and target
143 users. For example, L1 data can be used by fire analysts or researchers to evaluate suppression strategies and understand the
144 fire spread drivers or to evaluate burned area/fire perimeter mapping algorithms. L2 data is useful, for example, to calibrate
145 existing or build better fire spread models, while potential applications of L3 are improving fire danger rating, fire hazard
146 mapping and risk analysis. The overall flow of the data and methods is described in Figure 2.

147

148 **(Figure 2 near here)**

149

150 **2.2 Input Data (L0)**

151 To reconstruct the wildfire progressions, we used data acquired by satellites, from airborne sources and in the field by fire
152 personnel. Most of this data is currently integrated in a near-real time operational WEB-GIS fire monitoring platform (in
153 Portuguese “FEB Monitorização”, hereafter FEBMON) developed in 2018 by the Civil Protection Special Force (FEPC) and
154 the Portuguese National Authority for Emergency and Civil Protection (ANEPC). The data were complemented with official
155 fire data (e.g., ignition date and location) and information from external reports.

156 **2.2.1 Satellite data**

157 Satellite data was used to support the reconstruction of past wildfire spread. Currently, there are many sources of open-access
158 satellite data with capabilities to monitor wildfires over the entire globe. Their characteristics vary in resolution, ranging from
159 high (10-30 m) to low (4-5 km), and frequency of overpass, ranging from 5-15 days to every 15 minutes. To monitor wildfire
160 progression, satellites provide imagery and products that identify the location where a fire is actively burning at the time of
161 overpass (“thermal anomalies” or “active fire” products).

162

163 The Sentinel-2 Multispectral Instrument (MSI) and the Landsat 8\9 Operational Land Imager (OLI) provide images of the
164 Earth’s surface on average every 5 days when combined. Their spatial resolution ranges between 10 and 60 m depending on
165 the spectral band. PROBA-V has a lower number of spectral bands (4) when compared with other satellites used and provides
166 daily images at 300 m of spatial resolution and every 5 days with a 100 m spatial resolution. The VIIRS instrument aboard the
167 NPP and NOAA-20 satellites, collects data on average twice per day with a resolution varying between 375 m to 750 m,
168 depending on the spectral band. The Moderate-Resolution Imaging Spectroradiometer (MODIS) is an instrument onboard the
169 TERRA and AQUA satellites with spatial resolutions ranging from 250 m to 1000 m, depending on the spectral bands,



170 providing on average four daily revisits when combined. Sentinel-3 satellites have onboard the Sea and Land Surface
171 Temperature Radiometer (SLSTR) and the Ocean and Land Color Instrument (OLCI), with spatial resolutions ranging between
172 500 and 1000 m for the former, and 300 m for the latter. Data is acquired twice per day on average, but the OLCI does not
173 retrieve nighttime data.

174

175 We used L2 satellite imagery from the above-mentioned sensors to create false colour composites that could highlight burned
176 areas (low NIR, high SWIR reflectance), active flaming areas (high SWIR and/or TIR reflectance) and unburned vegetation
177 (high NIR reflectance). The bands used in the false colour composites depend on spectral characteristics of each sensor. Typical
178 false colour composites contain bands 12-8A-4 of Sentinel-2, bands 7-2-1 for MODIS and bands 1-2-4 for PROBA-V. Most
179 imagery was downloaded from Sentinel EO Browser (<https://apps.sentinel-hub.com/eo-browser/>), Worldview
180 (<https://worldview.earthdata.nasa.gov/>) and VITO-EODATA (<https://www.vito-eodata.be/PDF/>) which allow easy and fast
181 access to historical L2 data.

182

183 To complement the satellite imagery, we used the thermal anomaly products of VIIRS (VNP14IMGML-C1, Schroeder et al.,
184 2014, 2017) and MODIS (MCD14ML-C6, Giglio et al., 2003, 2016), with 375 m and 1 km resolution at nadir, respectively.
185 Data is available at fuoco.geog.umd.edu and FIRMS (<https://firms.modaps.eosdis.nasa.gov/>). These products allow estimating
186 the approximate location and timing of an active wildfire, and also provide an estimate of the fire radiative power (FRP), a
187 proxy of the radiant energy released per time unit and proxy for fuel consumption and fireline intensity. In addition, coarse
188 resolution data (~4 km) from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) sensor onboard the Meteosat
189 Second Generation (MSG) geostationary satellite, was used to characterise the temporal evolution of fire activity using FRP
190 estimates every 15' (Wooster et al., 2015). Data is available at <https://landsaf.ipma.pt/en/products/fire-products/frpgrid/>. The
191 FRP detections associated with each wildfire were identified using a spatial-temporal nearest distance algorithm. An empirical
192 threshold derived from the analysis of a selected number of wildfires was used to account for the satellite pixel geolocation
193 and temporal reporting uncertainties. For each wildfire, the Fire Radiative Energy (FRE), and associated uncertainties, were
194 estimated by integrating FRP detections over 30' periods and by assuming a constant rate of energy release (Eq. 1):

$$195 \quad FRE_i = 0.0009 \times (\sum_{k=1}^2 FRP_k), \quad (1)$$

196 where index i indicates the 30' bin, index k indicates the 15' FRP value in MW, and the 0.0009 factor converts the sum into
197 TJ.

198 2.2.2 Airborne data

199 Some aeroplanes and helicopters that operate during wildfires collect photos and videos. Data are collected during the initial
200 attack (i.e. up to 90 min after the alert) by the heli-brigades of the National Guard (GNR) using their mobile phones, and
201 occasionally, during extended attack. Aeroplanes, operated by FEPC\ANEPC since 2018, are equipped with a gimbal that



202 contains visible and thermal cameras, collecting photos and videos during extended attack. In addition, helicopters that
203 coordinate aerial suppression, also collect valuable information regarding fire progression. Both data sources collect data only
204 during daytime, with a very small number of exceptions, at relatively low altitudes.

205
206 These airborne data are systematically uploaded in real-time in FEBMON since 2018, providing high quality information
207 regarding the probable location of the fire start, active flaming zones, and specially wildfire progression. It is noteworthy to
208 mention that airborne footage is not synoptic, as different parts of the wildfire (e.g. left flank vs. right flank) are captured at
209 different moments. These, depending on the fire extent and operational priorities can be characterised by significant time lags.

210 **2.2.3 Ground data**

211 The FEBMON system is linked to user-friendly portable tools that allow collection of georeferenced ground data during
212 wildfires. These tools are typically installed in mobile phones and tablets and are used by fire personnel from several
213 organisations (e.g., fire fighters, forest service). Ground-collected data consists of three main types: i) photos and videos; ii)
214 points that identify active flaming combustion, inactive flaming or smouldering or locations requiring mop-up activities; iii)
215 polygons that delineate an area burned until the time of acquisition (i.e. fire progression).

216
217 Besides the data automatically linked to FEBMON, valuable ad-hoc information can be used to reconstruct wildfire spread,
218 such as additional photos and videos captured on the ground, and post-fire interviews. In sum, data collected by fire personnel
219 in the field provided valuable spatiotemporal information regarding wildfire spread, ignition and/or wildfire re-activation.

220 **2.2.4 Official fire data**

221 The Forest Service (in Portuguese, “Instituto da Conservação Natureza e das Florestas (ICNF)”) provides a fire database with
222 the final burned area perimeters for the entire country derived from a combination of field work and satellite data
223 (<https://geocatalogo.icnf.pt/>). We found some errors in the final perimeters that were corrected manually with Sentinel-2 or
224 Landsat 8/9 post-fire false colour composites (see section 2.2.1). In addition, for a very limited number of very large multi-
225 day wildfires, we used burned area perimeters provided by the Copernicus Emergency Management Service
226 (<https://emergency.copernicus.eu/mapping/>). The Forest Service also provides information regarding the wildfire start
227 location, mostly based on post-fire investigation done by GNR personnel (SGIF, <https://fogos.icnf.pt/sgif2010/>). Ignition data
228 have several known issues (Pereira et al., 2011) the most relevant of which, for the purposes of the present study, is the accuracy
229 of its exact location.

230
231 ANEPC manages the Operation Decision Support System (SADO) that includes information, such as i) date/hour of the
232 wildfire alert; ii) ignition location provided by first responders; and iii) a time log that seldom contains useful contextual
233 information on wildfire location at a given date/hour.



234 **2.2.4 Reports**

235 We also used ignition and fire progression data published in reports on the dynamics of the very large wildfires of June 2017,
236 including the Pedrogão Grande wildfire, and October 2017 (Guerreiro et al., 2017, 2018; Viegas et al., 2019). Regarding
237 Guerreiro et al. (2017, 2018), the primary data sources used to reconstruct the fire progression were satellite imagery, active
238 fire data and burned area perimeters provided by the Copernicus Emergency Management Service (see 2.2.1). Reports from
239 ANEPC and the Portuguese Institute for the Sea and the Atmosphere (IPMA, showing the fire plume evolution), GNR and the
240 Association for the Development and Industrial Aerodynamics (ADAI), were also used to identify fire arrival times and active
241 firelines. Additionally, other data sources allowed to reconstruct wildfire spread, such as: the official wildfire time log (see
242 2.2.4) , interviews (fire personnel involved in suppression, local residents), field work to identify the forward fire spread
243 direction based on scorched or charred foliage orientation, and other relevant data such as photos and videos. The fire spread
244 isochrones were determined through spatial interpolation methods (spline and inverse distance weighting), on high density
245 point clouds and experts' knowledge.

246

247 Viegas et al., (2019) reconstructed the extreme wildfires of October 2017 based on field work, interviews, photos/videos and
248 information contained in the official wildfire time log. Since the fire progression data were not provided by the authors, here
249 we used only very limited information regarding ignition location\time and general fire spread patterns, mostly to complement
250 data provided by Guerreiro et al., (2017, 2018).

251

252 We chose to include these fire progressions in our database, because they represent the most extreme wildfires that occurred
253 in mainland Portugal, under persistent cloud cover conditions that limited the acquisition of satellite data, and for that reason
254 they constitute relevant case studies, which otherwise would not be represented.

255 **2.3 Wildfire Progression (L1)**

256 Wildfire progression characterises the spatial and temporal evolution of the area burned in a specific fire event. It also contains
257 information regarding the ignition time and location, as well as, flaming zones that correspond to active areas during the
258 wildfire. These include spot fires and reactivation/rekindling areas. In Portugal, a rekindle is a reactivation of the wildfire after
259 its official conclusion and is considered a new incident. For simplicity, we will consider rekindles as reactivations throughout
260 the rest of the manuscript.

261

262 To robustly reconstruct wildfire progression, we combined the maximum available data from the different sources mentioned
263 above, with the aim of obtaining convergence of evidence. This allowed reducing the limitations and uncertainties of each
264 individual data source and building higher confidence in the derived wildfire progression.

265



266 Combining all the available data , we manually delimited the extent and time of the ignition, fire progression and active flaming
267 zones of each wildfire. The reconstruction was always made chronologically, i.e. starting from ignition and ending with the
268 progression prior to wildfire containment. Sentinel-2 and Landsat 8/9 pre-fire images were used to identify areas burned shortly
269 before the wildfire, and post-fire images were used to correct each progression polygon. As an example, Figure 3 shows how
270 different data sources were combined to derive the spread of the Castro Marim (2021) wildfire. All wildfire progression items
271 (L1) were defined as polygons, each with a set of different attributes (explained below).

272

273 **(Figure 3 near here)**

274

275 Ignition was defined as an area, instead of a point, to account for uncertainties in its location and to have a common data
276 typology for the entire database, in this case, vector polygons. We used mostly official ignition data and initial attack airborne
277 photos to define its location. This was complemented with expert knowledge and information from fire personnel to better
278 define ignition location. For a small set of wildfires (mostly nighttime ignitions), we also used satellite imagery and active-fire
279 data to identify the ignition area. All ignitions were compared with later fire spread patterns and with the final burned area to
280 reduce errors and guarantee consistency (e.g. ignition was contained in the final burned area). Regarding ignition time, the
281 official time of alert was compared with high frequency MSG-SEVIRI FRP detections, to confirm the alert time or, in a very
282 few cases, to anticipate if energy was released before the official ignition time. In addition, MSG-SEVIRI FRP were also useful
283 to identify (or confirm) the timing of reactivation. A clear example is shown in Figure 3, where the significant release of energy
284 around 11:30, combined with ground data, allowed identifying the location and time of the reactivation zone.

285

286 Active flaming zones were mostly derived from ground and/or high spatial resolution satellite imagery. Alternatively, they
287 were defined based on visual interpretation of multiple moderate resolution satellite imagery and often combined with active
288 fire data (mostly VIIRS due to its spatial resolution). Inconclusive visual interpretations were discarded, as well as active zones
289 that did not lead to any relevant subsequent fire spread. The ignition zone and all active flaming zones were always contained
290 within the subsequent fire spread polygon.

291

292 Wildfire progression was represented by a series of consecutive polygons delineating the temporal evolution of the area burned
293 by the wildfire. The number of polygons depended on fire size and data availability. The progression polygons were built using
294 as many data sources as possible, complementing each other in both space and time (see Figure 1). As an example: a common
295 feature found in the data was a pronounced fire spread during daytime, followed by very limited nighttime progression. In
296 these cases, first, the nighttime fire progression was delineated using active fire data (mostly VIIRS) and complemented with
297 ground data, when available. Second, satellite and/or airborne imagery acquired during the following morning were used to
298 perform any necessary adjustments in the nighttime spread polygon(s). Satellite-derived FRE estimates based on SEVIRI/MSG
299 were also used to identify if any substantial fire activity occurred between VIIRS/MODIS nighttime overpass and daytime



300 imagery (satellite and/or airborne). We assumed that fire activity decreased significantly when the wildfire released less than
301 0.5 TJ per 30' period, and anticipated the date/hour of the fire spread polygon accordingly. In smaller wildfires (<500 ha) this
302 threshold was set to 0.1 TJ. These thresholds were defined empirically (see Discussion section). The entire procedure reduced
303 the uncertainties associated with the delineation of the nighttime spread polygons. It should be noted that the fire behaviour
304 within the time span of each progression polygon was unknown and, therefore, was assumed to be free burning in a
305 homogeneous way (Storey et al., 2021). When data were insufficient to determine when a given area burned, the spread
306 polygon was flagged as “uncertain”.

307
308 Ignitions/active flaming zones were linked to the resultant spread polygon(s), by assigning a numeric label to a field called
309 “zp_link”, providing an explicit connection between both, and allowing to track the source of a given burned progression
310 polygon. When information was insufficient, for example, the start of the progression polygon was unknown, zp_link was
311 defined as “0”. After all ignition(s), fire progressions and active flaming zones were defined, each wildfire was divided into
312 burning periods. We assumed that each burning period contained relatively homogeneous fire runs that:

- 313
314 i) were ignited by the same set of ignitions or active flaming zones;
315 ii) did not exhibit large fire spread direction shifts (less than 45° of variation);
316 iii) were not impeded by barriers (e.g. previously burned area) and;
317 iv) did not exhibit significant changes in fire behaviour (e.g. large ROS variation).

318
319 Regarding the latter criterion, for example daytime and nighttime runs were usually separated in different burning periods even
320 if criteria (i)-(iii) were fulfilled. By definition, a new active flaming zone always marked the beginning of a new burning
321 period; however, not all burning periods started with an ignition or active flaming zone, since this depended on data availability.

322
323 When direct evidence of fire spotting was available (i.e. exact location/timing of the spot fire(s), typically from ground and/or
324 airborne data), if the fire front(s) rapidly (under 1 hour) coalesced with the original fire front, fire progression was merged
325 into a single polygon. In the remaining cases, typically associated with medium distance spotting and/or slow burning fire
326 fronts, the spotting location was defined as a new active flaming zone setting, defining a new burning period. When the exact
327 location/timing of the spot fire was not available, evidence of spotting consisted of observations of non-contiguous burned
328 areas that resulted from the same wildfire. These were typically separated by rivers, lakes and settlements. In these cases, due
329 to lack of data, the polygons separated from the major fire run were defined with zp_link=0 if the distance was larger than 200
330 m. No fire behaviour descriptors were calculated for these burned areas.

331
332 The definition of the burning period was always dependent on data availability and, in some cases, was subjective. For the
333 progressions derived using only satellite data, the length of the burning period was mostly determined by the timing of the



334 satellite overpass(es) and the FRE temporal evolution. For the progressions derived from more detailed data, the above-
335 mentioned criteria were easier to fulfil. In a few cases, uncertainties in fire progressions led to slightly overlapping periods.
336 An example is shown in the Results section and implications are addressed in the Discussion section.

337
338 After collecting input data for a large number of wildfires only those with at least one valid progression and a valid
339 ignition/active flaming zone were kept. We eliminated all suspicious cases where uncertainties were large, for example, due
340 to the presence of persistent smoke or clouds in the satellite images or absence of valid ground data. The L1 wildfire progression
341 database was defined by a set of polygons with attribute fields (details in section 3). The date/hour of each ignition(s), fire
342 spread and active flaming zones (if applicable) were approximated to the nearest 30' period.

343
344 Fire progression data from external reports were adapted to the rationale of the fire database described above. Findings from
345 different reports for the same wildfire were compared and satellite data was used to complement and improve the original fire
346 progressions.

347

348 **2.4 Wildfire behaviour (L2)**

349 The estimation of fire behaviour descriptors was supported by the use of spatial graphs. A graph is a mathematical structure
350 composed of nodes (N) and edges (E), which connect the nodes (Dale and Fortin, 2010). Based on the fire spread polygons
351 (L1) (Figure 4a), we built a spatial directed graph (or digraph) where each node refers to a spread polygon, and each edge
352 connects two spread polygons (i.e nodes), with a valid link (i.e. $z_p_link > 0$). These two nodes burned at different times, one
353 earlier (t_i) and the other later (t_j). The value of each edge was defined as the time elapsed between two nodes (Δt_{ij}) (Figure
354 4b). A node can have an inward edge (where fire is being transmitted from) and an outward edge (where fire is being
355 transmitted to).

356
357 First, the nodes were connected only if the associated fire progression polygons were contiguous, had the same z_p_link value
358 and burned at different timings. Second, only the edges corresponding to the shortest elapsed time between two nodes were
359 kept. The digraph allowed to formally structure the connections between fire spread polygons enabling the calculation of fire
360 behaviour descriptors.

361
362 To allow a better understanding of the methods used, a brief explanation based on the Ourique (2019) wildfire is provided. In
363 Figure 4, the number of the polygons on the left matches the number of nodes on the right. After its start (1), the wildfire
364 spread fast to the south and burned the area delimited by polygon 2 in about 120'. Fire behaviour changed after the head run,
365 and the left flank became the head and made a run to the southeast, burning the area represented by polygons 4, 5, 6 and 7, in
366 about 180'. This fire behaviour change observed at $t=120'$ determined the definition of two burning periods: one corresponding



367 to the initial head run, the other corresponding to head run from the left flank. The digraph was built with 7 nodes and 6 edges
368 with values ranging between 30' and 120'.

369

370 **(Figure 4 near here)**

371

372 Based on the fire progression (L1) and the corresponding di-graph, we calculated the following set of fire behaviour descriptors
373 (L2): forward ROS (m/h), spread direction ($^{\circ}$ from North), FGR (ha/h), and FRE (TJ). The polygons referring to areas burned
374 shortly before the fire analysed were removed from L2.

375

376 ROS was calculated for each node (N_j) with a valid inward edge (E_{ij}) connecting it to a prior node (N_i). By definition, the
377 forward ROS refers to the head of the fire and was calculated considering the longest distance line connecting two consecutive
378 fire progression polygons (i.e. nodes). representing the fastest spread (Storey et al., 2021). The ground distance (D_{ij}) between
379 each pair of polygons was calculated as follows:

380

- 381 • All ground distances between the polygon vertices of N_i and N_j were calculated, using the European Digital Elevation
382 Model (EU-DEM v1.1, <https://land.copernicus.eu/imagery-in-situ/eu-dem/eu-dem-v1.1>) resampled to 50 m spatial
383 resolution;
- 384 • For each vertex of the N_j polygon, only the shortest distance was kept and the corresponding pair of vertices, from
385 N_i and N_j , were stored;
- 386 • D_{ij} was defined as the maximum of all shortest distances between vertices.

387

388 The ROS was calculated by dividing the distance (D_{ij}) by the time elapsed between the pair of polygons (Δt_{ij}) and expressed
389 in m/h. We divided the ROS calculation in two distinct measures:

390

- 391 • Partial ROS (hereafter, ROSp) calculated between two consecutive polygons;
- 392 • Mean ROS (hereafter, ROSi), calculated between the ignition (or active flaming front) and a given spread polygon.

393

394 The spread direction was calculated using trigonometric rules considering the two above-mentioned vertices between two
395 polygons. The spread direction was calculated both for ROSp and ROSi, where the difference lies only on the origin polygon.
396 FGR was calculated dividing the burned area by each polygon/node (A_j) by the time elapsed between polygons (Δt_{ij}) and was
397 expressed in ha/h. An example of the calculation of these fire behaviour descriptors is shown in Figure 5.

398

399 **(Figure 5 near here)**



400

401 In addition to the standard fire behaviour descriptors, we also estimated the FRE for each progression polygon. This procedure
402 raised additional challenges. First, MSG-SEVIRI is affected by clouds and smoke, which can hinder the estimation of FRE for
403 some periods of the wildfires, or for their entire duration. Second, due to the coarse resolution of MSG-SEVIRI it was not
404 possible to calculate the FRE for each polygon directly. To circumvent this, FRE was calculated for each 30' bin from ignition
405 until the date/hour of the last wildfire spread polygon. In parallel, we estimated the area burned in each spread polygon every
406 30', using its start/end dates and assuming a constant FGR. Then, for each 30' bin, the total FRE was divided by weighting its
407 value by the proportion of area burned in each spread polygon. Finally, for each spread polygon the 30' FRE estimates were
408 summed only if they covered more than 70% of its duration (Δt_{ij}), to ensure that the total FRE was representative.

409

410 We also estimated the FRE flux rate ($\text{GJ ha}^{-1} \text{h}^{-1}$) for each spread polygon by dividing the estimated FRE by the corresponding
411 burned area extent and its duration (Δt_{ij}). As FRE is highly dependent on the extent burning at a given time window, the FRE
412 flux can provide estimates closer to “instantaneous” values required for other applications.

413 **2.5 Simplified Wildfire behaviour (L3)**

414 We calculated simplified metrics representing a mean fire behaviour across each burning period. This enables higher-level
415 analysis of the data, but at the cost of losing detail and making simplifications to the calculation of the fire behaviour metrics.

416

417 The simplified ROS corresponded to the ROSi estimated for the last spread polygon of a given burning period i.e. the average
418 ROS between the start and the end of each burning period. FGR was defined as the sum of the area burned in the period divided
419 by its duration. The total FRE was calculated considering all energy released by the polygons burning within the burning
420 period, if FRE estimates covered more than 70% of the area burned.

421 **2.6 Quality Control and Quality Assurance (QC/QA)**

422 All L1 to L2, and L2 to L3 processing was done using Matlab scripts complemented with quality controls checks to identify
423 errors in the original L1 data. These included simple checks to incorrect field names, incoherent data format (e.g., date/hour),
424 and consistency on the fire spread structure defined by the di-graphs, as for example: i) time elapsed between node was always
425 positive; and ii) every spread polygon with a positive z_{p_link} was always associated with a predecessor valid node (either of
426 “z” or “p” type), among others.

427

428 During the processing of L1 data to L2, we did frequent quality checks to identify potential errors, for example, null values of
429 ROS or FGR associated with valid fire spread polygons, fire progression polygons that did not have a known start/end date, or
430 did not have a known link to a preceding fire source (e.g., active flaming zone). In addition, we selected some wildfires and



431 made independent calculations of the ROS and FGR and compared them with the ones estimated using the developed Matlab
432 code. All these quality control steps assured that the data produced were reliable and of the best possible quality. The process
433 was iterative, requiring frequent corrections to the L1 data and the re-run of the quality check.

434
435 Finally, for each wildfire we defined a confidence flag that provides an overall information of how reliable the fire progression
436 data were. Although directly related to L1, ultimately it should also provide the user an estimate of the confidence associated
437 with L2 and L3. This was defined empirically based on the uncertainties that arose in the process of building the fire progression
438 polygons and was graded into a 5-level system where 1 refers to the lower quality and 5 to the highest quality (Table A1).

439 **3 Results**

440 **3.1 Overview of the PT-FireSprd database**

441 The PT-FireSprd database contains data for 80 large wildfires that occurred between 2015 and 2021. The individual wildfire
442 burned area extent ranges from 250 to 45,339 ha, with a mean and median area of 5,990 and 1,665 ha, respectively. The 80
443 wildfires were distributed throughout mainland Portugal, covering a wide range of environmental conditions (Figure 6). The
444 database spans a wide fire behaviour variability both between (e.g. Figure 6A,B,F) as well as within each wildfire (e.g. Figure
445 6C,E,D). The total burned area extent of the wildfires contained in the database was around 460,000 ha, which represents about
446 half of the area burned in the 2015-2021 period. On average, progression was reconstructed for 93% of the area burned by the
447 80 wildfires, leaving 7% deemed “uncertain”. Wildfire behaviour descriptors were estimated for 88% of the burned area extent
448 (ca. 400,000 ha). The time elapsed between two consecutive fire progression polygons ranged between 30’ and 14h30 with an
449 average value of 3h15. The mean duration of the burning periods was around 8h00, with a standard deviation of 4h50.

450
451 **(Figure 6 near here)**

452
453 A total of 1197 polygons with ROS and FGR estimates (L2) were derived from the progression data. We excluded very small
454 polygons (<25 ha) from further analysis, resulting in a dataset with 874 observations. Of the 1197 polygons, only 609 had FRE
455 estimates. Regarding L3 data, ROS and FGR were calculated for 241 burning periods (L3) and total FRE was only estimated
456 for 162 burning periods.

457
458 Overall, confidence in the database was lower for the earlier years (2015-2016) because input data was mostly from satellites.
459 In 2017, the quality increased due to the integration of i) ground data and ii) data from external reports that analysed the
460 extreme wildfires of June and October. From 2018 onwards, the integration of the monitoring aeroplanes, the creation of the
461 FEBMON system and the rapid availability of all the data that flows through it, significantly improved confidence of the
462 derived fire progressions.



463

464 The estimated forward ROS displayed a long-tail distribution (Figure 7, in log-scale) with a median value of 341 m/h and
465 average ROS of 746 m/h, representing large variability (std = 1071 m/h, cv = 143%). About 20% of the ROS values were
466 larger than 1000 m/h and about 9% were larger than 2000 m/h. The maximum observed ROS was 8956 m/h in the Lousã
467 wildfire of October 2017. The FGR distribution was highly skewed towards low values, with median and average values of 40
468 ha/h and 191 ha/h, respectively (sd = 438 ha/h, cv = 228%). About 10% of the observations had FGR larger than 500 ha/h and
469 only about 5% were larger than 1000 ha/h. The maximum observed FGR was 4436 ha/h in the Pedrogão Grande wildfire of
470 June 2017.

471

472 **(Figure 7 near here)**

473

474 The ROS distributions of the L2 and L3 datasets were similar. The largest differences were located in the lower and upper
475 tails, where the L3 ROS tends to be smoother due to the averaging procedure done over a longer time span. The FGR
476 distributions for L2 and L3 were also very similar, probably because all the polygon areas within a burning period are summed,
477 and the value does not result from an average. Differences were larger for more complex wildfires, for example with “finger
478 runs” (e.g. areas resulting from rapid propagation in a different direction than the dominant fire front).

479

480 We compared the histograms of L2 ROS and FGR for three aggregated confidence levels. The distribution of ROS estimates
481 for wildfires with lower confidence was slightly skewed towards lower values, when compared with higher confidence
482 estimates (Figure B1). The ROS distributions peak at 200 m/h, 500 m/h and 800 m/h for very low/low, moderate and high/very
483 high confidence, respectively, showing a clear relation between confidence and estimated ROS. Regarding FGR, very high
484 values above 500 ha/h were prevalent in wildfires with high and very high confidence progressions (Figure B2). Results are
485 similar if data from external reports for the extreme wildfires from June and October of 2017 are not included.

486

487 Estimated ROS and FGR were compared and percentiles 25, 50, 75, 90 and 97.5 were calculated for each variable
488 independently (Figure 8). The percentile values were simplified to enable a clear communication of results, especially between
489 researchers and fire personnel. The percentiles were translated into empirical classes, ranging from “very low” to “extreme”
490 fire behaviour. In general, as ROS increases so does the FGR. However, the relationship between ROS and FGR depends on
491 the morphology of the fire perimeter: elongated fast-spreading wildfires had relatively higher ROS and lower FGR (e.g. Figure
492 6B, C) and more complex burned area perimeters had relatively lower ROS and higher FGR (e.g. a flank run with an extensive
493 active fireline; see Figure 6A and the last polygons of Figures 6E and 6F). The dispersion tends to increase with higher
494 ROS/FGR values suggesting a progressively larger dependence on the burned area extent/perimeter. Identification of factors
495 determining such relationships is beyond the scope of this work. Nevertheless, wildfires with “Extreme” behaviour had both
496 very high values of ROS and FGR.



497

498 **(Figure 8 near here)**

499

500 Burned area extent is a relevant fire behaviour descriptor for researchers and fire management personnel. Analysis suggests
501 that the area burned by a wildfire is mostly determined by its FGR ($r=0.84$) rather than by the speed of the forward spread
502 ($r=0.62$; Figure 9a,b). The (cor)relations were lower using L2 data. As expected, FRE is highly correlated with burned area
503 extent ($r=0.85$, Figure 9c), and consequently of FGR. Correlation between ROS and average rate of energy release (TJ/h) is
504 lower ($r=0.30$, Figure 9d), however, there is a general direct relation between both descriptors.

505

506 **(Figure 9 near here)**

507

508 **2.2 Case study: The Castro Marim 2021 wildfire**

509 Here, we describe in detail the progression and behaviour of a specific wildfire to show how the PT-FireSprd database can be
510 used, for example, to analyse case studies, something often done by researchers and fire analysts.

511

512 The Castro Marim wildfire burned 5950 ha on the 16th and 17th of August of 2021. Figure 10 shows its reconstructed
513 progression (a) and associated ROS (b). Ignition occurred at nighttime (01:00) and a single run occurred towards SE until
514 approximately 08:30, defined as the first burning period. The mean ROS was 618 m/h, ranging between 321 and 957 m/h
515 (Figure 10c). The estimated FGR for the burning period was 43 ha/h, ranging between 33 and 77 ha/h, and the total FRE was
516 13 TJ (Figure 10d).

517

518 **(Figure 10 near here)**

519

520 Fire progression halted for about 3h until the wildfire reactivated around 11h30. It spread southwards until the head stopped
521 in an agricultural area around 19h30. In this second burning period, fire behaviour was significantly different from the first.
522 The mean ROS was ca. 1500 m/h, reaching a maximum value of 3720m/h between 16:30 and 17:30. On average, the fire grew
523 at a rate of 455 ha/h, however, significant variability was observed with values reaching 1236 ha/h coinciding with the ROS
524 peak. Framing the fire behaviour descriptors with the empirical classes represented in Figure 8, the behaviour in the second
525 burning period was often framed in the “Very High” class, i.e. between percentiles 90 and 97.5. As a consequence of the
526 behaviour exacerbation, the wildfire released around 38 TJ, with peaks of about 9 and 12 TJ observed during the afternoon.
527 The energy flux rate was highest between 16:00 and 16:30, coinciding with an abrupt increase in ROS (Figure 10d).

528



529 After the fire head stopped, a secondary head run stopped around 23:00 in a previously burned area (burning period 3). In the
530 follow-up, two left flank runs were observed, one until 02:30 and the other one, resulting from a reactivation, until 06:00, with
531 decreasing ROS, FGR and FRE. A secondary peak in the energy flux rate was estimated around 0:00, associated with an
532 increase in ROS and FGR.

533

534 Finally, in the Castro Marim wildfire burning periods 3 and 4 overlapped in time. A progression polygon in the rear/right flank
535 was delimited by fire personnel at 02:30, however the prior contiguous progression was identified at 16:30, suggesting a very
536 low burning flank, opposite to the fast burning part of the wildfire southwards. This overlap had no effect on the average ROS,
537 and only a very slight effect on the estimated FGR and FRE. However, users must be aware that burning periods seldom
538 overlap (~4% registered in the entire dataset), which may have implications in posterior analysis.

539 **4 Discussion**

540 **4.1 The PT-FireSprd database**

541 The PT-FireSprd is the first open access fire progression and behaviour database in the entire Mediterranean Europe. The
542 progression of 80 large wildfires that occurred in Portugal between 2015-2021 is reconstructed and fire behaviour descriptors
543 such as ROS, FGR and FRE are estimated, dramatically expanding the extant information (Palheiro et al., 2006; Rodriguez y
544 Silva & Molina-Martínez 2012; Fernandes et al., 2016). Wildfire progression was derived by converging evidence from
545 multiple data sources, which provides added credibility to the database. Wide variability in fire behaviour is covered, tackling
546 an important limitation pointed out by Cruz (2010). The approach presented will be used to update the database in the following
547 years for Portugal, and can be replicated in other countries, depending on data availability.

548

549 The large number of fire behaviour observations, both at the polygon level (L2) and at the burning period level (L3), provide
550 enough information for a wide variety of potential applications. For example, it can be used to: i) improve current knowledge
551 on the drivers affecting the behaviour of large wildfires; ii) calibrate existing or new models which ultimately should help to
552 better predict fire behaviour and support efficient fire management strategies (Alexander and Cruz, 2013a); iii) support the
553 construction of case studies by fire analysts and contribute to better training of fire personnel (Alexander and Thomas, 2003);
554 iv) contribute to improve operational fire suppression strategies; v) better understand how fire behaviour is linked to its effects
555 (Collins et al., 2009), and v) improve fire danger rating (Wotton, 2009). In addition, the fire behaviour classes described in
556 Figure 8 can assist fire suppression operations, including resources dispatching and decisions to fight or flee, or offensive vs
557 defensive strategies.

558

559 For several reasons, it is easier to collect information for larger wildfires than for smaller ones. The wide range in fire size
560 present in the PT-FireSprd database suggests that it is representative of wildfires burning under a broad range of conditions.



561 However, smaller wildfires (between 100 and 500 ha) are slightly under-represented in the database creating a potential bias.
562 This can be particularly relevant if one considers the proportion of smaller wildfires that occur every year. Thus, fire behaviour
563 descriptors may also be biased towards larger values which may have an implication, for example, on the fire behaviour classes
564 defined in Figure 8. Note that for typical fuel loads, say 15-20 t ha⁻¹ (Fernandes et al., 2016), the third class in Fig. 8 already
565 corresponds to fires very difficult to control directly (Hirsch and Martell 1996). Nevertheless, these classes should be
566 considered as a first exploratory approach with the aim of creating a simple and clear communication baseline between
567 researchers and fire personnel based on quantitative fire behaviour data. Ultimately, the database will allow framing the
568 behaviour of new wildfires according to historical patterns. Adding smaller wildfires to the PT-FireSprd database will certainly
569 help to better represent a wider range of fire behaviour.

570
571 Confidence in the wildfires of 2015-2016 was lower than for the most recent ones due to relevant advances in operational fire
572 monitoring resulting in better quality and higher quantity of fire data. Since 2018, the FEBMON system has improved and
573 grown, providing larger quantity and higher quality data, thus leading to more reliable and detailed fire progression
574 reconstructions. The distribution of the duration of the spread polygons between 2015 and 2021 (Figure B3) shows
575 heterogeneity of the database across time, but also the evolution introduced by the implementation of the FEBMON system.
576 Results suggest that estimates of ROS and FGR might be underpredicted in wildfires with lower confidence, most probably
577 due to the lack of data to thoroughly cover the afternoon, but especially the early night period (i.e. between VIIRS/MODIS
578 day and nighttime overpasses, Figure 1). This issue is further discussed in section 5.2. The user must take into account the
579 characteristics of the database and can choose to use the entire or part of the dataset based on the confidence flag or year of
580 the wildfire.

581
582 The PT-FireSprd database is flexible and open, allowing the users to subset the data based on their needs and requirements.
583 For example, users can decide to work with fire behaviour descriptors at the polygon level (L2) or at the burning period (L3),
584 or can create their own subset depending on their objectives. The dataset is heterogeneous which is reflected in two main
585 components: the duration of the spread polygons and the burning periods, and the confidence flag associated with each wildfire.

586
587 Regarding the duration, the average time elapsed between two progression polygons was 3h30 and 8h15 for the burning
588 periods. Durations were large in 2015 and 2016 (median values above 9h), decreased significantly in 2017 with the integration
589 of hourly isochrones from Guerreiro et al., (2017, 2018), and have had median durations below 2h since 2019 (Figure B3).
590 Gollner et al., (2015) argued that fire progression observations need to be made in real-time with a 10-metre spatial resolution
591 every 10' to meet the needs of fire behaviour forecasting. However, in operational context the current objective is to predict
592 fire behaviour time intervals larger or equal to 30' (Cruz and Alexander, 2013). Considering the average duration of the burning
593 periods, that represent a single fire run, the average time elapsed between progression observations represents a good
594 compromise and a clear advance in current data. Regardless, users can subset the database based on the duration of either the



595 progression polygons or the burning periods. L3 descriptors can be useful to provide more homogeneous and normalised fire
596 behaviour descriptors, dampening the effect of the large variability in L2 durations, allowing, for example, a better comparison
597 between wildfires.

598
599 Finally, preliminary results suggest that considering both ROS and FGR can improve understanding of wildfire dynamics. The
600 relation between both is related to perimeter morphology and extent, and future work is needed to better understand the
601 underlying factors. Most importantly, FGR was a better explanatory variable of burned area extent than ROS. The practical
602 consequence is that large burned areas can be generated by wildfires with a moderate forward ROS but with large FGR of the
603 entire perimeter, which in turn is highly influenced by spread duration and perimeter extent. This should have implications for
604 both the research and operational communities. FRE was estimated for a lower number of spread polygons and burning periods
605 when compared with ROS and FGR. This was most likely due to the impact of clouds and smoke on MSG detections and the
606 relatively conservative minimum number of observations threshold (75%). FRE and burned area extent were closely related,
607 however, relations between FRE and ROS were poor/moderate. One of the possible reasons may be related with the need to
608 consider the effect of the active perimeter extent when comparing both descriptors.

609 **4.2 Limitations and future improvements**

610 The generic limitations of the input data have been thoroughly described in Section 1. In particular for Portugal some
611 limitations of the data must be pointed out. Fire progression perimeters and fire points collected in the ground by fire personnel
612 have relevant spatio-temporal uncertainties. For example, there is often a lag between the date/hour a polygon is drawn in the
613 ground and the actual date/hour it burned completely. Another relevant issue is that of data acquisition / reporting errors done
614 by fire personnel, which may be reduced by improved training and experience. The number of users of the FEBMON system
615 has been growing in recent years and, with adequate training, it is expected that the quality and quantity of ground data will
616 increase in upcoming years. In fact, over 27,000 aerial and 2,500 ground photos were taken in the year 2022 which represents
617 a relevant increase compared to previous years.

618
619 Regarding airborne data, the discussion can be separated into two components. First, initial attack photos, which can be
620 extremely useful to draw initial fire progression and infer probable ignition areas, are not collected for every wildfire to which
621 a helicopter is dispatched, and sometimes are of poor quality. Additional training and increasing the awareness of fire personnel
622 for the relevance of the data they collect is necessary. Second, aeroplane data are acquired at relatively low altitude, precluding
623 a synoptic view of the wildfire. Time lags between data acquisition for different parts of the wildfire (e.g. left vs. right flanks)
624 may be large and introduce relevant spatio-temporal uncertainties in the delineation of the fire progression. In addition,
625 perimeters are drawn manually and depend on the training and experience of the fire expert. In upcoming years, the integration
626 of new airborne sensors, specially with multispectral capability, the ability to perform high-altitude scans and the use of
627 automatic perimeter delimitation procedures (e.g., Valero et al., 2018) should improve data quality and reduce the time lags of



628 airborne fire observations. With this new capacity, it will be possible to integrate deep learning processes in the data analysis,
629 increasing both the quantity and quality of the available fire data. This integration will also allow a well-organised structure in
630 data collection, management and analysis, improving decision-support systems. Finally, the use of UAVs during nighttime
631 (pioneered in 2022 in Portugal) will complement aeroplane/helicopter data during periods of low data availability.

632
633 Regarding official fire data, errors in the delineation of final burned area perimeters and in the ignition location, often located
634 outside of the fire perimeter, need to be corrected to increase the quality of the PT-FireSprd database. Regarding satellite data,
635 implementing (semi-) automatic algorithms to delimit fire perimeters (e.g., Chen et al., 2022) will increase the availability of
636 fire perimeters and reduce the uncertainties associated with manual perimeter delimitation. Improvements in the spatial
637 resolution geostationary satellites, such as the recently launched Meteosat Third Generation (MTG), will certainly improve
638 fire behaviour estimates, as already observed in HIMAWARI-8 and last generation GOES satellites.

639
640 Regarding methodological uncertainties, the major challenge was to assign the correct date/hour to a specific burned area. For
641 example, when raw data sources indicated that an area burned but active areas were absent or small, there were always
642 uncertainties as to when it actually burned completely, which could lead to a relevant ROS/growth rate underestimation. These
643 uncertainties were larger between dusk until VIIRS overpass(es) and between the later and dawn. One approach to reduce
644 these uncertainties was to use FRE data to monitor the daily cycle of fire activity and help to better define the start/end date of
645 a progression polygon. The method was empirical and future work is needed to better define the thresholds for setting the
646 ignition or reactivation times, as well as the end of a fire progression. Exploratory analysis done in a few wildfires of the PT-
647 FireSprd database suggest that FRE has a significant drop after the head of the fire stops, which may take several minutes/hours
648 until reaching the FRE thresholds used. This moment is commonly accompanied by a flank growth that burns slower and
649 releases lower amounts of energy. These fire dynamics probably explain why ROS was likely underestimated in low
650 confidence wildfires and why FGR was less affected by data confidence. Improvements can be achieved in the future, through
651 the use of more sophisticated methods (e.g. change point detection), more ground observations during the head to flank run
652 transition, and higher spatial resolution data from geostationary satellites. Part of these improvements can be used to partially
653 update the 2015-2021 wildfires of the PT-FireSprd database.

654
655 In terms of characterising uncertainties and its effects, future work should also adopt a metrological approach to propagate
656 uncertainties to the descriptors, providing useful information to users. By providing an uncertainty assessment, the PT-FireSprd
657 database would be on the pathway of Fiducial Reference Measurement (FRM) compliance.

658
659 The continuous update of the PT-FireSprd database will require a joint effort by researchers and fire personnel. The automation
660 of data collection procedures (discussed above), as well as dedicated training to fire personnel, are key factors to guarantee
661 both the quality as well as a sustainable update of the database. In the upcoming years, other fire behaviour descriptors could



662 be included such as type of spread (surface vs. crown fire), fireline intensity, flame size, spotting (including maximum distance)
663 and/or PyroCb occurrence. Finally, methods described in the current work can be, at least partially, applied to many other fire-
664 prone areas of the globe and contribute to the much-needed data on observed wildfire behaviour.

665 **5 Data Availability**

666 The dataset contains generic metadata file with relevant information for each wildfire (Table A2), such as the fire ID, official
667 incident ID (ANEPC, 13 digit number), fire name, municipality, civil parish, start date, duration (hours), extent (ha), among
668 others. The fire name was defined as Municipality_DDMMYYYY, where DD is day, MM month and YYYY the year. In
669 case more than one wildfire occurred in the same municipality on the same day, we added an additional string at the end of the
670 fire name (e.g. “_2”).

671
672 The dataset is then divided in 3 Levels, with three corresponding folders:

- 673 • Fire Spread (L1): Each year has a separate folder that contains one folder per wildfire labeled with the fire name. It
674 contains a polygon shapefile with the attributes listed in Table A3.
- 675 • Fire behaviour (L2): A single polygon shapefile that contains all wildfires and estimated fire behaviour metrics for
676 each individual fire spread polygon. The attributes are listed and explained in Table A4.
- 677 • Fire behaviour (L3): A single polygons shapefile that contains the simplified fire behaviour metrics calculated for
678 each burning period. The attributes are described in Table A5.

679
680 The generic metadata is connected to L1 data through the fire name field, and to L2 and L3 through the fire “ID” field.

681
682 The data are freely available at <https://doi.org/10.5281/zenodo.7495506> (last access: 30th December 2022; Benali et al. 2022).
683 We intend to update the database annually with wildfires from the current fire season and implement continuous improvements
684 to the procedure. Also, if additional information from past wildfires becomes available, we will update the database either by
685 changing existing fire spread polygons or by adding new wildfires. Updates for future years depend on the availability of input
686 data and associated funding.

687 **6 Conclusions**

688 The Portuguese Large Wildfire Spread Database (PT-FireSprd) is the first open access fire progression and behaviour database
689 available within Mediterranean Europe. It includes the reconstruction of the progression of 80 large wildfires that occurred in
690 Portugal between 2015 and 2021, that was derived by converging evidence from multiple data sources, which provides added
691 credibility to the database. PT-FireSprd contains a very large number of key fire behaviour observations, such as ROS, FGR



692 and FRE. Based on the statistical distribution of ROS and FGR, we defined 6 broad fire behaviour classes that can be easily
693 communicated to both research and management communities and support a wide number of applications, including better fire
694 management strategies. The PT-FireSprd has a large potential to contribute to the development of better fire behaviour
695 prediction tools, improve our current knowledge on wildfire dynamics, foster better operational training and contribute to
696 better decision-making. The approach will be used to continuously update the database in the following years for Portugal and
697 can be replicated in other countries/regions, depending on data availability. Improvements in data quality and the
698 implementation of automated methods are key factors for the regular update of the PT-FireSprd database in the future.

699 **Appendix A: Supporting material for the Methods**

700 (Table A1, Table A2, Table A3, Table A4 and Table A5 near here)

701 **Appendix B: Supporting material for the Results**

702 (Figure B1, Figure B2 and Figure B3 near here)

703 **Author Contribution**

704 AB and FS designed the study. AB, NG, HG, CM, JS carried out data processing and delimited fire progressions. BM carried
705 out FRE data processing. AB assembled the database, performed data analysis and wrote the first version of the manuscript.
706 All authors contributed to the interpretation of the results and writing of the manuscript.

707 **Competing interests**

708 The authors declare that they have no conflict of interest.

709 **Acknowledgements**

710 We thank Florian Briquemont for initial data processing and progression delimitation; FEPC personnel that provided relevant
711 fire to reconstruct some of the wildfires: Pedro Machado, Eduardo Marques, Marco Pires, Marco Lucas, Miguel Martins,
712 Daniel Santana and Vítor Caramelo. We would also like to thank other fire personnel that provided relevant fire data: João
713 Pedro Costa (AFOCELCA), José Silva (AFOCELCA), António Louro (CM-Mação), Sónia Oliveira (CM-Mação), Rui Lopes
714 (CBV Peso da Régua), Amélia Freitas (CM Caminha), Rui Pedro Fernandes (CBV Valença), Carlos Gomes (CBV Boticas),
715 António Ribeiro (ANEPC), Mário Silvestre (ANEPC), Emanuel Oliveira, and Elisio Pereira (CBV Porto de Mós). Finally, we
716 would like to thank ANEPC and FEPC for providing full access to their fire data enabling the development of the entire work.



717

718 This research was supported by the Forest Research Centre, a research unit funded by Fundação para a Ciência e a Tecnologia
719 I.P. (FCT), Portugal (UIDB/00239/2020), project foRester (PCIF/SSI/0102/2017) and FIRE-MODSAT II (PTDC/ASP-
720 SIL/28771/2017) also funded by FCT.

721

722 Akli Benali was funded by FCT through a CEEC contract (CEECIND/03799/2018/CP1563/CT0003). Nuno Guiomar was
723 funded by the European Union through the European Regional Development Fund in the framework of the Interreg V-A Spain-
724 Portugal program (POCTEP) under the CILIFO (Ref. 0753_CILIFO_5_E) and FIREPOCTEP (Ref.
725 0756_FIREPOCTEP_6_E) projects and by National Funds through FCT under the Project UIDB/05183/2020. Paulo
726 Fernandes contributed in the framework of the FCT-funded project UIDB/04033/2020. Ana Sá was supported under the
727 framework of the contract-program nr. 1382 (DL 57/2016/CP1382/CT0003).

728 **References**

729 Albini, F. A.: Wildland Fires: Predicting the behavior of wildland fires—among nature's most potent forces—can save lives,
730 money, and natural resources, *Am. Sci.*, 72(6), 590-597, 1984.

731 Alcasena, F., Ager, A., Le Page, Y., Bessa, P., Loureiro, C., and Oliveira, T.: Assessing wildfire exposure to communities and
732 protected areas in Portugal, *Fire*, 4(4), 82, doi:10.3390/fire404008, 2021.

733 Alexander, M., and Cruz, M. G.: Are the applications of wildland fire behaviour models getting ahead of their evaluation
734 again?, *Environ. Model. Softw.* (41), 65-71, doi:10.1016/j.envsoft.2012.11.001, 2013.

735 Alexander M. E., and Cruz M. G.: Evaluating a model for predicting active crown fire rate of spread using wildfire
736 observations, *Can. J. For. Res.*, 36, 3015–3028, doi:10.1139/x06-174, 2006.

737 Alexander, M. E., and Lanoville, R. A.: Wildfires as a source of fire behavior data: a case study from Northwest Territories,
738 Canada. 9th Conf. Fire and Forest Meteorology, April 21-24, San Diego, CA. American Meteorological Society, Boston, Mass,
739 86-93, 1987.

740 Alexander, M. E., and Thomas, D. A.: Wildland fire behavior case studies and analyses: Other examples, methods, reporting
741 standards, and some practical advice, *Fire Manag. Today*, 63(4), 4-12, 2003

742 Andela, N., Morton, D. C., Giglio, L., Paugam, R., Chen, Y., Hantson, S., van der Werf, G. R., and Randerson, J. T.: The
743 Global Fire Atlas of individual fire size, duration, speed and direction, *Earth Syst. Sci. Data*, 11(2), 529-552, doi:10.5194/essd-
744 11-529-2019, 2019.

745 Anderson, W. R., Cruz, M. G., Fernandes, P. M., McCaw, L., Vega, J. A., Bradstock, R. A., Fogarty, L. G., Gould, J. B.,
746 McCarthy, G. H., Marsden-Smedley, J. B., Matthews, S., Mattingley, G., Pearce, H. G., and van Wilgen, B. W.: A generic,
747 empirical-based model for predicting rate of fire spread in shrublands, *Int. J. Wildland Fire*, 24(4), 443-460,
748 doi:10.1071/WF14130, 2015.



- 749 Artés, T., Oom, D., De Rigo, D., Durrant, T. H., Maianti, P., Libertà, G., and San-Miguel-Ayanz, J.: A global wildfire dataset
750 for the analysis of fire regimes and fire behaviour, *Sci. Data*, 6(1), 1-11, doi:10.1038/s41597-019-0312-2, 2019.
- 751 Benali, A., Guiomar, N., Gonçalves, H., Mota, B., Silva, F., Fernandes, P.M., Mota, C., Penha, A., Santos, J., Pereira, J.M.C.,
752 and Sá, A.C.L.: The Portuguese Large Wildfire Spread Database (PT-FireSprd), <https://doi.org/10.5281/zenodo.7495506> ,
753 2022.
- 754 Briones-Herrera, C. I., Vega-Nieva, D. J., Monjarás-Vega, N. A., Briseño-Reyes, J., López-Serrano, P. M., Corral-Rivas, J. J.,
755 Alvarado-Celestino, E., Arellano-Pérez, S., Álvarez-González, J.G., Ruiz-González, A. D., Jolly, W. M., and Parks, S. A.:
756 Near real-time automated early mapping of the perimeter of large forest fires from the aggregation of VIIRS and MODIS
757 active fires in Mexico, *Remote Sens.*, 12(12), 2061, doi:10.3390/rs12122061, 2020.
- 758 Butler, B. W., and Reynolds, T. D.: Wildfire case study: Butte City, southeastern Utah, July 1, 1994, USDA For. Serv., Intermt.
759 Res. Stn., Ogden, UT. Gen. Tech. Rep. INT-GTR-351, doi:10.2737/INT-GTR-351, 1997.
- 760 Catchpole, W. R., Catchpole, E. A., Butler, B. W., Rothermel, R. C., Morris, G. A., and Latham, D. J.: Rate of spread of free-
761 burning fires in woody fuels in a wind tunnel, *Combust. Sci. Technol.*, 131(1-6), 1-37, doi:10.1080/00102209808935753,
762 1998.
- 763 Chen, Y., Hantson, S., Andela, N., Coffield, S. R., Graff, C. A., Morton, D. C., Ott, L.E., Foufoula-Georgiou, E., Smyth, P.,
764 Goulden, M. L., and Randerson, J. T.: California wildfire spread derived using VIIRS satellite observations and an object-
765 based tracking system, *Sci. Data*, 9(1), 1-15, doi:10.1038/s41597-022-01343-0, 2022.
- 766 Cheney, N. P., Gould, J. S., McCaw, W. L., and Anderson, W. R.: Predicting fire behaviour in dry eucalypt forest in southern
767 Australia, *For. Ecol. Manag.*, 280, 120-131, doi:10.1016/j.foreco.2012.06.012, 2012.
- 768 Cheney, N. P.: Fire behaviour during the Pickering Brook wildfire, January 2005 (Perth Hills Fires 71-80), *Conserv. Sci. West.*
769 *Aust.*, 7, 451–468, 2010.
- 770 Cheney, N., Gould, J., and Catchpole, W.: The Influence of Fuel, Weather and Fire Shape Variables on Fire-Spread in
771 Grasslands, *Int. J. Wildland Fire*, 3, 31, doi:10.1071/WF9930031, 1993.
- 772 Coen, J. L., and Riggan, P. J.: Simulation and thermal imaging of the 2006 Esperanza Wildfire in southern California:
773 application of a coupled weather–wildland fire model, *Int. J. Wildland Fire*, 23(6), 755-770, doi:10.1071/WF12194, 2014.
- 774 Collins, B. M., Miller, J. D., Thode, A. E., Kelly, M., van Wagtenonk, J. W., and Stephens, S. L.: Interactions among wildland
775 fires in a long-established Sierra Nevada natural fire area, *Ecosystems* 12, 114–128, doi:10.1007/s10021-008-9211-7, 2019.
- 776 Countryman, C. M.: The fire environment concept, USDA Forest Service, Pacific Southwest Range and Experiment Station,
777 Berkeley, California, USA, 1972.
- 778 Crowley, M. A., Cardille, J. A., White, J. C., and Wulder, M. A.: Generating intra-year metrics of wildfire progression using
779 multiple open-access satellite data streams, *Remote Sens. Environ.*, 232, 111295, doi:10.1016/j.rse.2019.111295, 2019.
- 780 Cruz, M. G., Alexander, M. E., and Kilinc, M.: Wildfire rates of spread in grasslands under critical burning conditions, *Fire*,
781 5(2), 55, doi:10.3390/fire5020055, 2022.



- 782 Cruz, M. G., Cheney, N. P., Gould, J. S., McCaw, W. L., Kilinc, M., and Sullivan, A. L.: An empirical-based model for
783 predicting the forward spread rate of wildfires in eucalypt forests, *Int. J. Wildland Fire*, 31(1), 81-95, doi:10.1071/WF21068,
784 2021.
- 785 Cruz, M. G., and Alexander, M. E.: The 10% wind speed rule of thumb for estimating a wildfire's forward rate of spread in
786 forests and shrublands, *Ann. For. Sci.*, 76(2), 1-11, doi:10.1007/s13595-019-0829-8, 2019.
- 787 Cruz, M. G., Alexander, M. E., Sullivan, A. L., Gould, J. S., and Kilinc, M.: Assessing improvements in models used to
788 operationally predict wildland fire rate of spread, *Environ. Model. Softw.*, 105, 54-63, doi:10.1016/j.envsoft.2018.03.027,
789 2018.
- 790 Cruz, M. G., Gould, J. S., Alexander, M. E., Sullivan, A. L., McCaw, W. L., and Matthews, S.: Empirical-based models for
791 predicting head-fire rate of spread in Australian fuel types, *Aust. For.*, 78(3), 118-158, doi:10.1080/00049158.2015.1055063,
792 2015.
- 793 Cruz, M. G., McCaw, W. L., Anderson, W. R., and Gould, J. S.: Fire behaviour modelling in semi-arid mallee-heath shrublands
794 of southern Australia, *Environ. Model. Softw.*, 40, 21-34, doi:10.1016/j.envsoft.2012.07.003, 2013.
- 795 Cruz, M. G., and Alexander, M. E.: Uncertainty associated with model predictions of surface and crown fire rates of spread,
796 *Environ. Model. Softw.*, 47, 16-28, doi:10.1016/j.envsoft.2013.04.004, 2013.
- 797 Cruz, M. G.: Monte Carlo-based ensemble method for prediction of grassland fire spread, *Int. J. Wildland Fire*, 19(4), 521-
798 530, doi:10.1071/WF08195, 2010.
- 799 Cruz, M. G., Alexander, M.E., and Wakimoto, R.H.: Development and testing of models for predicting crown fire rate of
800 spread in conifer forest stands, *Can. J. For. Res.*, 35(7), 1626-1639, doi:10.1139/x05-085, 2005.
- 801 Dale, M. R. T., and Fortin, M. J.: From graphs to spatial graphs, *Annu. Rev. Ecol. Evol. Syst.*, 21-38, doi:10.1146/annurev-
802 ecolsys-102209-144718, 2010.
- 803 Duff, T. J., Chong, D. M., and Tolhurst, K. G.: Quantifying spatio-temporal differences between fire shapes: Estimating fire
804 travel paths for the improvement of dynamic spread models, *Environ. Model. Softw.*, 46, 33-43,
805 doi:10.1016/j.envsoft.2013.02.005, 2013.
- 806 Fernandes, P. M., Sil, A., Ascoli, D., Cruz, M. G., Rossa, C. G., and Alexander, M. E.: Characterizing fire behavior across the
807 globe. In: Hood, S. M., Drury, S., Steelman, T., Steffens, R.[eds.]: Proceedings of the Fire Continuum-Preparing for the future
808 of wildland fire; 2018 May 21-24; Missoula, MT. Proceedings RMRS-P-78. Fort Collins, CO: US Department of Agriculture,
809 Forest Service, Rocky Mountain Research Station. p. 258-263., 78, 258-263, 2020.
- 810 Fernandes, P. M., Sil, A., Ascoli, D., Cruz, M. G., Alexander, M. E., Rossa, C. G., Baeza, J., Burrows, N., Davies, G. M.,
811 Fidelis, A., Gould, J. S., Govender, N., Kilinc, M., and McCaw, L.: Drivers of wildland fire behaviour variation across the
812 Earth. In: Viegas, D.X. (Ed.), *Advances in Forest Fire Research, Chapter 7 – Short contributions*, 1267-1270,
813 doi:10.14195/978-989-26-16-506_154, 2018.
- 814 Fernandes, P. M., Barros, A. M., Pinto, A., and Santos, J. A.: Characteristics and controls of extremely large wildfires in the
815 western Mediterranean Basin, *J. Geophys. Res. Biogeosci.*, 121(8), 2141-2157, doi:10.1002/2016JG003389, 2016.



- 816 Fernandes, P. M., Botelho, H.S., Rego, F.C., and Loureiro, C.: Empirical modelling of surface fire behaviour in maritime pine
817 stands, *Int. J. Wildland Fire*, 18(6), 698-710, doi:10.1071/WF08023, 2009.
- 818 Finney, M. A., McAllister, S. S., Forthofer, J. M., and Grumstrup, T. P.: *Wildland Fire Behaviour: Dynamics, Principles and*
819 *Processes*, CSIRO Pub., 2021.
- 820 Forestry Canada Fire Danger Group: Development and structure of the Canadian Forest Fire Behavior Prediction System. For.
821 Can., Ottawa, Ont. Inf. Rep. ST-X-3, 1992.
- 822 Frantz, D., Stellmes, M., Röder, A., and Hill, J.: Fire spread from MODIS burned area data: Obtaining fire dynamics
823 information for every single fire, *Int. J. Wildland Fire*, 25(12), 1228-1237, doi:10.1071/WF16003, 2016.
- 824 Giglio, L., Descloitres, J., Justice, C. O., and Kaufman, Y. J.: An enhanced contextual fire de- tecton algorithm for MODIS,
825 *Remote Sens. Environ.* 87, 273–282, doi:10.1016/S0034-4257(03)00184-6, 2003
- 826 Giglio, L., Schroeder, W. and Justice, C. O.: The collection 6 MODIS active fire detection algorithm and fire products. *Remote*
827 *Sens. Environ.*, 178, pp.31-41, doi:10.1016/j.rse.2016.02.054, 2016.
- 828 Gollner, M., Trouve, A., Altintas, I., Block, J., de Callafon, R., Clements, C., Cortes, A., Ellicott, E., Filippi, J. B., Finney, M.,
829 Ide, K., Jenkins, M.A., Jimenez, D., Lautenberger, C., Mandel, J., Rochoux, M., and Simeoni, A.: Towards data-driven
830 operational wildfire spread modeling: A report of the NSF-funded WIFIRE workshop, 2015.
- 831 Guerreiro, J., Fonseca, C., Salgueiro, A., Fernandes, P., Iglésias, E. L., Neufville, R., Mateus, P., Castellnou, M., Silva, J. S.,
832 Moura, J. M., Rego, F. C., and Caldeira, D.: Análise e apuramento dos factos relativos aos incêndios que ocorreram em
833 Pedrogão Grande, Castanheira de Pêra, Ansião, Alvaiázere, Figueiró dos Vinhos, Arganil, Góis, Penela, Pampilhosa da Serra,
834 Oleiros e Sertã, entre 17 e 24 de junho de 2017. Comissão Técnica Independente, Assembleia da República, Lisboa.
835 https://www.parlamento.pt/Documents/2017/Outubro/RelatórioCTI_VF.pdf, 2017.
- 836 Guerreiro, J., Fonseca, C., Salgueiro, A., Fernandes, P., Iglésias, E. L., Neufville, R., Mateus, P., Castellnou, M., Silva, J. S.,
837 Moura, J. M., Rego, F. C., and Caldeira, D.: Avaliação dos Incêndios ocorridos entre 14 e 16 de outubro de 2017 em Portugal
838 Continental. Comissão Técnica Independente, Assembleia da República, Lisboa.
839 <https://www.parlamento.pt/Documents/2018/Marco/RelatorioCTI190318N.pdf>, 2018.
- 840 Humber, M., Zubkova, M., and Giglio, L.: A remote sensing-based approach to estimating the fire spread rate parameter for
841 individual burn patch extraction, *Int. J. Remote Sens.*, 43(2), 649-673, doi:10.1080/01431161.2022.2027544, 2022.
- 842 Hirsch, K. G., and Martell, D. L. A review of initial attack fire crew productivity and effectiveness, *Int. J. Wildland Fire*, 6(4),
843 199-215, doi:10.1071/WF9960199, 1996.
- 844 Khanmohammadi, S., Arashpour, M., Golafshani, E. M., Cruz, M. G., Rajabifard, A., and Bai, Y.: Prediction of wildfire rate
845 of spread in grasslands using machine learning methods, *Environ. Model. Softw.*, 156, 105507,
846 doi:10.1016/j.envsoft.2022.105507, 2022.
- 847 Kilinc, M., Anderson, W., and Price, B.: *The Applicability of Bushfire Behaviour Models in Australia*. Victorian Government,
848 Department of Sustainability and Environment. DSE Schedule 5: Fire Severity Rating Project, Melbourne, VIC. Technical
849 Report 1, 2012.



- 850 McCaw, W. L., Gould, J. S., Cheney, N. P., Ellis, P. F. M., and Anderson, W. R.: Changes in behaviour of fire in dry eucalypt
851 forest as fuel increases with age, *For. Ecol. Manag.* 271, 170-181, doi:10.1016/j.foreco.2012.02.003, 2012
- 852 Oom, D., Silva, P. C., Bistinas, I., and Pereira, J. M. C.: Highlighting biome-specific sensitivity of fire size distributions to
853 time-gap parameter using a new algorithm for fire event individuation, *Remote Sens.*, 8(8), 663, doi:10.3390/rs8080663, 2016.
- 854 Palaiologou, P., Kalabokidis, K., Ager, A. A., and Day, M. A.: Development of Comprehensive Fuel Management Strategies
855 for Reducing Wildfire Risk in Greece, *Forests*, 11(8), 789, doi:10.3390/f11080789, 2020.
- 856 Palheiro, P. M., Fernandes, P. M., Cruz, M. G.: A fire behaviour-based fire danger classification for maritime pine stands:
857 comparison of two approaches, *For. Ecol. Manag.*, (234), S54, doi:10.1016/j.foreco.2006.08.075, 2006.
- 858 Parisien M. A., Parks S. A., Miller C., Krawchuk M. A., Heathcott M., and Moritz M. A.: Contributions of ignitions, fuels,
859 and weather to the burn probability of a boreal landscape, *Ecosystems* 14, 1141–1155, doi:10.1007/s10021-011-9474-2, 2011.
- 860 Parks, S. A.: Mapping day-of-burning with coarse-resolution satellite fire-detection data, *Int. J. Wildland Fire*, 23(2), 215-223,
861 doi:10.1071/WF13138, 2014.
- 862 Pereira, M. G., Malamud, B. D., Trigo, R. M., and Alves, P.I.: The history and characteristics of the 1980–2005 Portuguese
863 rural fire database, *Nat. Hazards Earth Syst. Sci.*, 11(12), 3343-3358, doi:10.5194/nhess-11-3343-2011, 2011.
- 864 Rodríguez y Silva, F., and Molina-Martínez, J.R.: Modeling Mediterranean forest fuels by integrating field data and mapping
865 tools, *Eur. J. For. Res.*, 131, 571–582, doi:10.1007/s10342-011-0532-2, 2012.
- 866 Rothermel, R.C.: A mathematical model for predicting fire spread in wildland fuels, Res. Pap. INT-115. Ogden, UT: U.S.
867 Department of Agriculture, Intermountain Forest and Range Experiment Station, 1972.
- 868 Sá, A. C., Benali, A., Fernandes, P. M., Pinto, R. M., Trigo, R. M., Salis, M., Russo, A., Jerez, S., Soares, P. M. M., Schroeder,
869 W., and Pereira, J. M. C.: Evaluating fire growth simulations using satellite active fire data, *Remote Sens. Environ.*, 190, 302-
870 317, doi:10.1016/j.rse.2016.12.023, 2017.
- 871 Salis, M., Del Giudice, L., Arca, B., Ager, A. A., Alcasena-Urdiroz, F., Lozano, O., Bacciu, V., Spano, D., and Duce, P.:
872 Modeling the effects of different fuel treatment mosaics on wildfire spread and behavior in a Mediterranean agro-pastoral area.
873 *J. Environ. Manage.*, 212, 490-505, doi:10.1016/j.jenvman.2018.02.020, 2018.
- 874 Santoni, P.-A., J.-B. Filippi, J.-H. Balbi, and Bosseur, F.: Wildland fire behaviour case studies and fuel models for landscape-
875 scale fire modeling, *J. Combust.*, 613424, doi:10.1155/2011/613424, 2011.
- 876 Schag, G. M., Stow, D. A., Riggan, P. J., Tissell, R. G., and Coen, J. L.: Examining landscape-scale fuel and terrain controls
877 of wildfire spread rates using repetitive airborne thermal infrared (ATIR) imagery, *Fire*, 4(1), 6, doi:10.3390/fire4010006,
878 2021.
- 879 Schroeder, W., Oliva, P., Giglio, L., and Csiszar, I. A.: The New VIIRS 375 m active fire detection data product: Algorithm
880 description and initial assessment, *Remote Sens. Environ.*, 143, 85-96, doi:10.1016/j.rse.2013.12.008, 2014.
- 881 Schroeder, W.: Visible Infrared Imaging Radiometer Suite (VIIRS) 375 m & 750 m Active Fire Detection Data Sets Based on
882 NASA VIIRS Land Science Investigator Processing System (SIPS) Reprocessed Data—Version 1. 2017.

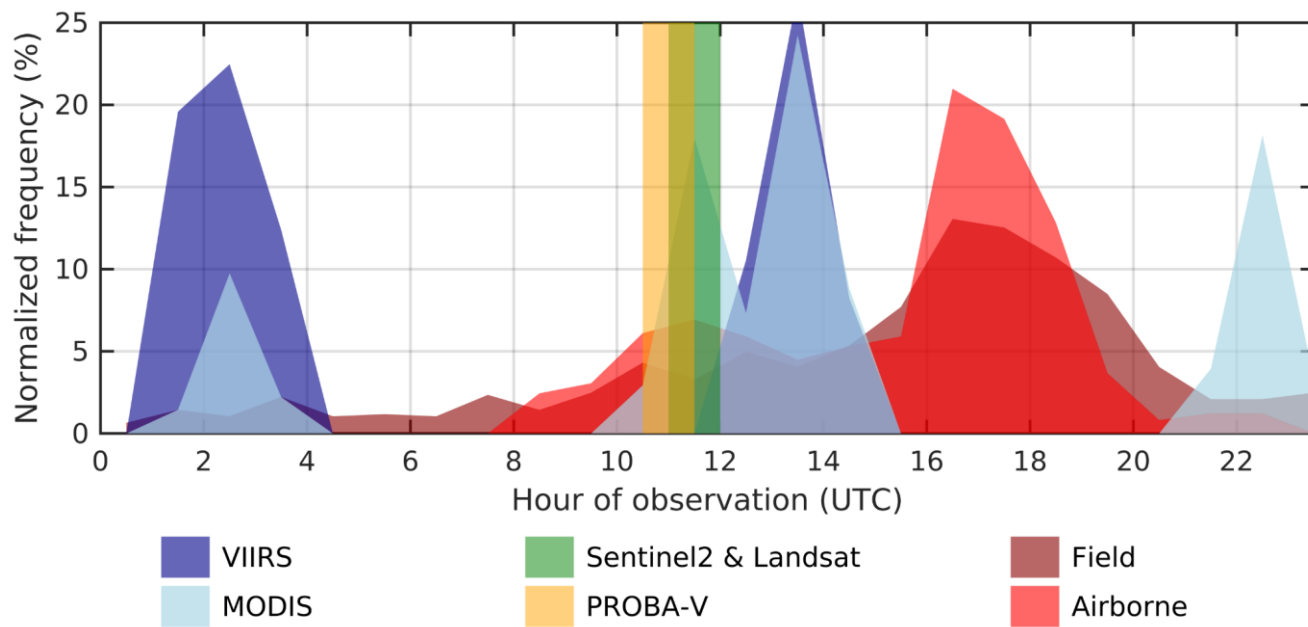


- 883 Scott, J. H., and Reinhardt, E. D.: Assessing Crown Fire Potential by Linking Models of Surface and Crown Fire Behavior,
884 US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO. Research Paper RMRS-
885 RP-29, doi:10.2737/RMRS-RP-29, 2001.
- 886 Sharples, J. J., McRae, R. H., and Wilkes, S. R.: Wind–terrain effects on the propagation of wildfires in rugged terrain: fire
887 channelling, *Int. J. Wildland Fire*, 21(3), 282–296, doi:10.1071/WF10055, 2012.
- 888 Sifakis, N. I., Iossifidis, C., Kontoes, C., and Keramitsoglou, I.: Wildfire detection and tracking over Greece using
889 MSG-SEVIRI satellite data, *Remote Sens.*, 3(3), 524–538, doi:10.3390/rs3030524, 2011.
- 890 Stocks, B. J., Alexander, M. E., Wotton, B. M., Steffner, C. N., Flannigan, M. D., Taylor, S. W., Lavoie, N., Mason, J. A.,
891 Hartley, G. R., Maffey, M. E., Dalrymple, G. N., Blake, T. W., and Cruz, M. G., Lanoville, R. A.: Crown fire behaviour in a
892 northern jack pine black spruce forest, *Can. J. For. Res.* 34, 1548–1560, doi:10.1139/x04-054, 2004.
- 893 Storey, M. A., Price, O. F., Sharples, J. J., and Bradstock, R. A.: Drivers of long-distance spotting during wildfires in south-
894 eastern Australia, *Int. J. Wildland Fire*, 29(6), 459–472, doi:10.1071/WF19124, 2020.
- 895 Storey, M. A., Bedward, M., Price, O. F., Bradstock, R. A., and Sharples, J. J.: Derivation of a Bayesian fire spread model
896 using large-scale wildfire observations, *Environ. Model. Softw.*, 144, 105127, doi:10.1016/j.envsoft.2021.105127, 2021.
- 897 Vaillant, N. M., Ewell, C.M., and Fites-Kaufman, J. A.: Capturing crown fire behavior on wildland fires - the Fire Behavior
898 Assessment Team in action, *Fire Manag. Today* 73(4):41–45, 2014.
- 899 Valero, M. M., Rios, O., Pastor, E., and Planas, E.: Automated location of active fire perimeters in aerial infrared imaging
900 using unsupervised edge detectors, *Int. J. Wildland Fire*, 27(4), 241–256, doi:10.1071/WF17093, 2018.
- 901 Veraverbeke, S., Sedano, F., Hook, S. J., Randerson, J. T., Jin, Y., and Rogers, B. M.: Mapping the daily progression of large
902 wildland fires using MODIS active fire data, *Int. J. Wildland Fire*, 23(5), 655–667, doi:10.1071/WF13015, 2014.
- 903 Viegas, D. X., Almeida, M. F., Ribeiro, L. M., Raposo, J., Viegas, M. T., Oliveira, R., Alves, D., Pinto, C., Rodrigues, A.,
904 Ribeiro, C., Lopes, S., Jorge, H., and Viegas, C. X.: Análise dos Incêndios Florestais Ocorridos a 15 de outubro de 2017,
905 Centro de Estudos sobre Incêndios Florestais (CEIF/ADAI/LAETA), 2019.
- 906 Wade, D. D., and Ward, D. E.: An analysis of the Air Force Bomb Range Fire. Res. Pap. SE–105. Asheville, NC, USDA
907 Forest Service, Southeastern Forest Experiment Station, 1973.
- 908 Wolfe, R. E., Roy, D. P., and Vermote, E.: MODIS land data storage, gridding, and compositing methodology: level 2 grid,
909 *IEEE Trans. Geosci. Remote Sens.* 36, doi:10.1109/36.701082, 1998.
- 910 Wooster, M. J., Roberts, G., Freeborn, P. H., Xu, W., Govaerts, Y., Beeby, R., He, J., Lattanzio, A., Fisher, D., and Mullen,
911 R.: LSA SAF Meteosat FRP products – Part 1: Algorithms, product contents, and analysis, *Atmos. Chem. Phys.*, 15, 13217–
912 13239, doi:10.5194/acp-15-13217-2015, 2015.
- 913 Wotton, B. M.: Interpreting and using outputs from the Canadian Forest Fire Danger Rating System in research applications,
914 *Environ. Ecol. Stat.*, 16(2), 107–131, doi:10.1007/s10651-007-0084-2, 2009.

915
916



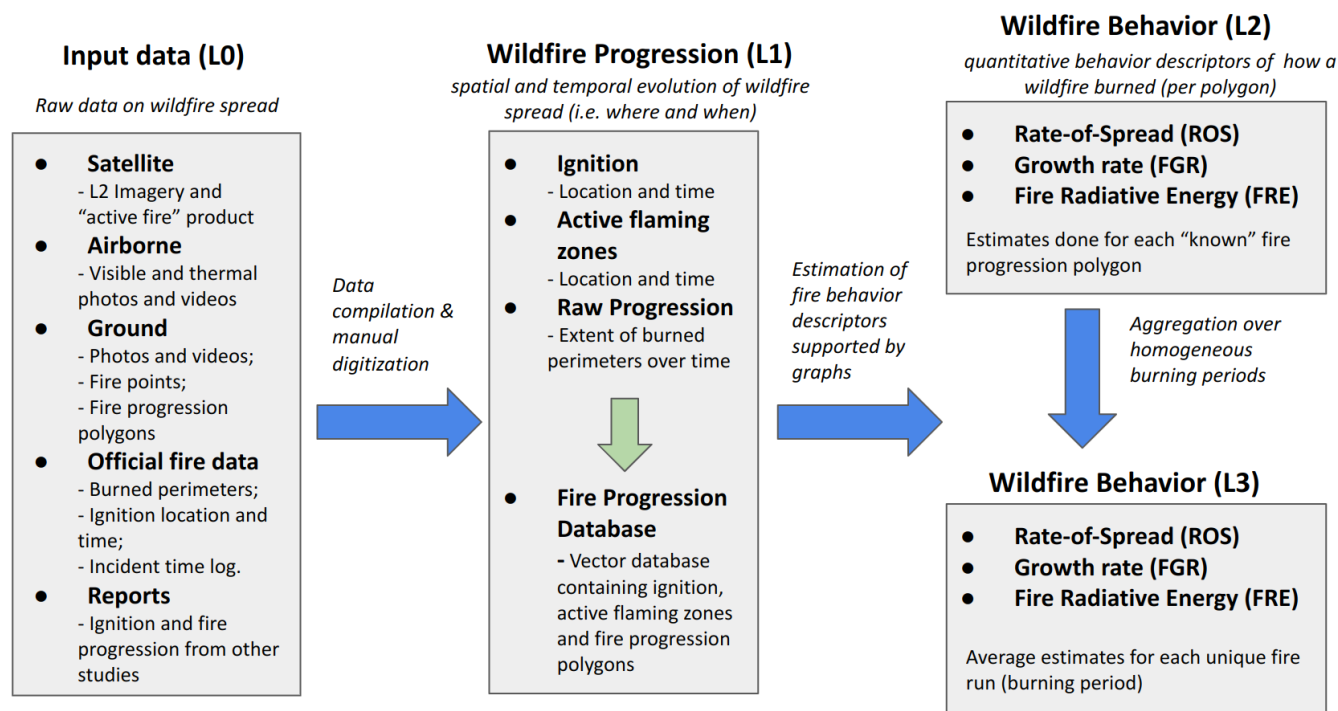
917 **Figures**



918
919

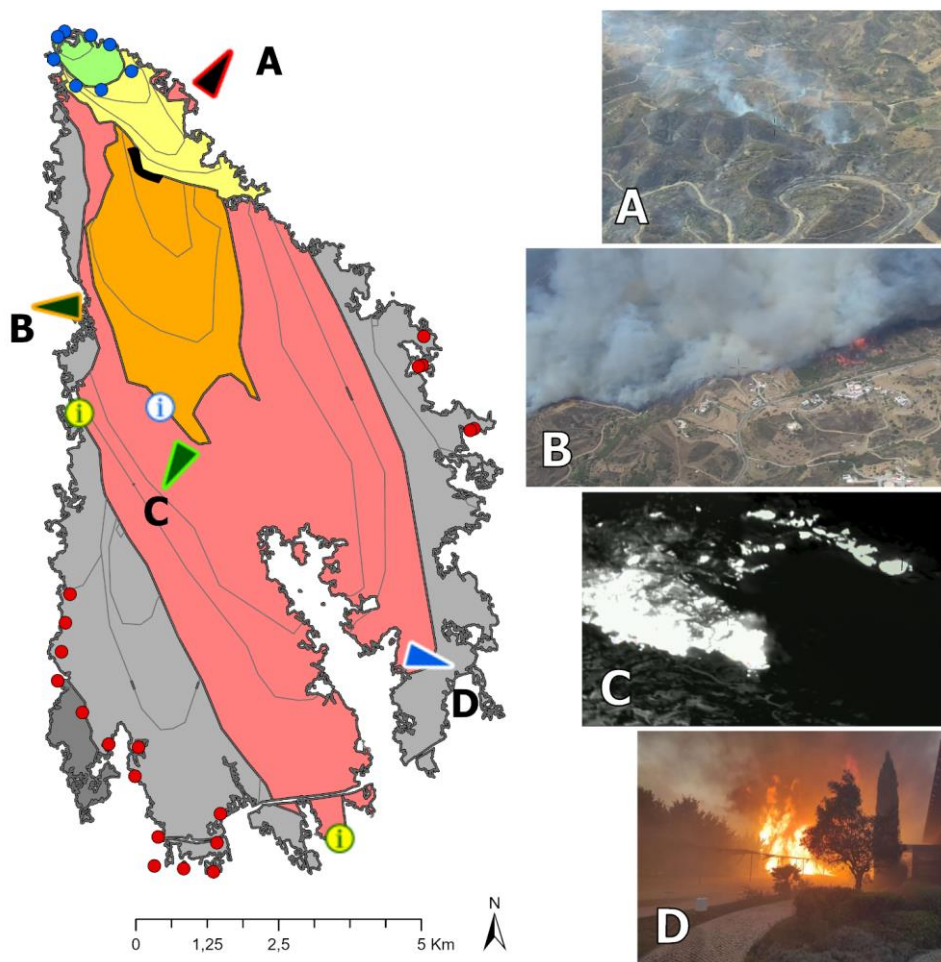
920 **Figure 1: Hourly frequency of observations in active wildfires acquisitions for satellite, field and airborne data. The data used refers to the year 2019 as an example. The frequency is normalised by dividing the number of observations by the total of each data source. Sentinel-2, Landsat and PROBA-V refer to the temporal windows and not the frequency, since all of the data are acquired in a very short window. The time windows of Sentinel-3 are similar to those of MODIS. MSG-SEVIRI data are not represented since it has a 15' frequency. Acronyms are described in the Data and Methods section.**

925



926
 927
 928
 929
 930

Figure 2: Flowchart that represents an overview of the data and methods used in the development of the PT-FireSprd database.



Input data

Satellite data (VIIRS thermal anomalies)

- 2021-08-16 03:09
- 2021-08-17 02:47

Airborne and fire operatives data

- ▼ A – Airplane, reactivation (2021-08-16 11:32)
- ▼ B – Airplane, right flank (2021-08-16 16:26)
- ▼ C – Airplane (thermal), fire front (2021-08-16 16:38)
- ▼ D – Operatives, fire front (2021-08-16 19:30)

Reports

- ⓘ Location reported in timeline (2021-08-16 16:18)
- ⓘ Locations reported in timeline (2021-08-16 19:30)

Estimated Fire Progression

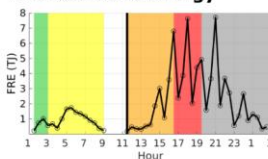
Ignition/active flaming zones

- Reactivation Zone (2021-08-16 11:30)

Fire perimeters

- 2021-08-16 03:00
- 2021-08-16 09:00
- 2021-08-16 16:30
- 2021-08-16 19:30
- 2021-08-17 03:00
- 2021-08-17 12:00
- Intermediate perimeters

Fire Radiative Energy



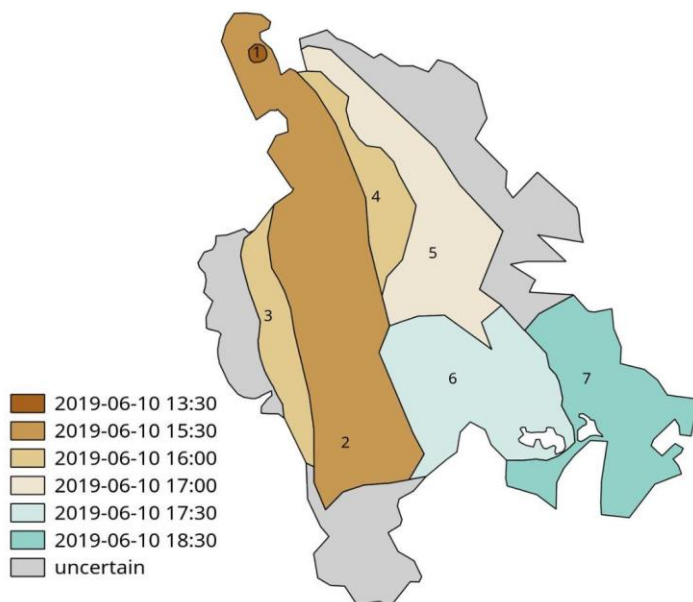
931
 932

933
 934

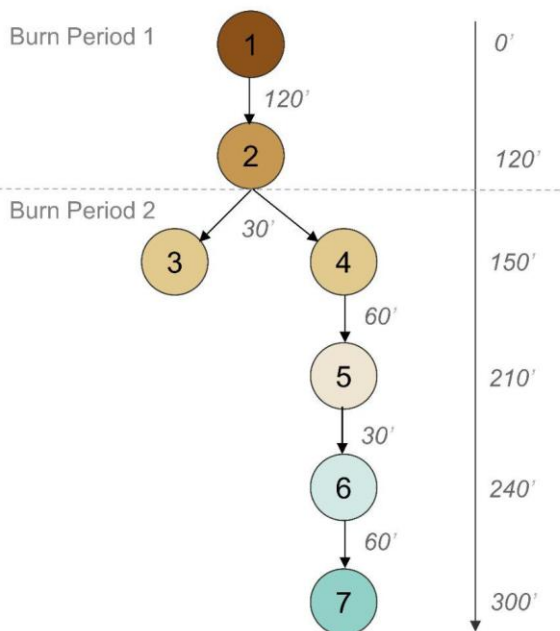
Figure 3: Example of multi-source data integration to derive fire perimeters and reconstruct the progression of the Castro Marim (2021) wildfire. The lines represent different progression polygons. Photos A, B, C, D were kindly provided by ANEPC\FEPC



a) Fire progression

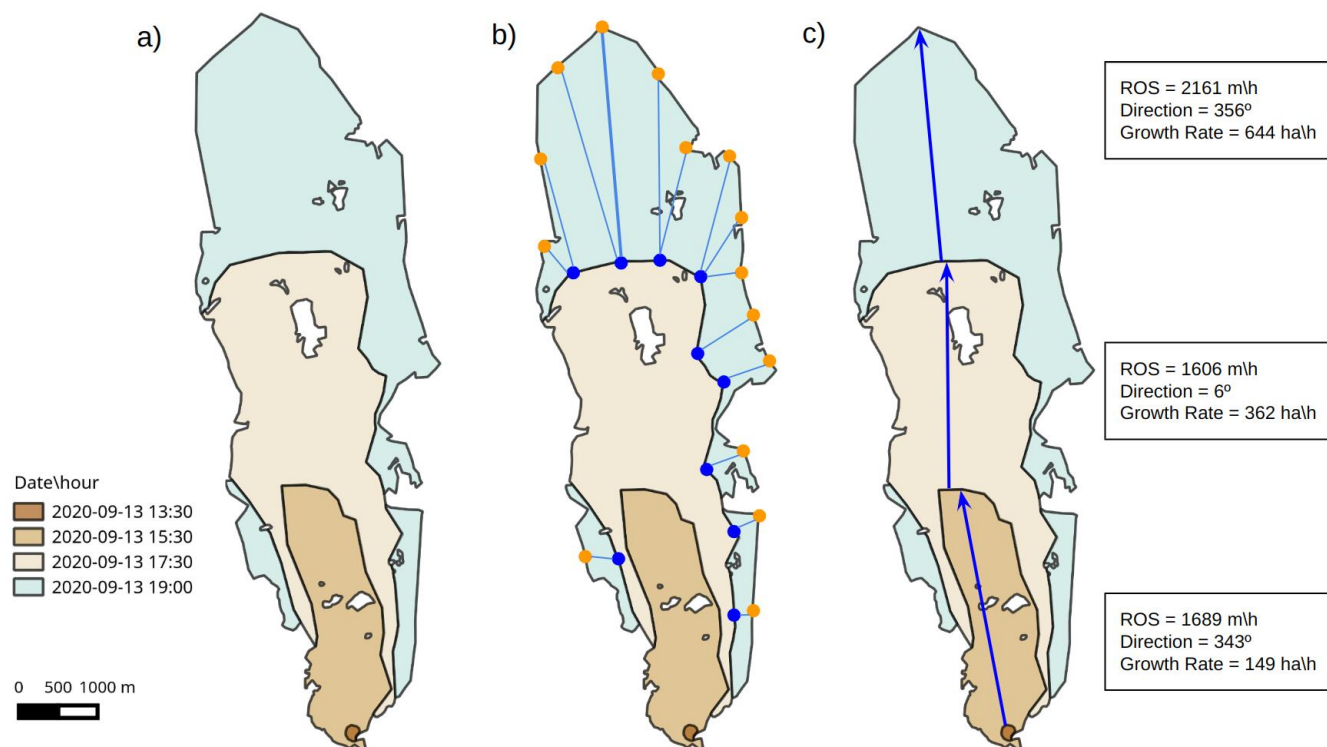


b) Di-graph structure



935
936
937
938

Figure 4: Example of how the estimated fire progression (a) of the Ourique 2019 wildfire was used to build the digraph (b).



939

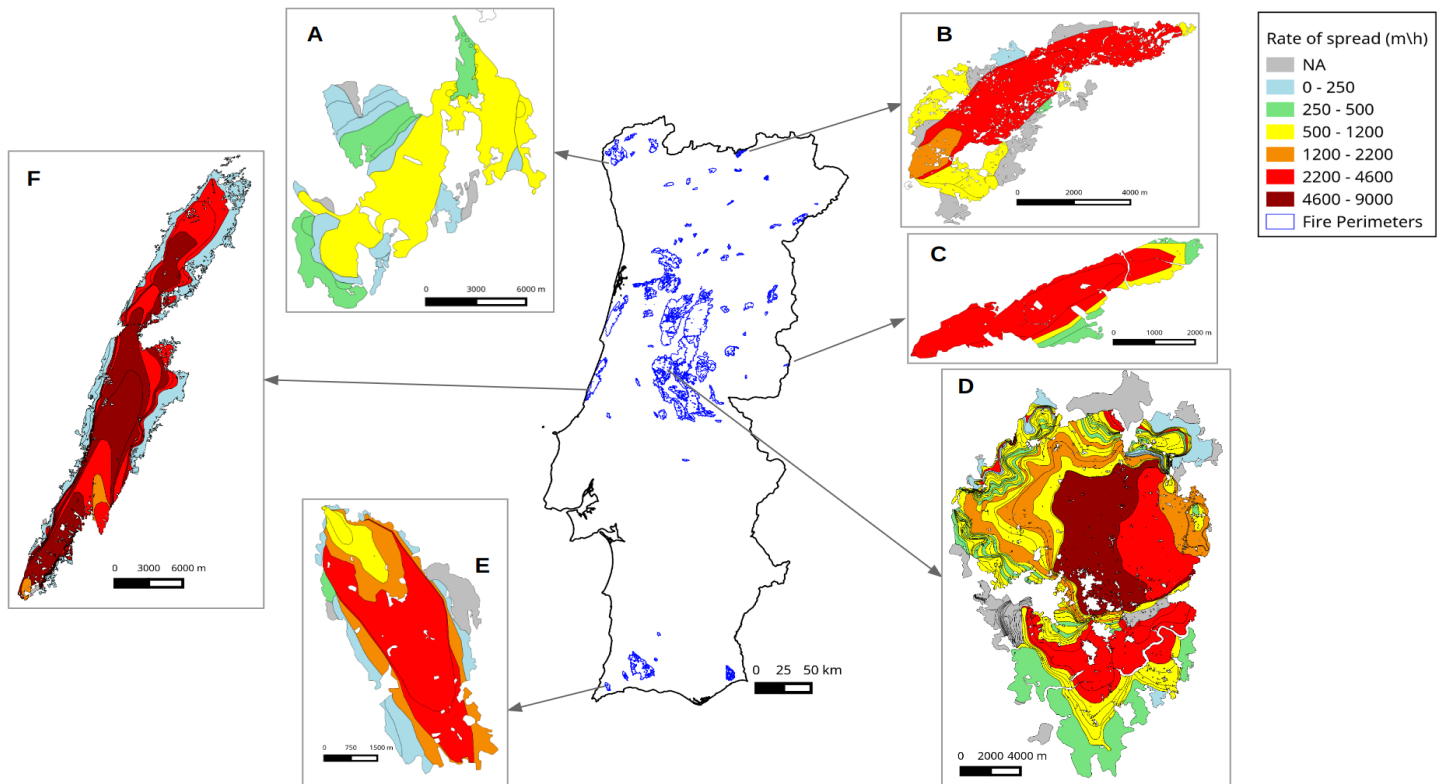
940

941

942

943

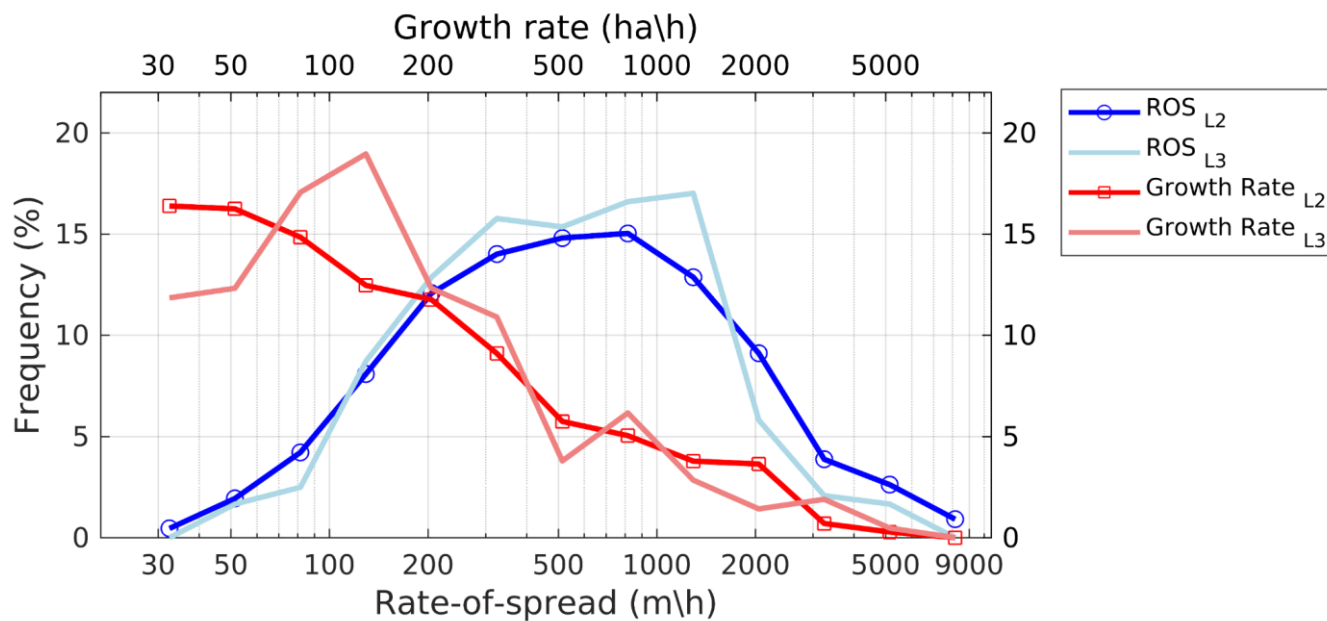
Figure 5: Example of how the fire behaviour descriptors are calculated based on the Proença-a-Nova (2020) wildfire: a) partial fire progression; b) procedure to calculate the distance for each vertex of the pair of consecutive polygons; and c) estimated main spread axis and associated fire behaviour descriptors.



44
45
46
47
48

Figure 6: Overall spatial distribution of the wildfire perimeters in the PT-FireSprd database, with examples of ROS estimates for 6 wildfires: A-Paredes de Coura (2016); B-Chaves (2020); C-Idanha-a-Nova (2020); D-Pedrógão Grande (2017); E-Aljezur (2020); F-Alcobaça (2017).

949
950
951



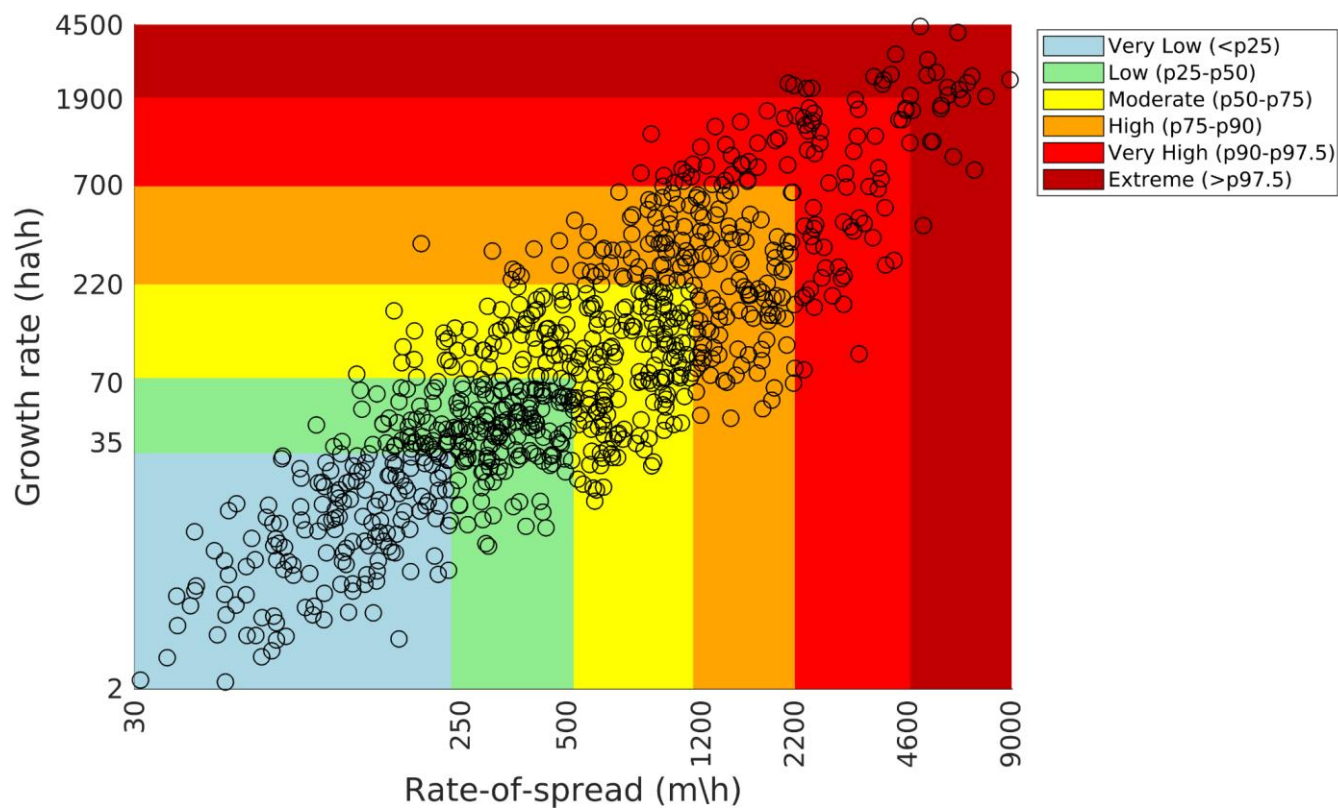
952

953

954

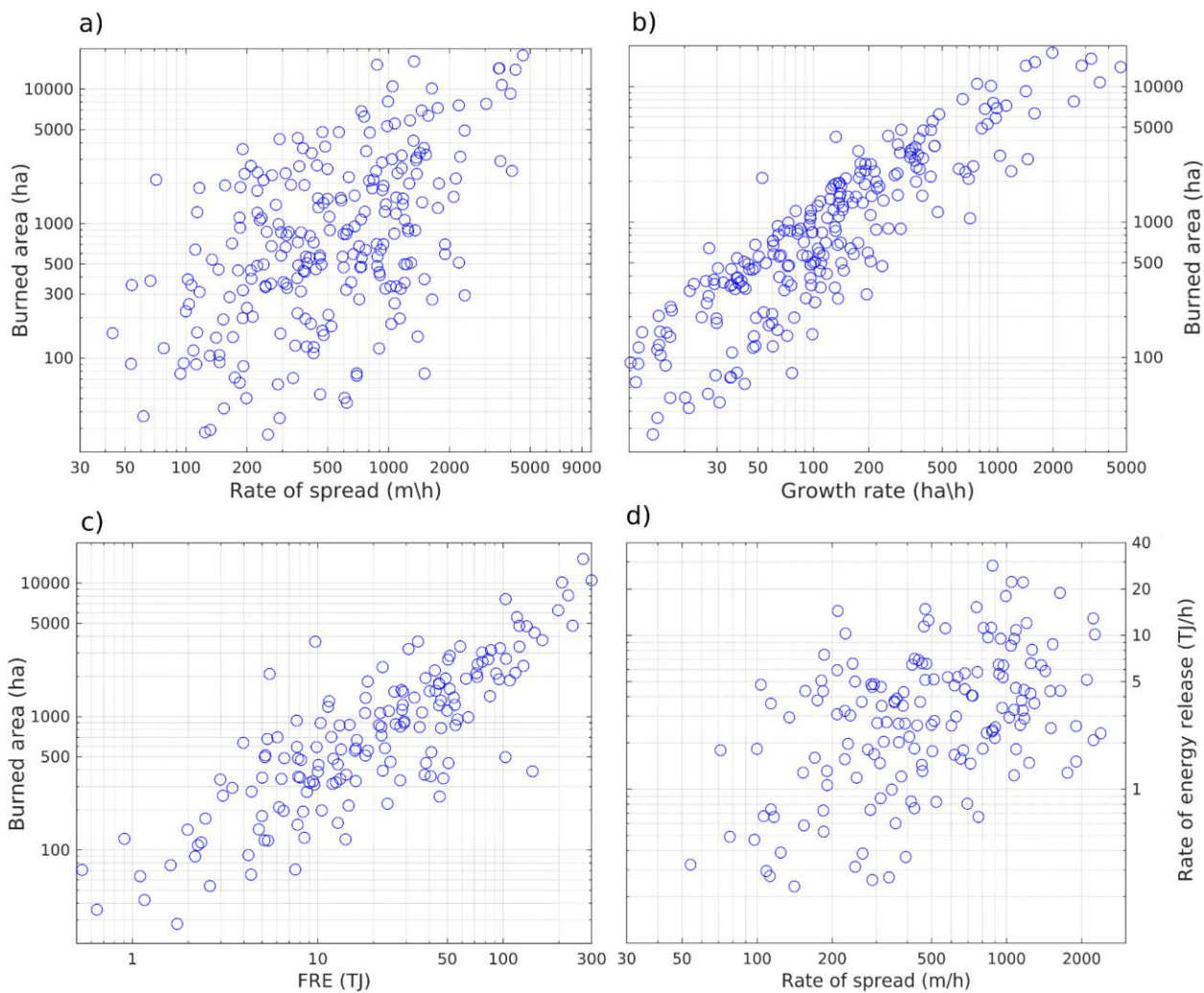
Figure 7: Histogram of the estimated ROS and FGR for L2 and L3 data (in log-scale). Each point represents the frequency in evenly spaced bins on a logarithmic scale.

955



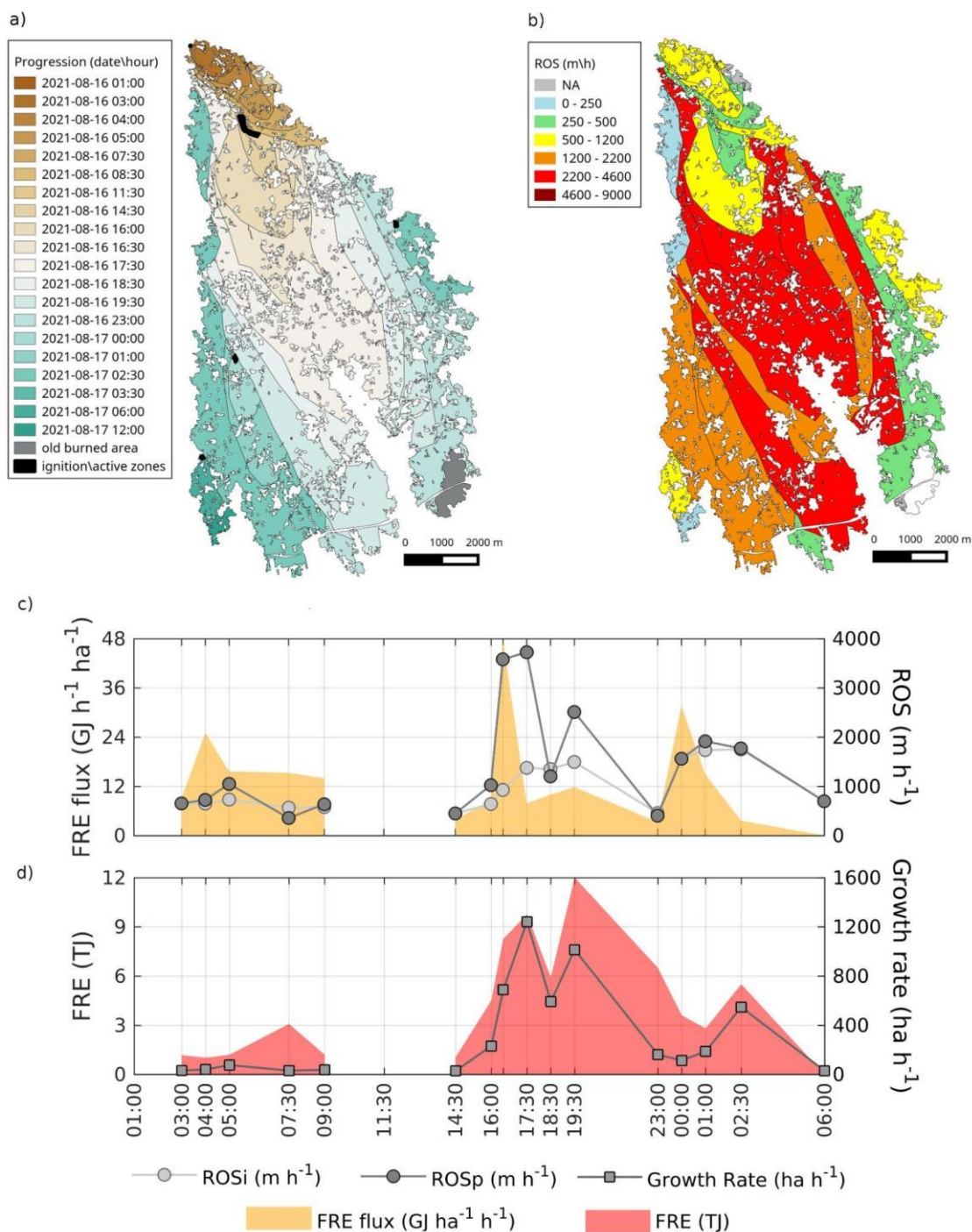
956
957

958 **Figure 8: Distribution of the estimated partial rate-of-spread (ROSp) and FGR (L2). Each point represents a wildfire progression**
959 **with at least 25 ha of extent. The percentiles were calculated for each variable separately (n=874).**



960
961

962 **Figure 9: Comparison between simplified wildfire behaviour descriptors (L3): burned area extent and ROS (a), burned area extent**
963 **and FGR (b), burned area extent and FRE (c), and ROS and average rate of energy release (d). The latter was calculated dividing**
964 **the total FRE by the burning period duration.**



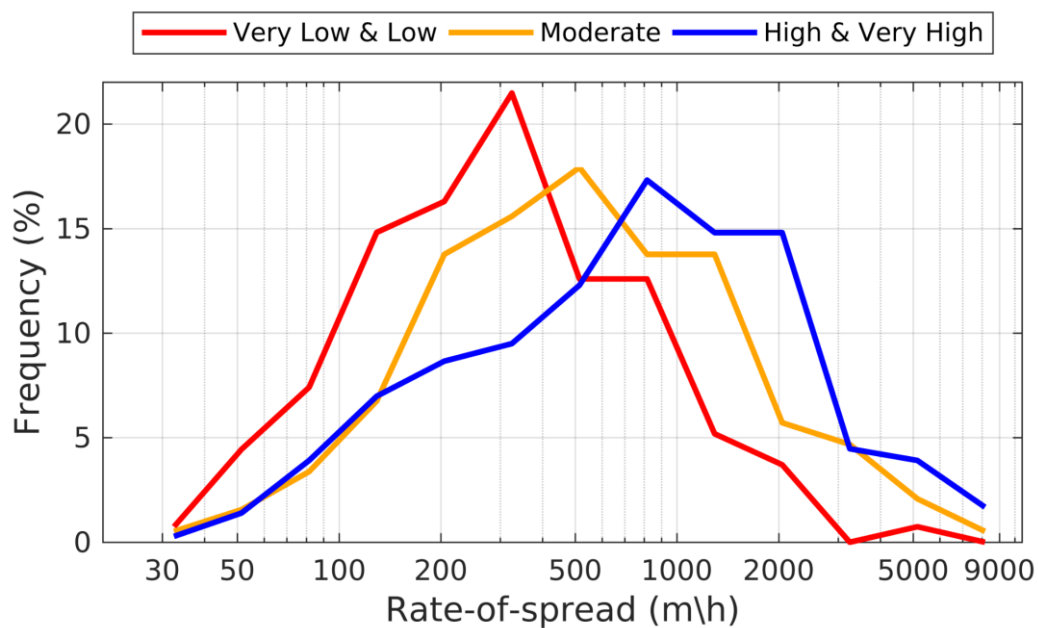
965
 966

967
 968
 969

Figure 10: The Castro Marim (2021) wildfire progression (a). Wildfire behaviour descriptors include: the spatial distribution of ROS (b); the temporal distribution of ROS and FRE flux rate (c); and the temporal distribution of FRE and FGR (d). Plots (c) and (d) start at 01:00 of the 16th of August and end at 06:00 of the 17th of August.

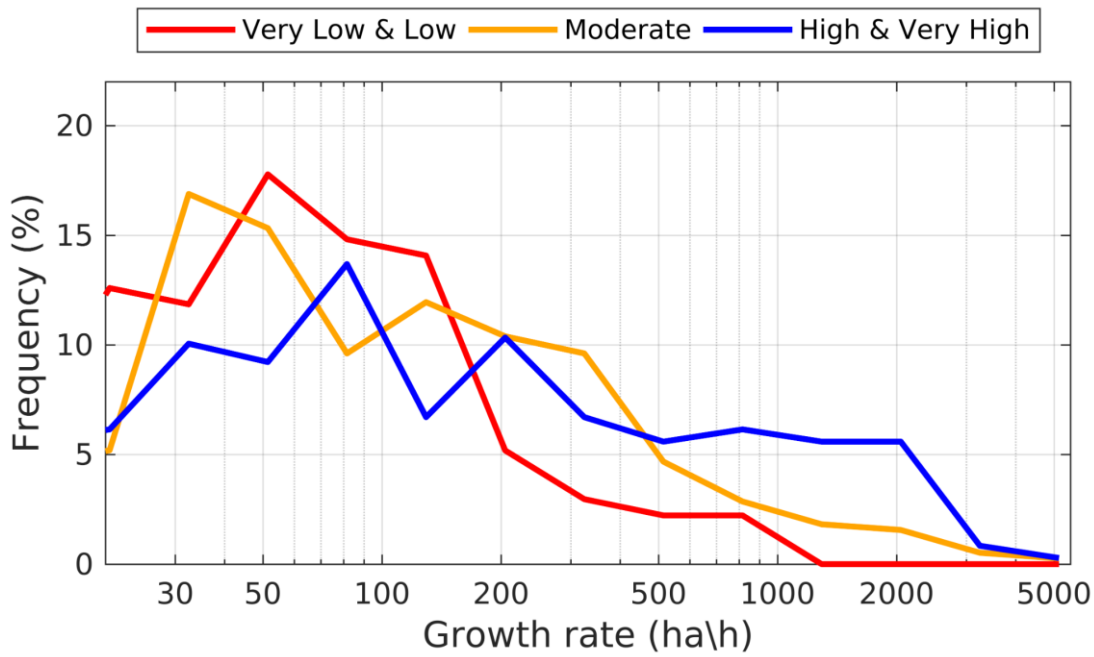


970



971

972 **Figure B1: Histogram of the estimated ROS (L2) for three aggregated levels of confidence. L2 ROS estimates were used and the**
973 **confidence flags are explained in Table A1.**



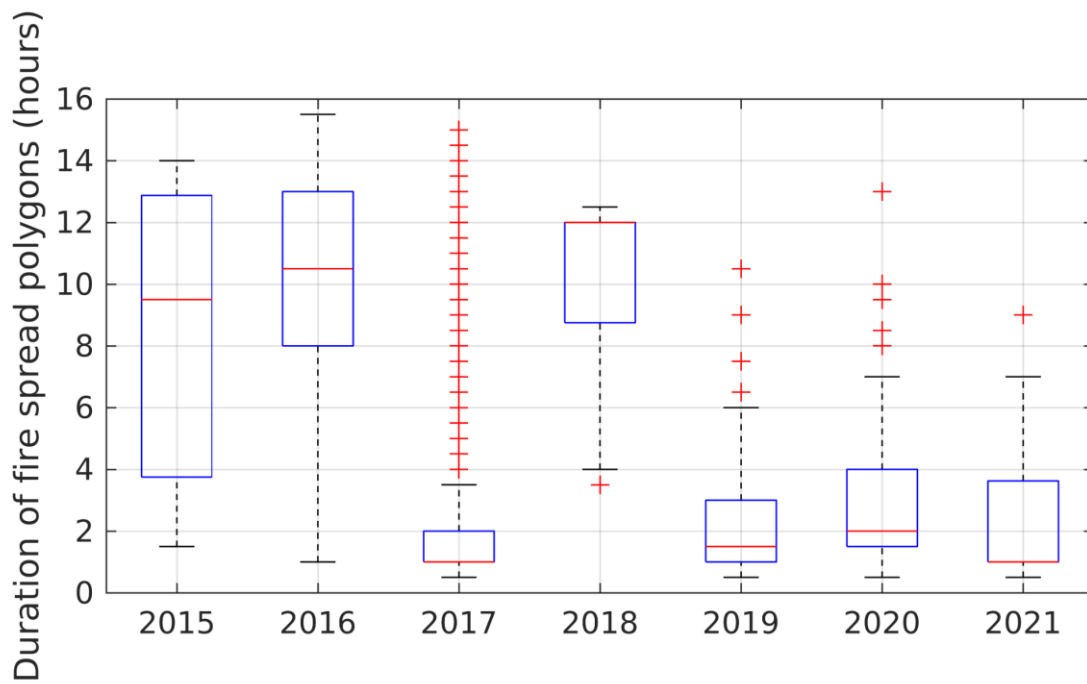
974

975
976

Figure B2: Histogram of the estimated FGR for three levels of confidence. L2 FGR estimates were used and the confidence flags are explained in Table A1.



977



978

979 **Figure B3: Distribution of the duration of the progression polygons divided by years**



980 **Tables**

981 **Table A1. Confidence flag value, class and interpretation. The flag is defined for each wildfire.**

Flag value	Flag Class	Interpretation
1	Very Low	The major fire progressions were observed only with satellite data, with important associated uncertainties.
2	Low	The major fire progressions were observed only with satellite data with moderate uncertainties
3	Moderate	The major fire progressions were observed with satellite data with low/moderate uncertainties and complemented with other sources.
4	High	The major fire progressions were at least partially observed with ground and airborne data, with relevant uncertainties associated (e.g. the exact hour of an important progression, or a flank position, etc)
5	Very High	The major fire progressions were observed with ground and airborne data with low uncertainties

982



983

Table A2. Database metadata list for L1

ID	Fire Name	Municipality	Civil Parish	Start Date	End Date	Extent (ha)	Confidence flag	ANEPC incident ID	P1	P2
1	Gouveia_10082015	Gouveia	Mangualde da Serra	2015-08-10	2015-08-12	2513	2	2015090024014	99	86
2	Oleiros_03082015	Oleiros	Alvaro	2015-08-03	2015-08-04	853	2	2015050020535	100	95
3	VilaNovadeCerveira_08082015	Vila Nova de Cerveira	Candemil	2015-08-08	2015-08-09	2988	3	2015160019994	87	87
4	Agueda_08082016	Águeda	Préstimo	2016-08-08	2016-08-12	7317	1	2016010058351	99	63
5	Anadia_10082016	Anadia	V.N. de Monsarros	2016-08-10	2016-08-12	3370	2	2016010059055	97	80
6	ArcosdeValdevez_08082016	Arcos de Valdevez	Cabana Maior	2016-08-08	2016-08-11	5806	1	2016160022311	93	71
7	Arouca_08082016	Arouca	Janarde	2016-08-08	2016-08-14	23547	2	2016010058554	97	96
8	Boticas_05092016	Boticas	Codecoso	2016-09-05	2016-09-07	1694	3	2016170021732/ 2016170021835	97	97
9	CabeceirasdeBasto_06092016	Cabeceiras de Basto	Rio Douro	2016-09-06	2016-09-07	1336	2	2016030067614	100	100
10	Caminha_09082016	Caminha	Argela	2016-08-09	2016-08-11	1628	1	2016160022551	99	61
11	Cinfaes_07082016	Cinfães	Cinfães	2016-08-07	2016-08-08	567	1	2016180042605	95	95
12	Cinfaes_08082016	Cinfães	Oliveira do Douro	2016-08-08	2016-08-09	756	2	2016180042656	100	100
13	FreixodeEspadaaCinta_06092016	Freixo de Espada a Cinta	Freixo Espada à Cinta e Mazouco	2016-09-06	2016-09-07	5194	3	2016040027372	99	97
14	Moncao_06092016	Monção	Riba de Mouro	2016-09-06	2016-09-07	656	2	2016160025950	71	58
15	Moncao_09082016	Monção	Barroças e Taias	2016-08-09	2016-08-11	1115	1	2016160022460	77	77
16	ParedesdeCoura_07082016	Paredes de Coura	Meixedo	2016-08-07	2016-08-12	10457	2	2016160022456	100	96
17	PontedeLima_08082016	Ponte de Lima	Calheiros	2016-08-08	2016-08-09	739	1	2016160022390	91	75
18	SeverdoVouga_09082016	Sever do Vouga	Pessegueiro do Vouga	2016-08-10	2016-08-12	1818	3	2016010058973	96	94
19	VieiradoMinho_10082016	Vieira do Minho	Rossas	2016-08-10	2016-08-11	1637	2	2016030060428	99	96
20	Resende_17082017	Resende	S. Martinho de Mouros	2017-08-17	2017-08-21	544	1	2017180043566	84	38
21	RibeiradePena_15082017	Ribeira de Pena	Cerva	2017-08-15	2017-08-16	507	1	2017170021591	100	100
22	CastroDaire_05102017	Castro Daire	Almfala	2017-10-05	2017-10-05	701	2	2017180054022	99	99
23	Mortagua_07102017	Mortagua	Espinho	2017-10-07	2017-10-08	961	2	2017180054507	99	99
24	Mirandela_16072017	Mirandela	Alvites	2017-07-16	2017-07-17	949	2	2017040020105	100	88
25	Pombal_06102017	Pombal	Abiul	2017-10-06	2017-10-07	1225	2	2017100054724	100	100
26	TorredeMoncorvo_18072017	Torre de Moncorvo	Acoreira	2017-07-18	2017-07-18	1536	3	2017040020365	100	100
27	Guarda_23082017	Guarda	Fernão Joanes	2017-08-23	2017-08-25	3457	3	2017090026098	91	91
28	Serta_08092017	Serta	Pedrogao Pequeno	2017-09-08	2017-09-09	4177	3	2017050027511	100	100



29	Abrantes_09082017	Abrantes	Aldeia do Mato	2017-08-09	2017-08-10	4357	3	2017140045924	83	79
30	CasteloBranco_23072017	Castelo Branco	Santo André das Tojeiras	2017-07-23	2017-07-28	4569	3	2017050023219	97	85
31	Serta_15102017_2	Serta	Pedrógão Pequeno	2017-10-15	2017-10-16	2320	3	2017050030728	54	54
32	CasteloBranco_13082017	Castelo Branco	Louriçal do Campo	2017-08-13	2017-08-15	6173	2	2017050025136	100	96
33	PampilhosadaSerra_06102017	Pampilhosa da Serra	Fajao	2017-10-06	2017-10-09	7217	2	2017060044928	97	96
34	Guarda_17072017	Guarda	Rochoso	2017-07-17	2017-07-18	7523	2	2017090021641	88	88
35	FigueiradaFoz_15102017	Figueira da Foz	Quiaios	2017-10-15	2017-10-17	15141	4	2017060046330	100	97
36	Oleiros_23082017	Oleiros	Cambas	2017-08-23	2017-08-25	7985	3	2017050026111	88	67
37	Gois_17062017	Gois	Alvares	2017-06-17	2017-06-22	15852	3	2017060026571	100	99
38	Alcobaca_15102017	Alcobaca	Pataias	2017-10-15	2017-10-16	18575	4	2017100056537 /2017100056554	100	100
39	Arganil_15102017	Arganil	Coja	2017-10-15	2017-10-16	31970	3	2017060046312 /2017090031521	100	99
40	Serta_15102017	Serta	Figueiredo	2017-10-15	2017-10-17	30974	4	2017050030693	97	97
41	Alvaiazere_11082017	Alvaiazere	Pussos	2017-08-11	2017-08-19	23715	2	2017100043917/ 2017050025201	99	52
42	PedrogaoGrande_17062017	Pedrogao Grande	Pedrogao Grande	2017-06-17	2017-06-19	29456	4	2017100032538	92	91
43	Serta_23072017	Serta	Várzea dos Cavaleiros	2017-07-23	2017-07-27	33401	3	2017050023195	97	96
44	Lousa_15102017	Lousã	Vilarinho	2017-10-15	2017-10-17	45249	4	2017060046260	100	95
45	Agueda_15102017	Agueda	Albitelhe	2017-10-15	2017-10-16	9095	3	2017180056272	83	78
46	OliveiraFrades_15102017	OliveiraFrades	Varzielas	2017-10-15	2017-10-17	9297	3	2017180056290	99	97
47	Monchique_03082018	Monchique	Monchique	2018-08-03	2018-08-08	26227	3	2018080033743	93	82
48	Agueda_05092019	Agueda	Macinhata do Vouga	2019-09-05	2019-09-06	1602	3	2019010072794	89	84
49	Alijo_24072019	Alijo	Vila Verde	2019-07-24	2019-07-24	574	5	2019170019467	100	100
50	Baiao_04092019	Baião	Teixeira	2019-09-05	2019-09-06	728	3	2019130150620	75	73
51	Nisa_01082019	Nisa	Tolosa	2019-08-01	2019-08-01	712	5	2019120016787	99	98
52	Ourique_10062019	Ourique	Monte Lavarjao	2019-06-10	2019-06-10	554	5	2019020015472	75	75
53	Penedono_21072019	Penedono	Beselga	2019-07-21	2019-07-23	736	4	2019180039496	99	99
54	Sabugal_29082019	Sabugal	Vale Mourisco	2019-08-29	2019-08-29	578	5	2019090029579	100	100
55	Serta_13092019	Sertã	Marmeleiro	2019-09-13	2019-09-14	676	4	2019050028005	100	90
56	Tomar_03082019	Tomar	São Pedro Tomar	2019-08-03	2019-08-03	511	4	2019140045796	86	73
57	Valenca_04092019	Valença	Cerdal	2019-09-04	2019-09-05	642	1	2019160026115	83	83
58	Valpacos_13092019	Valpaços	Ervões	2019-09-13	2019-09-13	738	2	2019170026369	56	56
59	ViladeRei_20072019	Vila de Rei	Fundada	2019-07-20	2019-07-22	9305	3	2019050022178	99	99
60	MirandadoCorvo_13092019	Miranda do Corvo	Moinhos	2019-09-13	2019-09-14	540	3	2019060042282	96	96



61	Fundao_07082020	Fundão	Capinha	2020-08-07	2020-08-08	472	4	2020050018968	87	85
62	Silves_06072020	Silves	Boião	2020-07-06	2020-07-06	520	4	2020080025576	77	77
63	Avis_21072020	Avis	Montes Juntos	2020-07-21	2020-07-21	698	5	2020120014122	95	95
64	IdanhaaNova_30062020	Idanha-a-Nova	Salvaterra do Extremo	2020-06-30	2020-06-30	728	4	2020050015270	100	100
65	SaoJoaoPesqueira_10072020	São João da Pesqueira	Riodades	2020-07-10	2020-07-11	770	4	2020180031783	97	94
66	Fundao_06082020	Fundao	Bogas Baixo	2020-08-06	2020-08-06	749	5	2020050018872	96	96
67	PortoMos_06092020	Porto de Mós	Codacal	2020-09-06	2020-09-07	998	4	2020100046280	97	91
68	OliveiraFrades_07092020	Oliveira de Frades	Antelas	2020-09-07	2020-09-08	1902	3	2020180044235	86	73
69	Aljezur_19062020	Aljezur	Bordeira	2020-06-19	2020-06-20	2243	5	2020080023014	99	93
70	Sernancelhe_06082020	Sernancelhe	Lapa	2020-08-06	2020-08-06	2213	5	2020180037681	100	100
71	Chaves_30072020	Chaves	Vila Verde da Raia	2020-07-30	2020-07-31	2508	3	2020170018342	83	82
72	Oleiros_25072020	Oleiros	Sardeiras de Baixo	2020-07-25	2020-07-27	5564	3	2020050017687	95	92
73	ProencaaNova_13092020	Proenca-a-Nova	Cunqueiros	2020-09-13	2020-09-14	14568	4	2020050022403	91	91
74	CasteloBranco_29082020	Castelo Branco	Ponsul	2020-08-29	2020-08-29	315	4	2020050021105	100	92
75	CastroDaire_07092020	Castro Daire	Cujo	2020-09-07	2020-09-07	452	4	2020180044155	76	76
76	Odemira_18082021	Odemira	João Martins	2021-08-18	2021-08-19	944	5	2021020019189	100	98
77	CastroMarim_16082021	Castro Marim	Pernadeira	2021-08-16	2021-08-17	5956	5	2021080035488	100	99
78	Monchique_17072021	Monchique	Tojeiro	2021-07-17	2021-07-18	1900	4	2021080029244	99	99
79	FreixoEspadaaCinta_20082021	Freixo de Espada à Cinta	Lagoaça	2021-08-20	2021-08-20	412	4	2021040023667	71	71
80	Mogadouro_20072021	Mogadouro	Tó	2021-07-20	2021-07-20	253	5	2021040019425	99	98

984 p1: stands for percentage of known fire progression (%); p2: stands for percentage fire behaviour descriptors calculated (%)



985 **Table A3. Attribute fields of the fire progressions (L1)**

Field	Description	Possible values
id	Polygon ID	>0
type	Type of Spread Polygon	p - wildfire progression ; z - ignition or active flaming zone ; a - previously burned area
date_hour	Date and hour of the polygon	yyyy-mm-dd hh:mm; uncertain ; na (not applicable)
source	Source of the data	fserv - forest service ; sat - satellite data ; airb - airborne data; fops - fire personnel; ek - expert knowledge; rep - external reports
zp_link	Numerical link between a ignition or active flaming zone (“z”) polygon and a wildfire progression (“p”) polygon	1,2,3... - the link between types "p" and "z" with known dates and hours; 0 - used for type "a" or when progression in "uncertain" or when the link between "p" and "z" is unknown
burn_period	Burning period	1,2,3,..; 0 for the same cases as “zp_link”.

986



987 **Table A4. Attribute fields of the fire behaviour database (L2)**

Field	Description	Possible values
fid	Fire ID	1-80*
fname	Fire Name	Municipality_StartDate (e.g. Gouveia_10082015)
year	Year	2015-2021*
type	Type of Spread Polygon	p - wildfire progression ; z - ignition or active flaming zone ; a - previously burned area
sdate	Start date and hour of the polygon	yyyy-mm-dd hh:mm; uncertain ; na (not applicable)
edate	End date and hour of the polygon	yyyy-mm-dd hh:mm; uncertain ; na (not applicable)
inidoy	Start day-of-year of the polygon (hours in decimal values)	1 to 366; -1 for uncertain progression polygons, polygons with unknown zp_link and previously burned areas
endday	End day-of-year of the polygon (hours in decimal values)	1 to 366; -1 for uncertain progression polygons, polygons with unknown zp_link and previously burned areas
source	Source of the data	fserv - forest service ; sat - satellite data ; airb - airborne data; fops - fire personnel; ek - expert knowledge; rep - external reports
zp_link	Numerical link between a ignition or active flaming zone ("z") polygon and a wildfire progression ("p") polygon	1,2,3... - the link between types "p" and "z" with known dates and hours; 0 - used for type "a" or when progression in "uncertain" or when the link between "p" and "z" is unknown
burn_period	Burning period	1,2,3,...; 0 for the same cases as "zp_link".
area	Burned area extent (ha)	> 0 for progression polygons, -1 for ignition or active flaming zones.
growth_rate	Fire growth rate (ha/h)	>0 for progression polygons with zp_link value >0; -1 for previously burned areas or uncertain progression polygons
ros_i	Average rate-of-spread (m/h) calculated since ignition\active flaming areas or a progression marking the start of the burning period	>0 for progression polygons with zp_link value >0; -1 for previously burned areas or uncertain progression polygons
ros_p	Parcial rate-of-spread (m/h) calculated between consecutive ignition\active flaming areas and progression polygon, or between two consecutive progression polygons	>0 for progression polygons with zp_link value >0; -1 for previously burned areas or uncertain progression polygons
spdir_i	Spread direction associated with "ros_i" (° from North)	0 to 359.99; -1 for the same cases in "ros_i"
spdir_p	Spread direction associated with "ros_p" (° from North)	0 to 359.99; -1 for the same cases in "ros_p"
duration_i	Duration (hours) associated with the "ros_i" metric	>0 known progression polygons; -1 for ignition\active flaming zones, previously burned áreas or uncertain progression polygons



duration_p	Duration (hours) associated with the “ros_p” metric	>0 known progression polygons; -1 for ignition/active flaming zones, previously burned areas or uncertain progression polygons
qc	Confidence flag for each wildfire	See table A1
FRE	Fire Radiative Energy (TJ)	>0 for known progressions with at least 70% of FRE observations between “sdate” and “edate”; - 1 for the remaining polygons
FRE_flux	Fire Radiative Energy flux (TJ ha ⁻¹ h ⁻¹)	>0 for known progressions with at least 70% of FRE observations between “sdate” and “edate”; - 1 for the remaining polygons
FRE_perc	Percentage of FRE observations between “sdate” and “edate”	Between 0 and 100 for known progression polygons; -1 for the remaining.

988 * values will change when the database will be updated with new wildfires.



989

Table A5. Attribute fields of the simplified fire behaviour database (L3)

Field	Description	Possible values
fid	Fire ID	1-80*
fname	Fire Name	Municipality_StartDate (e.g. Gouveia_10082015)
burn_period	Burning period	≥1
year	Year	2015-2021*
sdate	Start date and hour of the burning period	yyyy-mm-dd hh:mm; “na” for burning periods which only have progression polygons with unknown “zp_link” (see Table A4)
edate	End date and hour of the burning period	yyyy-mm-dd hh:mm; “na” for burning periods which only have progression polygons with unknown “zp_link” (see Table A4)
inidoy	Start day-of-year of the burning period (hours in decimal values)	1 to 366; -1 for burning periods which only have progression polygons with unknown “zp_link” (see Table A4)
endday	End day-of-year of the burning period (hours in decimal values)	1 to 366; -1 for burning periods which only have progression polygons with unknown “zp_link” (see Table A4)
qc	Confidence flag for each wildfire	See table A1
area	Burned area extent (ha)	>0
growth_rate	Average fire growth rate (ha/h)	>0; -1 for burning periods which only have progression polygons with unknown “zp_link” (see Table A4)
ros	Average rate-of-spread (m/h)	>0; -1 for burning periods which only have progression polygons with unknown “zp_link” (see Table A4)
max_ros	Maximum rate-of-spread (m/h) observed in the burning period	>0; -1 for burning periods which only have progression polygons with unknown “zp_link” (see Table A4)
spdir	Spread direction associated with “ros_i” (° from North)	0 to 359.99; -1 for burning periods which only have progression polygons with unknown “zp_link” (see Table A4)
duration	Duration (hours) of the burning period	>0; -1 for burning periods which only have progression polygons with unknown “zp_link” (see Table A4)
FRE	Fire Radiative Energy (TJ)	>0 for known progressions with at least 70% of the area burned during the burning period covered with FRE estimates; - 1 for the remaining polygons
FRE_flux	Fire Radiative Energy flux (TJ ha ⁻¹ h ⁻¹)	>0 for known progressions with at least 70% of the area burned during the burning period covered with FRE estimates; - 1 for the remaining polygons
FRE_perc	Percentage of FRE observations between “sdate” and “edate”	Between 0 and 100

990 * values will change when the database will be updated with new wildfires.