SPECIAL ISSUE | RESEARCH PAPER
https://doi.org/10.1071/WF22123

International Journal of WILDLAND FIRE
International Association of Wildland Fire

Modelling the daily probability of lightning-caused ignition in the Iberian Peninsula

Marcos Rodrigues[A,B,*], Adrián Jiménez-Ruano[A,B], Pere Joan Gelabert[C], Víctor Resco de Dios[D,E,F], Luis Torres[G], Jaime Ribalaygua[G] and Cristina Vega-García[C]

For full list of author affiliations and declarations see end of paper

*Correspondence to:
Marcos Rodrigues
Department of Geography and Land Management, University of Zaragoza, Pedro Cerbuna 12, 5009, Zaragoza, Spain
Email: rmarcos@unizar.es

Received: 30 June 2022
Accepted: 27 December 2022
Published: 24 January 2023

Cite this:
Rodrigues M *et al.* (2023)
International Journal of Wildland Fire
32(3), 351–362. doi:10.1071/WF22123

ABSTRACT

Background. Lightning is the most common origin of natural fires, being strongly linked to specific synoptic conditions associated with atmospheric instability, such as dry thunderstorms; dry fuels are required for ignition to take place and for subsequent propagation. **Aims.** The aim was to predict the daily probability of ignition by exploiting a large dataset of lightning and fire data to anticipate ignition over the entire Iberian Peninsula. **Methods.** We trained and tested a machine learning model using lightning strikes (>17 million) in the period 2009–2015. For each lightning strike, we extracted information relating to fuel condition, structural features of vegetation, topography, and the specific characteristics of the strikes (polarity, intensity and flash density). **Key results.** Naturally triggered ignitions are typically initiated at higher elevations (above 1000 m above sea level) under conditions of low dead fuel moisture (<10–13%) and moderate live moisture content (Drought Code > 300). Negative-polarity lightning strikes (−10 kA) appear to trigger fires more frequently. **Conclusions and implications.** Our approach was able to provide ignition forecasts at multiple temporal and spatial scales, thus enhancing forest fire risk assessment systems.

Keywords: fire danger, forecast, fuel moisture, Iberian Peninsula, ignition probability, lightning strike, machine learning, wildfires.

Introduction

Lightning is the most frequent source of natural fires worldwide, shaping fire regimes and ecosystems over millennia prior to the advent of human influence and use of fire (Bowman *et al.* 2009; Fernandes *et al.* 2022). In Europe, natural-caused fires represent a small percentage of the total number of fires, typically below 5–10% (San-Miguel-Ayanz *et al.* 2012). In contrast to anthropogenic fires, which occur in a spatial pattern close to the human footprint (e.g. road networks, agricultural lands or the wildland–urban interface), the spatial location of natural fires is often clustered in specific hotspots of intense lightning activity and fuel availability (Ganteaume *et al.* 2013; Nampak *et al.* 2021). These locations experience above-average burned area attributed to lightning-caused fires, which are particularly abundant in certain regions of the Iberian Peninsula compared with other Mediterranean regions (Dijkstra *et al.* 2022).

Lightning fires frequently start in regions distant from human habitation and trigger multiple simultaneous events, which adds an additional layer of difficulty to suppression (Rodrigues *et al.* 2019*a*) and increases their likelihood of growing into large fire events when fuel and relief conditions facilitate fire spread (Pineda and Rigo 2017). The stages conducive to wildfire from a lightning source – strike, ignition, survival and arrival – depend on a specific set of factors and conditions (Anderson 2002). In short, a flash with sufficient peak current and duration is needed to initiate the smouldering phase. Fuel moisture, availability and type modulate the chances of survival, along with fire-prone weather conditions (e.g. lack of or small rainfall episodes). Under such conditions, lightning-caused smouldering may begin to flame and initiate a wildfire event (Pineda and Rigo 2017).

Most climate predictions envisage a decrease in fuel moisture content while fuel accumulation due to land abandonment is creating the perfect conditions for larger and more extreme fires across the Western Mediterranean Basin (Moreira *et al.* 2020, p. 202). Although further research is needed, there is already evidence of an increase in the occurrence of natural fires related to the increased frequency of dry storms and heat waves (Coogan *et al.* 2020; Li *et al.* 2020), potentially increasing fire frequency, size and intensity (Turco *et al.* 2018; Dupuy *et al.* 2020; Barros *et al.* 2021). During the summer of 2022, unprecedented lightning-caused fire activity was observed in some Mediterranean-type regions like California or Spain while some studies warn about lightning-caused fires becoming larger (Dowdy and Mills 2012) and more frequent (Cattau *et al.* 2020) in Australia. Understanding the linkages and feedbacks between environmental and atmospheric conditions that ultimately relate to lightning-caused fires is relevant to better understand fire regimes (Barros *et al.* 2021; Cochrane and Bowman 2021) or develop fire risk assessments (Chuvieco *et al.* 2014).

Lightning-caused ignitions have attracted considerable attention worldwide, which has translated into a wide array of studies devoted to understanding lightning patterns and causes, their link with fire ignition or their future evolution under climate warming (Price and Rind 1994*a*, p. 199, 1994*b*). Without being exhaustive, we found examples in the USA (Cattau *et al.* 2020; Jiménez-Ruano *et al.* 2022), Canada (Wotton and Martell 2005; Wotton *et al.* 2010), China (Hu and Zhou 2014), Australia (Dowdy and Mills 2012) and Europe (Dijkstra *et al.* 2022). In the Iberian Peninsula, the region under study in the present work, lightning-caused fires have also been investigated historically (Vazquez and Moreno 1998; Castedo-Dorado *et al.* 2011; Nieto *et al.* 2012; Chuvieco *et al.* 2014; Couto *et al.* 2020). Numerous studies have focused on addressing the probability of fires caused by lightning, humans or both. Most studies dealing with natural fires are based on historical compilations of natural fires coupled with environmental and climatic variables (Nieto *et al.* 2012; Rodríguez-Pérez *et al.* 2020) and meteorological data to determine the atmospheric conditions conducive to lightning fires (Pérez-Invernón *et al.* 2021; Soler *et al.* 2021; Pineda *et al.* 2022) or the lightning features that foster fire events (Pineda *et al.* 2014, 2022; Couto *et al.* 2020; Soler *et al.* 2021). The interactions between atmospheric models and the rate of spread of wildfires have also been investigated by Couto *et al.* (2020). However, there is still a lack of homogeneity in the methods, variables and protocols to predict, and therefore anticipate, the occurrence of natural wildfires. Here, we advocate an approach that decouples the conditions that favour lightning strikes from the environments that favour ignition.

The present work aimed to predict the probability of ignition of natural fires by translating lightning records into spatial-temporal patterns of lightning ignition potential on the Iberian Peninsula. For this purpose, we analysed a large lightning strike dataset (more than 17 million observations) to ascertain the conditions under which a lightning strike may trigger an ignition, i.e. the probability of ignition conditional on the occurrence of the lightning strike itself. The model was initially trained and validated using data covering all peninsular Spain plus the Balearic Islands, and we subsequently investigated its ability to anticipate ignitions in Portugal. Several variables related to fuel moisture content, vegetation structure and topography, as well as some factors related to polarity, intensity and density of impacts were used. The methodology is based on Random Forest modelling and other related techniques (bootstrapping, resampling and data stratification) to assess the accuracy of the probability obtained. Most of the existing approaches have focused on either the atmospheric conditions surrounding lighting and ignitions (Pérez-Invernón *et al.* 2021; Soler *et al.* 2021; Pineda *et al.* 2022) or on developing stationary likelihood maps (Nieto *et al.* 2012; Rodríguez-Pérez *et al.* 2020). Instead, here we delve into the spatial and landscape circumstances enabling natural wildfires. In this way, daily and moderate-resolution lightning ignition prediction for the whole Iberian Peninsula can be provided.

Materials and methods

Dependent variable: lightning strikes and fire ignitions

The core of our approach leverages a large and comprehensive dataset of cloud-to-ground lightning strikes recorded in the period 2009–2015. The dataset was compiled by the Spanish Meteorological Agency (AEMET) and provided by Meteogrid SL. AEMET's lightning detection system for the Spanish territory is composed of 14 detectors on the Spanish mainland, 1 in the Balearic Islands and 5 in the Canary Islands (the latter were not used in this work); Meteo-France provides data from 10 of its detectors and the Portuguese Meteorological Agency adds 4 extra sensors. All data from these 34 detectors are consolidated in the calculation and tracking system at the AEMET's climate service. The system is capable of detecting lightning strikes with a location accuracy of the order of 1 hm and a detection efficiency of 90% of total lightning strikes. The dataset used in this study consists of more than 17 million records in the period 2009–2020, reporting flash intensity, polarity and density. Lightning strikes were paired with actual lightning ignition data compiled by national agencies. Information about fire ignitions (location, date, cause and size) was retrieved from the Spanish (Estadística General de Incendios Forestales, EGIF; MAAyMA 2015) and Portuguese (Dispositivo Especial de Combate a Incêndios Florestais, DECIF; Pereira *et al.* 2011) databases. The Spanish dataset was used to identify those lightning strikes that started a fire while the Portuguese records were retained to evaluate the applicability of the model to the specific conditions of Portugal.

The response variable was constructed using historical ignitions in Spain. Approximately 2500 fires were matched to a lightning strike occurring within a radius of 1000 m and no more than 3 days after the ignition date (maximum hold-over period according to Pineda *et al.* 2014). For each fire event, we retrieved all lightning strikes within a 3-day window. Then, we ranked lightning strikes by proximity to the fire's ignition point. The closest lightning was associated with that ignition and further considered as lightning that triggered a fire when building the response variable. Only those flashes within a 1000-m distance from an ignition source were considered owing to the uncertainty in flash detection. From this information, we built a binary response variable combining strikes that triggered a fire (1 or presence) or were not related to a fire (0 or absence). The response variable comprised both lightning strikes on days where a wildfire occurred (fire days henceforth) and a balanced set of strikes from days without observed lightning-caused fires (non-fire days). We implemented a stratified sampling procedure to prevent undesired effects from the largely unbalanced response (more 0 s than 1 s) and minimise spatial autocorrelation.

Independent variables

To predict the probability of ignition, we employed factors related to vegetation, topography and others directly linked to lightning features. Live fuel moisture content was characterised using the Drought Code (DC; Van Wagner 1987) as proxy, calculated following the standard procedure described in Wagner and Pickett (1985). DC was retrieved from the Copernicus Emergency System, which produces daily gridded (at 25 km resolution) historical data for a variety of meteorological fire danger indices, including the Canadian Fire Weather Index, to which DC belongs. We calculated the daily dead fuel moisture content (dFMC) as described in Nolan *et al.* (2016), previously used successfully on the Iberian Peninsula (Boer *et al.* 2017; Resco de Dios *et al.* 2022). The method is based on the exponential decline in dFMC with increasing vapour pressure deficit (VPD) and, in turn, it calculates VPD from MODIS Land Surface Temperature (MOD11A1 Collection 6). The dFMC measures 1–10 h fine fuels moisture while DC relates to seasonal drought conditions potentially affecting the moisture content of live vegetation (Viegas *et al.* 2001). The dFMC was calculated using surface temperature and relative humidity daily data from the ERA5-Land dataset at a 9 km spatial resolution (Muñoz Sabater 2019). Secondly, elevation and relief curvature (derived from elevation via topographic position index; Weiss 2001) were selected to depict the influence exerted by the terrain in the ignition of natural-cause fires. We retrieved elevation data from the NASADEM global digital elevation model at 30 m spatial resolution (NASA JPL 2020). The last landscape feature incorporated into the model was vegetation height, acting as a proxy for vegetation structure and cover. We used the global forest canopy height map using a novel dataset calibrated combining Global Ecosystem Dynamics Investigation (GEDI) footprints of canopy structure (Dubayah *et al.* 2020) and Landsat Operational Land Imager spectral imagery. All variables were regridded from their original resolution to 1×1 km. Topographic variables and vegetation height were aggregated by their median value into the 1×1 km destination cell while dFMC and DC were resampled using bilinear interpolation downscaling from 9×9 to 25×25 km. Our model also accounts for lightning-related features, such as strike intensity and polarity, and density of flashes, as provided by the lightning detection network.

Model calibration with Random Forest

The model was trained using Random Forest (Breiman 2001), a popular and effective algorithm for wildfire distribution modelling (Bar Massada *et al.* 2013; Rodrigues and de la Riva 2014; Vecín-Arias *et al.* 2016; Su *et al.* 2018). We trained a total of 100 models exploring different iterations of the response variable. In each realisation, we extracted a balanced sample of 1-presence (flashes igniting a fire; 1/3 of the final sample) and 0-absence (lightning strikes non-conducive to fire; 2/3 of the final sample distributed as 1/3 in fire days and another 1/3 in non-fire days). The process was stratified using a 100×100 km grid so that at each grid cell, we extracted a balanced sample of 1-presence and 0-absence to reduce the potential spatial autocorrelation (Wang *et al.* 2016). It must be noted that we only analysed the April-to-October period, coinciding with the brunt of the fire season.

Model calibration was conducted using the *caret* package (Kuhn 2008) in the R environment for statistical computing (R Core Team 2021). For each model realisation, we split data into a calibration (70% of data) and testing dataset (30% of data). During the calibration phase, we implemented a 5-fold repeated cross-validation (three repetitions), optimising model parameters: number of trees (*ntree*; from 100 to 2000 trees with steps of 100) and number of variables to use at each split (*mtry*; consecutive values from 1 to 6). The minimum node size (*nodesize*) parameter was kept constant at its default value for regression (5) because it had little influence on performance. Once the model was calibrated, we evaluated the importance and explanatory sense of each covariate using partial dependence plots (Greenwell 2017), and the overall performance of each model via area under the receiver operating characteristic curve (AUC) (Bradley 1997), using the test sample in the latter and the *out-of-bag* sample in the first. Out of the set of 100 models, we selected the one attaining average prediction performance (AUC = 0.828; $mtry = 2$, $ntree = 1000$ and $nodesize = 5$) as the candidate model to forecast and map the daily evolution of lightning-related ignition probability. The model was subsequently evaluated calculating the AUC from the Portuguese fire data, which consisted in 105 lightning fires

(>1 ha) in the period 2009–2015. Like the procedure followed in the calibration stage, we created multiple random samples of 0-absence to account for the non-occurrence of fires in Portugal. In this case, we created a balanced sample of 105 non-fire locations on fire days, placing as many absences as fires were reported. This procedure was repeated 100 times to check the potential effect of the random placement of 0-absence. Finally, to perform daily spatial predictions over the entire Iberian Peninsula, we used daily information about dFMC and DC. Vegetation height and topography were kept constant. We set constant values of intensity, polarity and density of strikes equal to their respective average value over the study area and period (i.e. -10 kA and five discharges, respectively).

Results

Lightning caused fires in the Iberian Peninsula

Fig. 1 illustrates the comparison of fire size and seasonal distribution of lightning-caused fires in the region. Lightning-caused fires were significantly larger than human-caused fires either due to arson or negligence (Fig. 1*a*, *b*). Note that accidental and negligence fires were grouped together owing to differences in the cause classification systems between Portuguese and Spanish agencies. Both the average fire size (burned area divided by number of fire events) and the 95th percentile of the size distribution show the potential for larger burned area size that lightning fires have. The average fire size of lightning fires was at least two times larger than human-related ignitions, even three times in the case of Portugal (Table 1). The interannual comparison of fire size revealed the prominent role lightning fires played in the waves of fires during the years 1994 (Spain), 2003 and 2017 (Portugal). Their intra-annual distribution is clearly concentrated in the summer months (June–September), though in Spain they peak in August and July, whereas in Portugal, they are more evenly distributed during these summer months. (Fig. 1*c*). In any case, they seldom occur outside that temporal window, being scarce in April and October, though the agencies still reported some fires in both countries in these 2 months (1% of total).

The Random Forest models denoted a high predictive capability (calculated test samples) in Spain (mean AUC of 0.82; Fig. 1*a*) and the feasibility of extrapolating them to similar regions, as exemplified in Portugal (AUC = 0.74; Fig. 1*b*). The sensitivity of the models' performance and extrapolation potential to the sampling of the response variable (in the case of Spain) and to the random location of the background sample (in Portugal) was low. AUC values ranged between 0.80 and 0.85 in the first case and between 0.70 and 0.78 in the latter. The AUCs showed frequency distributions close to Gaussian, peaking at the means of the reported AUCs.

The major contributing factors (Fig. 2*c*, *d*), DC and dFMC, are related to fuel dryness and prolonged drought periods. DC and dFMC showed very clear profiles of association with predicted ignition probability. The former is related to increasing probability up to a DC value of 300 for which the predicted probability increases from 0.28 to 0.49, levelling off at 0.42 beyond the 600 mark. dFMC shows the reverse profile, decreasing rapidly from its maximum at 0.5 probability at 6% moisture content, decreasing to 0.34 towards 13%. After that threshold, the probability continues to show a general decline, stabilising at ~0.3 when fine dead fuels have more than 20% moisture content. Ignition probability was strongly linked to elevation and vegetation height, although to a lesser extent than for the former factors. The probability of ignition by lightning increased with elevation, from 27 at sea level to 0.4 at 1000 m elevation. Likewise, vegetation layers 10–15 m tall were more prone to ignition from lightning (0.45–0.48). Lastly, less important but still contributing was the density of sensor-detected bolts and their intensity and polarity. The probability of ignition increased almost linearly with respect to the density of detected flashes, with a probability greater than 0.4 when more than 10 flashes occurred. Polarity and intensity showed the most erratic profile. Negative flashes at -10 kA showed the highest probability. Most of the flashes that started a fire were observed around this range, as can be seen in the percentile lines on the x-axis in Fig. 2*h*. Relief curvature was not selected as a meaningful driver in any model.

Fig. 3 exemplifies the predicted probability over the Iberian Peninsula on 2 days with contrasting meteorological conditions. In general, the spatial patterns reflect the conditions described above, i.e. a higher probability in mountain ranges with a taller canopy layer. However, comparison between the pairs of predictions highlights the leading role played by meteorological conditions and fuel moisture. The late May 2013 prediction exemplifies the situation under mild off-season conditions, while the early June 2015 prediction illustrates the effect of a heat wave episode that boosts the probability of ignition across the Peninsula. The temporal variability can be also seen in Fig. 4*c*–*h*, with increasingly hazardous conditions from May to October. Likewise, the predicted probability varies yearly (Fig. 4*a*, *b*), though an underlying spatial pattern emerges from the comparison of all temporal aggregations matching the spatial footprint of the stationary drivers, i.e. elevation and vegetation height.

Discussion

The Iberian Peninsula is the most fire-affected region in Europe (Camia *et al.* 2013). In this work, we produced a daily forecast of natural-caused wildfire ignition probability across the Iberian Peninsula, a region that despite being dominated by human-caused fires stands out as one of the main hotspots of lightning fires in the Western Mediterranean

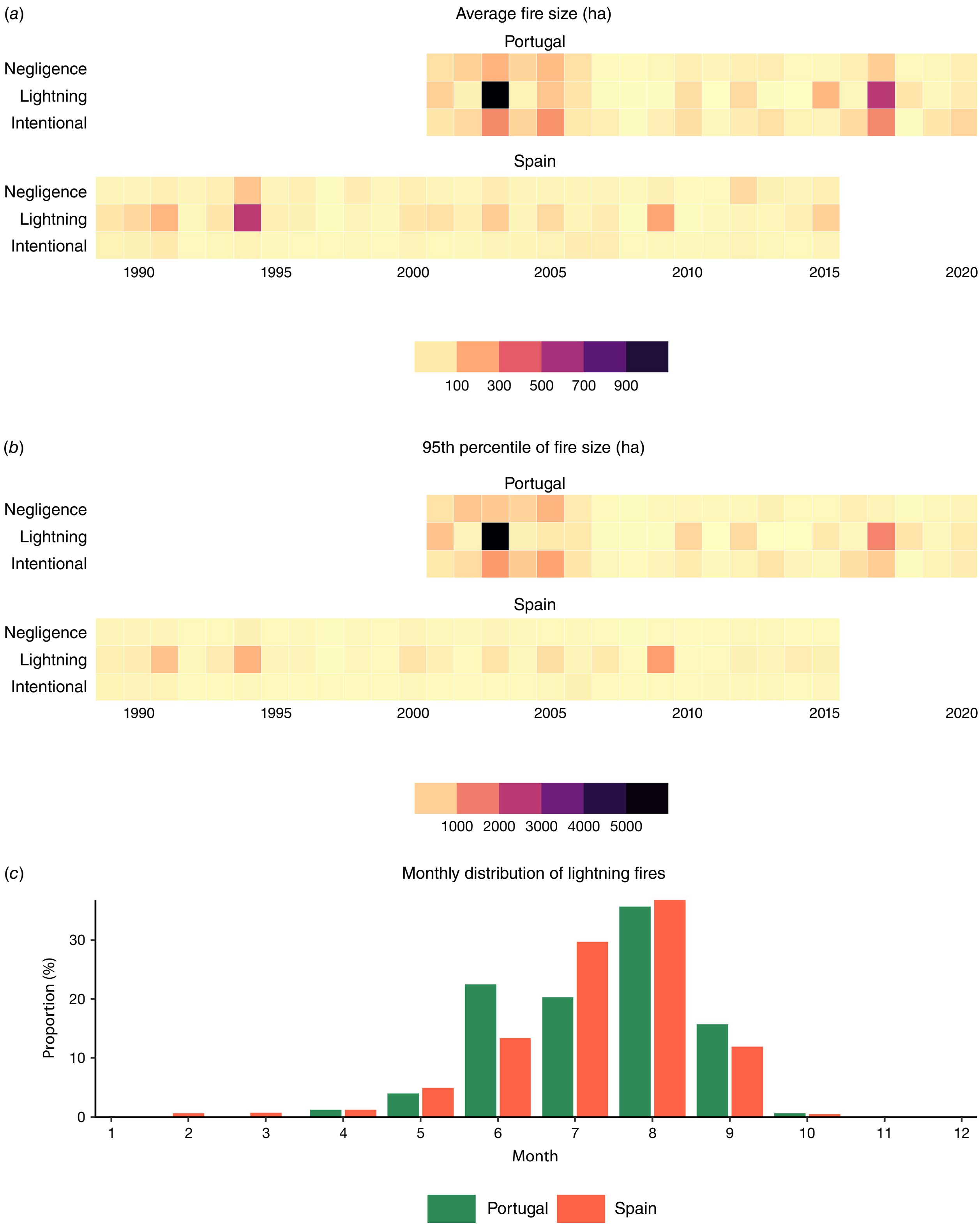

Fig. 1. (*a*) Average fire size (ha) by cause of ignition, year, and country. (*b*) 95th percentile of fire size (ha) by cause of ignition, year, and country. (*c*) Monthly distribution of lightning-caused fires by country. Only fires larger than 1 ha and with a known source of ignition were considered. The period of the data was chosen based on those years with reliable data in terms of classification schemes and compilation completeness (Pereira *et al.* 2011; MAAyMA 2015).

Table 1. Summary of fire activity by cause of ignition and country.

Country	Cause	*N*	Area (ha)	Avg. size (ha)	95th P size (ha)
Portugal (2001–2020)	Arson	9713	832 295	85.7	225
	Lightning	329	103 665	315.1	1022
	Negligence	15 108	637 720	42.2	106
Spain (1988–2015)	Arson	94 123	2 099 354	27.4	69.8
	Lightning	3538	331 352	88.7	176
	Negligence	32 710	1 097 640	39.0	75

Only fires larger than 1 ha and with a known source of ignition where considered. *N*, number of fires; Area, total burned area in hectares; Avg. size, average fire size (Area/*N*) in hectares; 95th P size, 95th percentile of fire size in hectares. Only fires larger than 1 ha were included.

Basin. The benefits from managing natural ignitions are well documented, for example, in guiding fire reintroduction in heavily fire-excluded forests resilient or adapted to fire (Barros *et al.* 2021; Rodrigues *et al.* 2022).

An insight into former models of lightning-caused fires on the Iberian Peninsula

Previous studies have set the basis for developing our methodological scheme to assess the probability of natural fire ignition over multiple temporal and spatial scales. Most of these were developed using binary classification approaches such as logit regression, assembling ignition locations along with environmental and climate-related drivers to derive the likelihood of ignition and identify the main driving forces (Castedo-Dorado *et al.* 2011; Nieto *et al.* 2012; Chuvieco *et al.* 2014). Castedo-Dorado *et al.* (2011) assessed the probability of ignition in the Spanish province (NUTS3, see https://ec.europa.eu/eurostat/web/nuts/background) of León, one of the major hotspots of lightning-caused fires. They employed binary logistic regression, analysing the stationary probability in the period 2002–2007, pointing to the confluence of thunderstorms over coniferous forests in high-altitude ranges as key factors, and reaching an AUC of 0.79. Using a similar approach but applied to two contrasting regions (the Autonomous Regions – NUTS2 – of Madrid and Aragón), Nieto *et al.* (2012) reached similar conclusions though with different levels of performance, the highest standing at AUC = 0.70. The method by Nieto *et al.* (2012) was subsequently extended over the entire Spanish mainland in Chuvieco *et al.* (2014). Their model was calibrated using categorical data on climate types, the number of lightning strikes and the Duff Moisture Code from the Canadian Fire Weather Index, giving an accuracy of 64.2%. Our proposal builds on these studies by delving into the temporal dimension.

About the suitability and performance of the proposed approach

We adopted a novel approach based on lightning strikes (Moris *et al.* 2020), in contrast to fire-based approaches (Amatulli *et al.* 2007; Nieto *et al.* 2012). That is, instead of building the response variable from the presence or absence of lightning-caused fires, we analysed the type of lightning and the landscape features conducive to fire, thus separating the ignition potential from the atmospheric conditions surrounding the occurrence of lightning and thunderstorms. In terms of risk assessment, we believe the strategy of breaking down ignition likelihood and the probability of lightning occurrence offers several advantages. Decoupling both phenomena allows us to better understand their relative importance while enabling embedding different models for lightning or thunderstorm forecast, or even real observations of lightning flashes. That is particularly important for short-term predictions. In turn, long-term modelling would also benefit from such an approach. Our forecast of ignition probability requires few weather-related factors (the same as are necessary for the calculation of the Canadian Fire Weather Index: rainfall, relative humidity, temperature and wind speed; Van Wagner 1987); thus, their prediction under climate scenarios (either Shared Socioeconomic Pathways or the former Representative Concentration Pathways) is already feasible. Along the same lines, several models are available to forecast lightning strikes (Woodard *et al.* 2014), hence giving the possibility to combine them with the ignition component. Furthermore, we translated model outcomes into spatial explicitly and scalable predictions as recommended in the literature (Vecín-Arias *et al.* 2016). Our modelling approach holds a high predictive accuracy (~0.82 AUC; Fig. 2*a*), which is noteworthy given the rare nature of thunderstorm-driven fires (Fernandes *et al.* 2021). Moreover, we demonstrated its suitability to perform forecasts in nearby fire-prone regions, evidenced by the successful validation in Portugal (Fig. 2*b*).

The driving factors of lightning-caused fires on the Iberian Peninsula

The observed relationships between ignition drivers and predicted likelihood are in line with their expected behaviour based on former modelling endeavours in the region. That is, low fuel moisture fosters ignition and higher chances of ignition are observed in mountain ranges. The

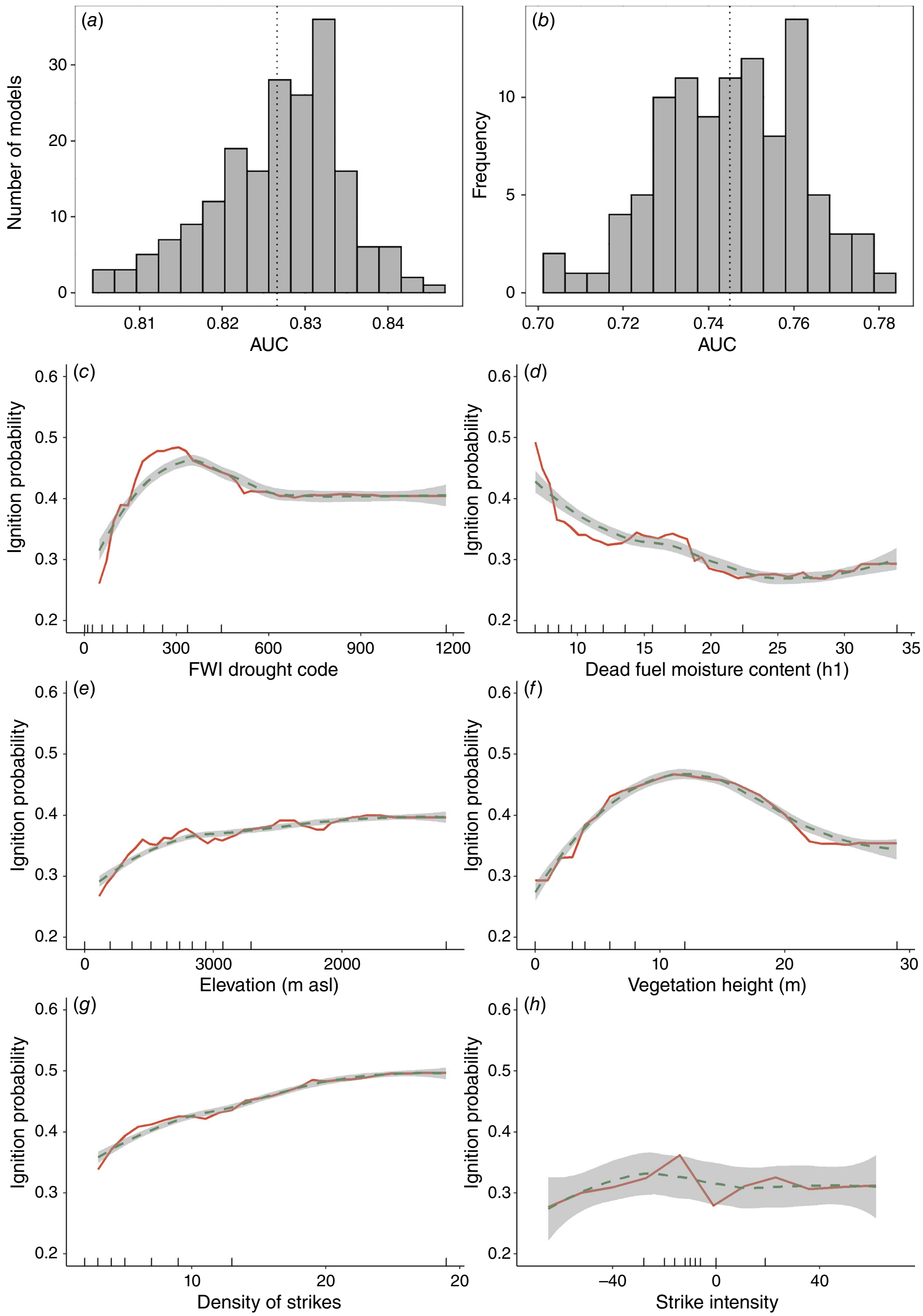

Fig. 2. Summary of variable and model performance. (*a*) AUC value of the original model for Spain. (*b*) AUC values of the validation in Portugal. (*c–h*) Partial dependence plots of the explanatory factors obtained for the average model. The brown solid line indicates the average value of the response across models. The dashed green line with grey shading shows the LOESS smoothed response and the 95 percent confidence interval range from the 100 model iterations, respectively. Rug lines along the x-axis indicate the 10-fold percentile position of the variable.

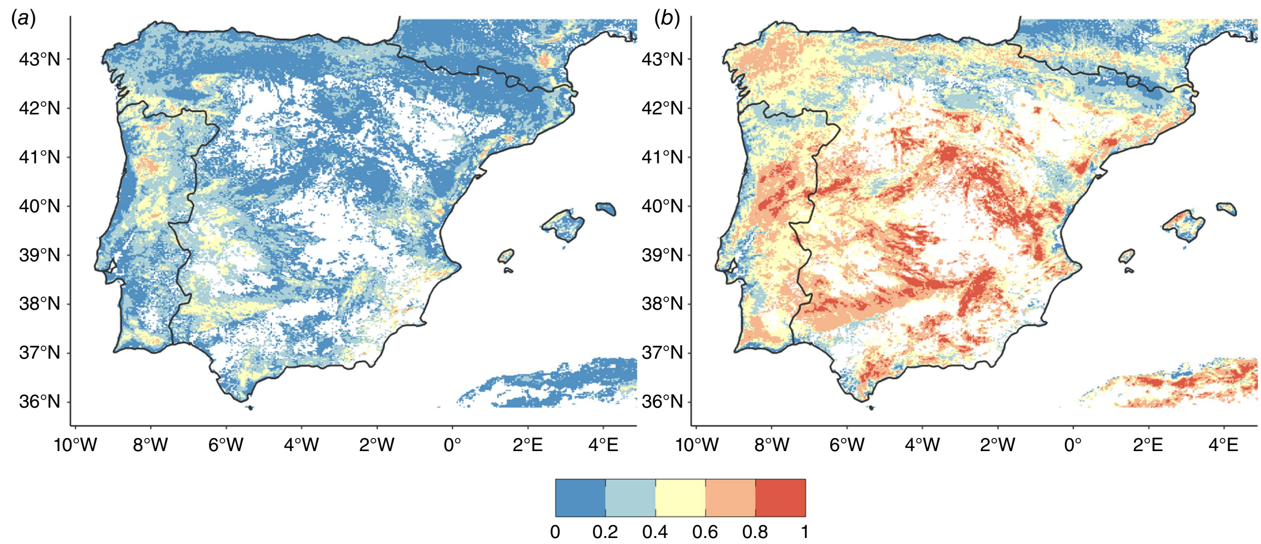

Fig. 3. Spatial distribution of lightning ignition probability in the Iberian Peninsula: (*a*) prediction for milder conditions during 26 May 2013; (*b*) fire-prone conditions during a heat wave episode on 7 June 2015.

important role of moisture content of the live/dead fuels in causing the ignition and propagation of new fire events has previously been documented in many works (Resco de Dios *et al.* 2021, 2022; Rodrigues *et al.* 2021; Baranovskiy and Kirienko 2022). Dry thunderstorms with a low precipitation rate preceded the largest lightning-caused fires recorded in the Mediterranean Basin (Pérez-Invernón *et al.* 2021). Indeed, the lack of rainfall is key during the survival phase of a fire (Pineda and Rigo 2017; Soler *et al.* 2021) and searching for these conditions has been at the core of many studies focused on the atmospheric and synoptic conditions surrounding lightning fires. But, according to our findings, it is the influence of fuel moisture and seasonal drought (DC) that seem more decisive, perhaps because they inherently involve reduced or complete lack of rainfall. It is nonetheless worth noting that the critical DC threshold was identified at a value of 300. According to the European Forest Fire Information System, this indicates low-to-moderate fire weather danger conditions, which suggests that lightning fires occur even during fairly short drought anomalies.

The type of vegetation was claimed as the most relevant factor for lightning ignition by Rodríguez-Pérez *et al.* (2020). We also found a clear association with vegetation, in terms of the height of the vegetation layer (though further research using fuel types would be useful), but it was outperformed by fuel moisture metrics and meteorological drought proxies (dFMC and DC). Topography also played a role, with most ignitions occurring at high elevations. Increased probability of ignition has been consistently found across the coastal and hinterland mountain ranges (the Spanish Plateau, Sierra da Estrela or the Iberian Mountains, among others). The altitude gradient in the Iberian Peninsula is related to the distribution and type of vegetation communities. But the arrangement of the relief, which runs parallel to the coast (east–west) in the Mediterranean and the Plateau, has been recognised to modulate ignition by influencing air humidity through Foehn effects (Rodrigues *et al.* 2019*b*) and is also linked to storm fronts and thunderstorm episodes (Soler *et al.* 2021).

Our findings were non-conclusive in terms of the role of the characteristics of the fire-causing flash. We did not find a strong signal in terms of intensity or polarity, though negative-current flashes seemed to be more strongly tied to fire ignition. Previous models in central Spain support increased ignition likelihood linked to the mean peak current of negative flashes (Vecín-Arias *et al.* 2016), while some studies conducted in Catalonia (northeast Spain) found no specific evidence of the role played by polarity (Pineda *et al.* 2014)

Conclusions

In this work, we produced a daily forecast of the probability of natural wildfires in the Iberian Peninsula. We analysed a large dataset of lightning strikes, consisting of more than 17 million observations retrieved from a comprehensive ground-based network of sensors. We combined lightning and wildfire data with environmental variables to train and test a binary Random Forest model.

We identified the critical thresholds promoting lightning-related ignitions across the Iberian Peninsula. Lightning fires tend to occur when dead fuels are below 10–13% moisture content, with moderate drought conditions (DC > 300). We observed a large temporal variability linked to these drivers, but also a structural spatial pattern related to topography and vegetation structure. Natural fires concentrate in

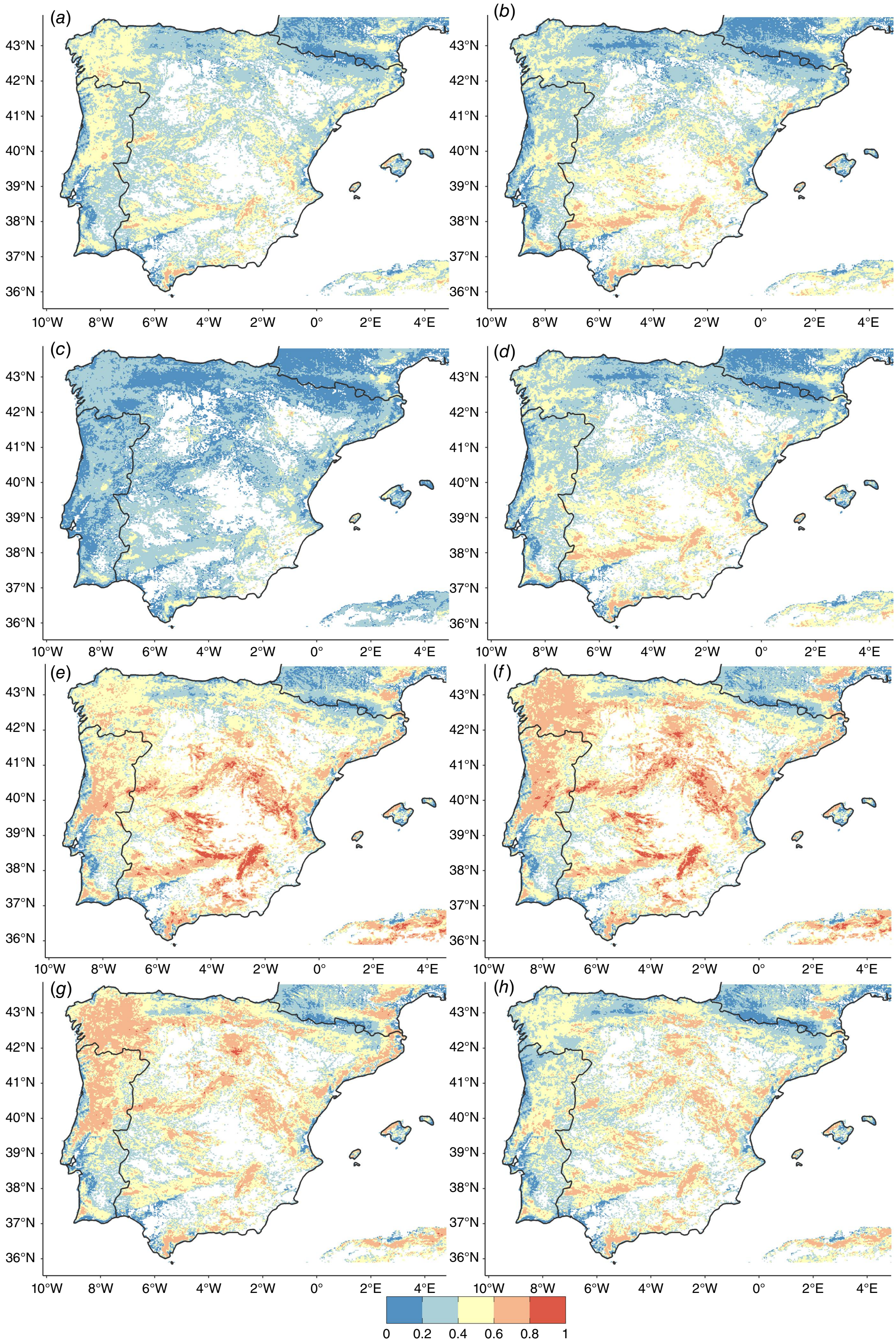

Fig. 4. Spatial distribution of predicted probability of lightning ignitions in the Iberian Peninsula. Panels *a* and *b* display the annual average in 2013 (*a*) and 2012 (*b*). Panels *c–h* show the monthly averages in the period 2009–2015 from May (*c*) to October (*h*).

mountain ranges above 1000 m with sufficient fuel load (vegetation height between 10 and 15 m tall). Our results revealed the major role played by fuel moisture content, the factor most sensitive to climate warming in the coming decades. This highlights the need for further research and consideration of natural ignitions in hazard mitigation plans, even though their contribution to fire activity is currently moderate.

References

Amatulli G, Peréz-Cabello F, de la Riva J (2007) Mapping lightning/human-caused wildfires occurrence under ignition point location uncertainty. *Ecological Modelling* **200**, 321–333. doi:10.1016/j.ecolmodel.2006.08.001

Anderson K (2002) A model to predict lightning-caused fire occurrences. *International Journal of Wildland Fire* **11**, 163–172. doi:10.1071/WF02001

Bar Massada A, Syphard AD, Stewart SI, Radeloff VC (2013) Wildfire ignition-distribution modelling: a comparative study in the Huron–Manistee National Forest, Michigan, USA. *International Journal of Wildland Fire* **22**, 174–183. doi:10.1071/WF11178

Baranovskiy NV, Kirienko VA (2022) Forest Fuel drying, pyrolysis and ignition processes during forest fire: a review. *Processes* **10**, 89. doi:10.3390/pr10010089

Barros AMG, Day MA, Preisler HK, Abatzoglou JT, Krawchuk MA, Houtman R, Ager AA (2021) Contrasting the role of human- and lightning-caused wildfires on future fire regimes on a central Oregon landscape. *Environmental Research Letters* **16**, 064081. doi:10.1088/1748-9326/ac03da

Boer MM, Nolan RH, Resco De Dios V, Clarke H, Price OF, Bradstock RA (2017) Changing weather extremes call for early warning of potential for catastrophic fire. *Earth's Future* **5**, 1196–1202. doi:10.1002/2017EF000657

Bowman DMJS, Balch JK, Artaxo P, Bond WJ, Carlson JM, Cochrane MA, D'Antonio CM, DeFries RS, Doyle JC, Harrison SP, Johnston FH, Keeley JE, Krawchuk MA, Kull CA, Marston JB, Moritz MA, Prentice IC, Roos CI, Scott AC, Swetnam TW, van der Werf GR, Pyne SJ (2009) Fire in the Earth System. *Science* **324**, 481–484. doi:10.1126/science.1163886

Bradley AP (1997) The Use of the Area Under the ROC Curve in the Evaluation of Machine Learning Algorithms. *Pattern Recognition* **30**, 1145–1159. doi:10.1016/S0031-3203(96)00142-2

Breiman L (2001) Random forests. *Machine Learning* **45**, 5–32. doi:10.1023/A:1010933404324

Camia A, Durrant T, San-Miguel-Ayanz J (2013) 'Harmonized classification scheme of fire causes in the EU adopted for the European Fire Database of EFFIS.' (Publications Office of the European Union: Luxembourg) Available at https://publications.jrc.ec.europa.eu/repository/handle/JRC80682

Castedo-Dorado F, Rodríguez-Pérez JR, Marcos-Menéndez JL, Álvarez-Taboada MF (2011) Modelling the probability of lightning-induced forest fire occurrence in the province of León (NW Spain). *Forest Systems* **20**, 95–107. doi:10.5424/fs/2011201-9409

Cattau ME, Wessman C, Mahood A, Balch JK (2020) Anthropogenic and lightning-started fires are becoming larger and more frequent over a longer season length in the USA. *Global Ecology and Biogeography* **29**, 668–681. doi:10.1111/geb.13058

Chuvieco E, Aguado I, Jurdao S, Pettinari ML, Yebra M, Salas J, Hantson S, de la Riva J, Ibarra P, Rodrigues M, Echeverría M, Azqueta D, Román MV, Bastarrika A, Martínez S, Recondo C, Zapico E, Martínez-Vega FJ (2014) Integrating geospatial information into fire risk assessment. *International Journal of Wildland Fire* **23**, 606–619. doi:10.1071/WF12052

Cochrane MA, Bowman DMJS (2021) Manage fire regimes, not fires. *Nature Geoscience* **14**, 455–457. doi:10.1038/s41561-021-00791-4

Coogan SCP, Cai X, Jain P, Flannigan MD (2020) Seasonality and trends in human- and lightning-caused wildfires ≥2 ha in Canada, 1959–2018. *International Journal of Wildland Fire* **29**, 473–485. doi:10.1071/WF19129

Couto FT, Iakunin M, Salgado R, Pinto P, Viegas T, Pinty JP (2020) Lightning modelling for the research of forest fire ignition in Portugal. *Atmospheric Research* **242**, 104993. doi:10.1016/j.atmosres.2020.104993

Dijkstra J, Durrant T, San-Miguel-Ayanz J, Veraverbeke S (2022) Anthropogenic and lightning fire incidence and burned area in Europe. *Land* **11**, 651. doi:10.3390/land11050651

Dowdy AJ, Mills GA (2012) Characteristics of lightning-attributed wildland fires in south-east Australia. *International Journal of Wildland Fire* **21**, 521–524. doi:10.1071/WF10145

Dubayah R, Blair JB, Goetz S, Fatoyinbo L, Hansen M, Healey S, Hofton M, Hurtt G, Kellner J, Luthcke S, Armston J, Tang H, Duncanson L, Hancock S, Jantz P, Marselis S, Patterson PL, Qi W, Silva C (2020) The global ecosystem dynamics investigation: high-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing* **1**, 100002. doi:10.1016/j.srs.2020.100002

Dupuy J-l, Fargeon H, Martin-StPaul N, Pimont F, Ruffault J, Guijarro M, Hernando C, Madrigal J, Fernandes P (2020) Climate change impact on future wildfire danger and activity in southern Europe: a review. *Annals of Forest Science* **77**, 35. doi:10.1007/s13595-020-00933-5

Fernandes P, Guiomar N, Davim DA (2022) Modelling the Behavior and Extent of Mid-Holocene Lightning-Caused Fires in Portugal *Environmental Sciences Proceedings* **17**, 89. doi:10.3390/environsciproc2022017089

Fernandes PM, Santos JA, Castedo-Dorado F, Almeida R (2021) Fire from the sky in the anthropocene. *Fire* **4**, 13. doi:10.3390/fire4010013

Ganteaume A, Camia A, Jappiot M, San-Miguel-Ayanz J, Long-Fournel M, Lampin C (2013) A review of the main driving factors of forest fire ignition over Europe. *Environmental Management* **51**, 651–662. doi:10.1007/s00267-012-9961-z

Greenwell BM (2017) pdp: An R package for constructing partial dependence plots. *The R Journal* **9**, 421–436. doi:10.32614/RJ-2017-016

Hu T, Zhou G (2014) Drivers of lightning- and human-caused fire regimes in the Great Xing'an Mountains. *Forest Ecology and Management* **329**, 49–58. doi:10.1016/j.foreco.2014.05.047

Jiménez-Ruano A, Jolly WM, Freeborn PH, Vega-Nieva DJ, Monjarás-Vega NA, Briones-Herrera CI, Rodrigues M (2022) Spatial predictions of human and natural-caused wildfire likelihood across Montana (USA). *Forests* **13**, 1200. doi:10.3390/f13081200

Kuhn M (2008) Building Predictive Models in R Using the caret Package. *Journal of Statistical Software* **28**, 1–26. doi:10.18637/jss.v028.i05

Li Y, Mickley LJ, Liu P, Kaplan JO (2020) Trends and spatial shifts in lightning fires and smoke concentrations in response to 21st century climate over the national forests and parks of the western United States. *Atmospheric Chemistry and Physics* **20**, 8827–8838. doi:10.5194/acp-20-8827-2020

MAAyMA (2015) 'Estadística general de incendios forestales.' (Ministerio de Agricultura, Alimentación y Medio Ambiente, Centro de Coordinación de la Información Nacional sobre Incendios Forestales: Madrid) [In Spanish]

Moreira F, Ascoli D, Safford H, Adams MA, Moreno JM, Pereira JMC, Catry FX, Armesto J, Bond W, González ME, Curt T, Koutsias N, McCaw L, Price O, Pausas JG, Rigolot E, Stephens S, Tavsanoglu C, Vallejo VR, Van Wilgen BW, Xanthopoulos G, Fernandes PM (2020) Wildfire management in Mediterranean-type regions: paradigm change needed. *Environmental Research Letters* **15**, 011001. doi:10.1088/1748-9326/ab541e

Moris JV, Conedera M, Nisi L, Bernardi M, Cesti G, Pezzatti GB (2020) Lightning-caused fires in the Alps: Identifying the igniting strokes. *Agricultural and Forest Meteorology* **290**, 107990. doi:10.1016/j.agrformet.2020.107990

Muñoz Sabater J (2019) ERA5-Land hourly data from 2001 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Available at doi:10.24381/CDS.E2161BAC

Nampak H, Love P, Fox-Hughes P, Watson C, Aryal J, Harris RMB (2021) Characterizing spatial and temporal variability of lightning activity associated with wildfire over Tasmania, Australia. *Fire* **4**, 10. doi:10.3390/fire4010010

NASA JPL (2020) NASADEM Merged DEM Global 1 arc second V001. Available at doi:10.5067/MEASURES/NASADEM/NASADEM_HGT.001

Nieto H, Aguado I, García M, Chuvieco E (2012) Lightning-caused fires in central Spain: Development of a probability model of occurrence for two Spanish regions. *Agricultural and Forest Meteorology* **162–163**, 35–43. doi:10.1016/j.agrformet.2012.04.002

Nolan RH, Resco de Dios V, Boer MM, Caccamo G, Goulden ML, Bradstock RA (2016) Predicting dead fine fuel moisture at regional scales using vapour pressure deficit from MODIS and gridded weather data. *Remote Sensing of Environment* **174**, 100–108. doi:10.1016/j.rse.2015.12.010

Pereira MG, Malamud BD, Trigo RM, Alves PI (2011) The history and characteristics of the 1980–2005 Portuguese rural fire database. *Natural Hazards and Earth System Sciences* **11**, 3343–3358. doi:10.5194/nhess-11-3343-2011

Pérez-Invernón FJ, Huntrieser H, Soler S, Gordillo-Vázquez FJ, Pineda N, Navarro-González J, Reglero V, Montanyà J, Van Der Velde O, Koutsias N (2021) Lightning-ignited wildfires and long continuing current lightning in the Mediterranean Basin: Preferential meteorological conditions. *Atmospheric Chemistry and Physics* **21**, 17529–17557. doi:10.5194/acp-21-17529-2021

Pineda N, Rigo T (2017) The rainfall factor in lightning-ignited wildfires in Catalonia. *Agricultural and Forest Meteorology* **239**, 249–263. doi:10.1016/j.agrformet.2017.03.016

Pineda N, Montanyà J, van der Velde OA (2014) Characteristics of lightning related to wildfire ignitions in Catalonia. *Atmospheric Research* **135–136**, 380–387. doi:10.1016/j.atmosres.2012.07.011

Pineda N, Altube P, Alcasena FJ, Casellas E, Segundo HS, Montanyà J (2022) Characterising the holdover phase of lightning-ignited wildfires in Catalonia. *Agricultural and Forest Meteorology* **324**, 109111. doi:10.1016/j.agrformet.2022.109111

Price C, Rind D (1994*a*) Possible implications of global climate change on global lightning distributions and frequencies. *Journal of Geophysical Research* **99**, 10823–10831. doi:10.1029/94JD00019

Price C, Rind D (1994*b*) Modeling Global Lightning Distributions in a General Circulation Model. *Monthly Weather Review* **122**, 1930–1939. doi:10.1175/1520-0493(1994)122<1930:MGLDIA>2.0.CO;2

R Core Team (2021) R: A Language and Environment for Statistical Computing. (R Foundation for Statistical Computing: Vienna, Austria) Available at https://www.R-project.org/

Resco de Dios V, Hedo J, Cunill Camprubí À, Thapa P, Martínez del Castillo E, Martínez de Aragón J, Bonet JA, Balaguer-Romano R, Díaz-Sierra R, Yebra M, Boer MM (2021) Climate change induced declines in fuel moisture may turn currently fire-free Pyrenean mountain forests into fire-prone ecosystems. *Science of the Total Environment* **797**, 149104. doi:10.1016/j.scitotenv.2021.149104

Resco de Dios V, Cunill Camprubí À, Pérez-Zanón N, Peña JC, Martínez del Castillo E, Rodrigues M, Yao Y, Yebra M, Vega-García C, Boer MM (2022) Convergence in critical fuel moisture and fire weather thresholds associated with fire activity in the pyroregions of Mediterranean Europe. *Science of the Total Environment* **806**, 151462. doi:10.1016/j.scitotenv.2021.151462

Rodrigues M, de la Riva J (2014) An insight into machine-learning algorithms to model human-caused wildfire occurrence. *Environmental Modelling & Software* **57**, 192–201. doi:10.1016/j.envsoft.2014.03.003

Rodrigues M, Alcasena F, Vega-García C (2019*a*) Modeling initial attack success of wildfire suppression in Catalonia, Spain. *Science of The Total Environment* **666**, 915–927. doi:10.1016/j.scitotenv.2019.02.323

Rodrigues M, González-Hidalgo JC, Peña-Angulo D, Jiménez-Ruano A (2019*b*) Identifying wildfire-prone atmospheric circulation weather types on mainland Spain. *Agricultural and Forest Meteorology* **264**, 92–103. doi:10.1016/j.agrformet.2018.10.005

Rodrigues M, Mariani M, Russo A, Salis M, Galizia LF, Cardil A (2021) Spatio-temporal domains of wildfire-prone teleconnection patterns in the western Mediterranean Basin. *Geophysical Research Letters* **48**, e2021GL094238. doi:10.1029/2021GL094238

Rodrigues M, Zúñiga-Antón M, Alcasena F, Gelabert P, Vega-Garcia C (2022) Integrating geospatial wildfire models to delineate landscape management zones and inform decision-making in Mediterranean areas. *Safety Science* **147**, 105616. doi:10.1016/j.ssci.2021.105616

Rodríguez-Pérez JR, Ordóñez C, Roca-Pardiñas J, Vecín-Arias D, Castedo-Dorado F (2020) Evaluating lightning-caused fire occurrence using spatial Generalized Additive Models: A case study in Central Spain. *Risk Analysis* **40**, 1418–1437. doi:10.1111/risa.13488

San-Miguel-Ayanz J, Schulte E, Schmuck G, Camia A, Strobl P, Liberta G, Giovando C, Boca R, Sedano F, Kempeneers P, McInerney D, Withmore C, Santos de Oliveira S, Rodrigues M, Durrant T, Corti P, Oehler F, Vilar L, Amatulli G (2012) Comprehensive monitoring of wildfires in Europe: the European Forest Fire Information System (EFFIS). In 'Approaches to Managing Disaster - Assessing Hazards, Emergencies and Disaster Impacts'. (Ed. J Tiefenbacher) pp. 87–105. (InTech)

Soler A, Pineda N, San Segundo H, Bech J, Montanyà J (2021) Characterisation of thunderstorms that caused lightning-ignited wildfires. *International Journal of Wildland Fire* **30**, 954–970. doi:10.1071/WF21076

Su Z, Hu H, Wang G, Ma Y, Yang X, Guo F (2018) Using GIS and Random Forests to identify fire drivers in a forest city, Yichun, China. *Geomatics Natural Hazards & Risk* **9**, 1207–1229. doi:10.1080/19475705.2018.1505667

Turco M, Rosa-Cánovas JJ, Bedia J, Jerez S, Montávez JP, Llasat MC, Provenzale A (2018) Exacerbated fires in Mediterranean Europe due to anthropogenic warming projected with non-stationary climate-fire models. *Nature Communications* **9**, 3821. doi:10.1038/s41467-018-06358-z

Van Wagner CE (1987) Development and structure of the Canadian Forest Fire Weather Index System. Forestry Technical Report 35. (Canadian Forestry Service, Headquarters: Ottawa)

Vazquez A, Moreno JM (1998) Patterns of lightning-, and people-caused fires in peninsular Spain. *International Journal of Wildland Fire* **8**, 103–115. doi:10.1071/WF9980103

Vecín-Arias D, Castedo-Dorado F, Ordóñez C, Rodríguez-Pérez JR (2016) Biophysical and lightning characteristics drive lightning-induced fire occurrence in the central plateau of the Iberian Peninsula. *Agricultural and Forest Meteorology* **225**, 36–47. doi:10.1016/j.agrformet.2016.05.003

Viegas DX, Piñol J, Viegas MT, Ogaya R (2001) Estimating live fine fuels moisture content using meteorologically-based indices. *International Journal of Wildland Fire* **10**, 223–240. doi:10.1071/WF01022

Wagner CEV, Pickett TL (1985) Equations and FORTRAN program for the Canadian Forest Fire Weather Index System. 33. (Canadian Forest Service) doi:citeulike-article-id:14026112

Wang J-F, Zhang T-L, Fu B-J (2016) A measure of spatial stratified heterogeneity. *Ecological Indicators* **67**, 250–256. doi:10.1016/j.ecolind.2016.02.052

Weiss A (2001) Topographic position and landforms analysis. In '21st Annual ESRI International User Conference', San Diego, CA. (Ecoregional Data Management Team The Nature Conservancy, Northwest Division 217 Pine St. Suite 1100 Seattle WA 98103)

Woodard SC, Rosenthal Y, Miller KG, Wright JD, Chiu BK, Lawrence KT (2014) Antarctic role in northern hemisphere glaciation. *Science* **346**, 847–851. doi:10.1126/science.1255586

Wotton BM, Martell DL (2005) A lightning fire occurrence model for Ontario. *Canadian Journal of Forest Research* **35**, 1389–1401. doi:10.1139/x05-071

Wotton BM, Nock CA, Flannigan MD (2010) Forest fire occurrence and climate change in Canada. *International Journal of Wildland Fire* **19**, 253–271. doi:10.1071/WF09002

Data availability. Data about lightning strikes are managed and distributed by the Spanish Meteorological Agency (AEMET), and were acquired by METEOGRID SL to conduct this work. Data about lightning-related ignitions were provided by the Spanish and Portuguese agencies. Spanish data come from the EGIF database (Estadística General de Incendios Forestales; https://www.miteco.gob.es/es/biodiversidad/servicios/banco-datos-naturaleza/informacion-disponible/incendios-forestales.aspx), available on request to the Spanish Ministry for Ecological Transition and Demographic Challenge while Portuguese fire data can be accessed on the ICNF website (Instituto da Conservação da Natureza e das Florestas; http://www2.icnf.pt/portal/florestas/dfci/inc/estat-sgif). We retrieved elevation data from the NASADEM global digital elevation model (NASA JPL 2020). We used the global forest canopy height map by Dubayah *et al.* (2020).

Conflicts of interest. The authors declare no conflicts of interest.

Declaration of funding. This work was financed by the projects FirEUrisk – Developing a Holistic, Risk-wise Strategy for European Wildfire Management, which received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101003890; and FIREPATHS (PID2020-116556RA-I00), funded by the Spanish Ministry of Science and Innovation.

Acknowledgements. The authors would like to thank the anonymous reviewers whose suggestions and comments have helped to improve this work.

Author affiliations

[A]Department of Geography and Land Management, University of Zaragoza, Pedro Cerbuna 12, 5009, Zaragoza, Spain.

[B]GEOFOREST Group, University Institute for Research in Environmental Sciences of Aragon (IUCA), University of Zaragoza, Pedro Cerbuna 12, 5009, Zaragoza, Spain.

[C]Department of Agriculture and Forest Engineering, University of Lleida, Alcalde Rovira Roure 191, 25198, Lleida, Spain.

[D]Department of Crop and Forest Sciences, University of Lleida, Alcalde Rovira Roure 191, 25198, Lleida, Spain.

[E]Joint Research Unit CTFC-AGROTECNIO-CERCA Center, Alcalde Rovira Roure 191, 25198, Lleida, Spain.

[F]School of Life Science and Engineering, Southwest University of Science and Technology, 59 Qinlong Road, 621010, Mianyang, China.

[G]MeteoGRID SL, Calle de Almansa 88, 28040, Madrid, Spain.