Electric Power Systems Research 231 (2024) 110296

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Survival functions of holdover time of lightning-ignited wildfires

Jose V. Moris [a,*], Davide Ascoli [a], Hugh G.P. Hunt [b]

[a] Department of Agricultural, Forest and Food Sciences, University of Turin, Grugliasco, Italy
[b] Johannesburg Lightning Research Laboratory, School of Electrical and Information Engineering, University of Witwatersrand, Johannesburg, South Africa

ARTICLE INFO

Keywords:
Lightning
Wildfire
Holdover time
Probability distribution fitting
Gamma distribution
Survival analysis

ABSTRACT

Lightning-ignited wildfires (LIWs) can go unreported for hours, days or even weeks before being reported. This is due to the fact that some LIWs present an intermediate phase between ignition and detection characterized by a smoldering combustion. Holdover time is generally defined as the time between lightning-induced ignition and fire detection. This study aims at obtaining survival functions to estimate the probability of a LIW reaching a certain holdover time. To this end, we fitted nine different probability distributions (exponential, chi-squared, log-normal, log-logistic, F, gamma, Weibull, Pareto, and Gompertz) to data from a database gathering 42 frequency distributions of holdover time obtained from more than 152,375 LIWs in 13 countries from 1921 to 2020. Gamma distributions provide the best fits to the observed holdover times. Accordingly, we estimated several survival functions derived from gamma distributions fitted to holdover time data. The survival functions are monotonically decreasing functions characterized by high probabilities for short holdover times and low probabilities for long holdover times. These survival functions can be used for holdover times of LIWs occurring globally as well as in boreal, Mediterranean and temperate coniferous forest biomes. Survival functions are intended to provide a more reliable way to assess holdover time-based probabilities of lightning causing wildfires.

1. Introduction

Lightning fire research is not a new topic [1], but recently it is receiving more attention [2,3]. Cloud-to-ground (CG) lightning discharges are the most frequent ignition source of natural wildfires worldwide [4,5]. In the absence of humans, vegetation fires have been occurring in our planet since over 400 million years ago, nearly all presumably ignited by lightning [6,7]. Lightning protection has historically focused on buildings and other structures (e.g., distribution lines, wind turbines), as well as human injuries and casualties [8,9]. Although systems for protecting individual trees from lightning exist [10], limiting the number of lightning-ignited wildfires (LIWs) is not feasible due to the vast territories where LIWs can start. Consequently, the management of LIWs is usually part of a strategic wildfire planning [11], although in some regions (e.g., dominated by boreal forests) LIWs are recognized as a fundamental ecosystem process [12].

LIWs are an important component of the fire regime (i.e., the spatial and temporal patterns of fires and their effects in a given area over a given time period [13]) in remote and mountainous areas along the globe. For example, in Canada LIWs account for approximately half of the wildfires but are responsible for more than 90 % of the burned area [14]. In the western United States, around 40 % of the wildfires and 70 % of the burned area are due to LIWs [15], while in south-east Australia LIWs account for around 30 % of the wildfires and 90 % of the burned area [16]. Moreover, LIWs can be a major component of wildfire crises in Australia [17], large wildfires in southern Europe [18], and extreme wildfire events in South Africa [19] to name a few examples.

LIWs present some particular characteristics in comparison with human-caused wildfires. First, in temperate and boreal latitudes LIWs mostly occur in summer months, when the majority of lightning events take place and weather conditions are more conducive for fire ignition and spread. While there may be a positive correlation between the number of lightning events and LIWs in boreal regions [2], weather is a strong driver of the occurrence and size of LIWs [20], which explains the high variability in LIW activity between years [21], and also within the same fire season [22]. In fact, LIWs are frequently associated with dry thunderstorms and dry lightning (i.e., lightning that occurs without a significant amount of precipitation) [15,23]. On the other hand, LIWs tend to occur, on average, in more remote locations, delaying and complicating the detection, access and suppression of these fires [24,

* Corresponding author.
E-mail address: moris.josev@gmail.com (J.V. Moris).

https://doi.org/10.1016/j.epsr.2024.110296
Received 21 July 2023; Received in revised form 1 December 2023; Accepted 1 March 2024

25]. In addition, LIWs seem to start more often in specific vegetation types, such as forests dominated by certain coniferous species [26,27]. All these factors increase the likelihood of LIWs becoming larger and more intense than anthropogenic wildfires [28]. The projected increase, within some regions, in the total amount of lightning [29] and lightning with Long-Continuing-Currents (LCC) [3], in combination with climate change [20], point to fire regimes with more and larger LIWs in the next decades, in which extreme fire seasons associated with LIWs may likely occur more frequently [22,30].

The holdover phenomenon is a particular characteristic of many LIWs. The physical process of a LIW is often divided into three distinct phases: ignition, survival, and arrival [31,32]. It is generally accepted that lightning with LCC is the main precursor of ignition of wildland fuels [33,34]. LIWs may have a survival phase between ignition and fire detection, also known as holdover or smoldering phase, characterized by the smoldering combustion (i.e., slow, low-temperature, flameless burning) of the soil organic layers [24,32]. Therefore, the lightning-induced ignition may spread almost immediately as a surface fire or survive as a smoldering fire in the organic soil. When the conditions become more favorable for fire spread (e.g., drier fuels), a transition from smoldering to flaming combustion takes place reaching the arrival phase, in which the fire is finally detected [26,32]. Since the existence and duration of the survival phase in LIWs is generally unknown, holdover time is usually defined as the time between lightning-induced fire ignition and fire detection [24,32]. Holdover times can be heterogeneous, ranging from a few minutes to several days and even weeks, although the majority of LIWs may be detected in less than 24 h, and holdover time data seem to follow right-skewed distributions (i.e., frequency tends to decrease non-linearly with time) [32].

Survival analysis refers to a set of statistical methods for analyzing the length of time until the occurrence of a certain event [35,36,37]. Survival analysis is frequently used to analyze censored data, i.e., when the exact times are not known. Morin et al. [38] used survival analysis to model the control time of forest fires (i.e., the time interval from the start of the initial attack to the time that the fire is declared as being under control). We are not aware of previous applications of survival analysis to holdover time data, although this kind of data seem well suited for such analysis [39]. In fact, survival analysis could help increase our limited knowledge on the initial behavior of LIWs given the lack of field observations available to study the ignition, survival and arrival phases. For example, survival analysis may become relevant to explore the drivers of holdover time [26], improve the predictions of models of LIW occurrence [24,31], and refine the methods to find the lightning responsible for LIWs [40].

In this paper, we extend the work presented by Moris et al. [39] that analyzed multiple holdover time datasets to find suitable statistical distributions for this phenomenon. Our goal is to obtain survival functions that allow us to estimate the probability of a LIW reaching a certain holdover time (i.e., the duration from lightning ignition to fire detection). To do so, we explore what probability distributions fit best the data included in a global holdover time database [32]. Thus, in this study we use a parametric survival approach. Furthermore, we explore other aspects to better understand the data from the holdover time database. The remainder of the paper is organized as follows: Section 2 briefly describes the holdover time database and the methods used to analyze the data. Section 3 presents the main results, including survival functions of holdover time. These results are discussed in Section 4, and Section 5 presents the main conclusions. Lastly, the Supplementary Materials include additional results (Appendix 1) as well as the holdover time data and scripts to reproduce the results presented in this paper within the R statistical environment (Appendix 2).

2. Materials and methods

2.1. Holdover time data

We used exclusively data download from a global database on holdover times of LIWs [41]. Here, we briefly introduce the database, but for a detailed description refer to Moris et al. [32]. The database contains three data files: censored data, non-censored data, and ancillary data. Censored data are the main component of the database and consist of 42 frequency distributions (i.e., 42 datasets) reporting the number or relative frequency of LIWs per interval of holdover time in different study areas and periods. Censored data do not include exact holdover times but the lower and upper bounds of the time intervals surrounding the holdover times. Thus, the censored data can be visualized as histograms (Fig. 1). The duration of the time intervals in which the frequency distributions are divided can vary substantially between datasets (from few minutes to several days; Fig. 1), and also within the same dataset. The censored data were collected from 29 different studies with observations from more than 152,375 LIWs covering 13 countries in five continents from 1921 to 2020. Table S1 (Appendix 1) presents a summary of the 42 datasets included in the database.

On the other hand, the database contains nine non-censored datasets in which the exact estimated values of holdover time (e.g., in seconds) for each single LIW are known (i.e., no censoring). These nine datasets also appear as hourly frequency distributions within the censored data file. Finally, the ancillary data include diverse types of information for each dataset regarding data identification (e.g., id code), spatial features (e.g., country, biome), temporal features (e.g., start and end years), fires (e.g., number of LIWs), lightning (e.g., lightning location network), methodology to derive the holdover times (e.g., criteria used to select the most likely igniting lightning), and data entry (e.g., data collector). These ancillary data are fundamental to understand data diversity and differences between datasets. All data from the database are already harmonized to facilitate the analyses.

2.2. Survival analyses

Let the holdover time, T, be a continuous and non-negative random variable, $[0,\infty)$, that represents the period elapsing between lightning-induced ignition and fire detection. The survival function, $S(t)$, is defined as the complement of the cumulative distribution function (CDF), $F(t)$, or the integral of the probability density function (PDF), $f(t)$: $S(t) = P\ (T > t) = 1 - F(t) = \int_t^\infty f(x)\ dx$. In the context of holdover time, the survival function represents the probability that a LIW has not been detected during a specified period of time t. Survival functions are monotonically decreasing functions, $S(x) \le S(t)$ for all $x > t$, and probabilities range from one at the origin, $S(0) = 1$, to zero as time approaches infinity, $S(\infty) = 0$ (Fig. 2) [35,36].

In order to build survival functions from holdover time data, the first step was to fit different probability distributions to the available data from the holdover time database. First of all, we assigned a “reliability class” to each of the 42 censored datasets (Table S1). This was a grouping into three classes based on factors that may affect the fits, such as the number of LIWs or the number of records (i.e., number of time intervals in which a censored dataset is divided) [39]. The lowest reliability (class 1) was assigned to datasets with (i) $<$ 150 LIWs, (ii) or $<$ 4 records, (iii) or those with a frequency of LIWs in the second day or hour similar or higher than the frequency observed in the first day or hour, respectively (13 datasets). On the contrary, the highest reliability (class 3) was assigned to datasets in which the holdover times were estimated using lightning data from lightning location networks (20 studies), while class 2 was allocated to the remaining ones (9 datasets).

In this part of the analysis, we selected nine candidate probability distributions (exponential, chi-squared, log-normal, log-logistic, *F*,

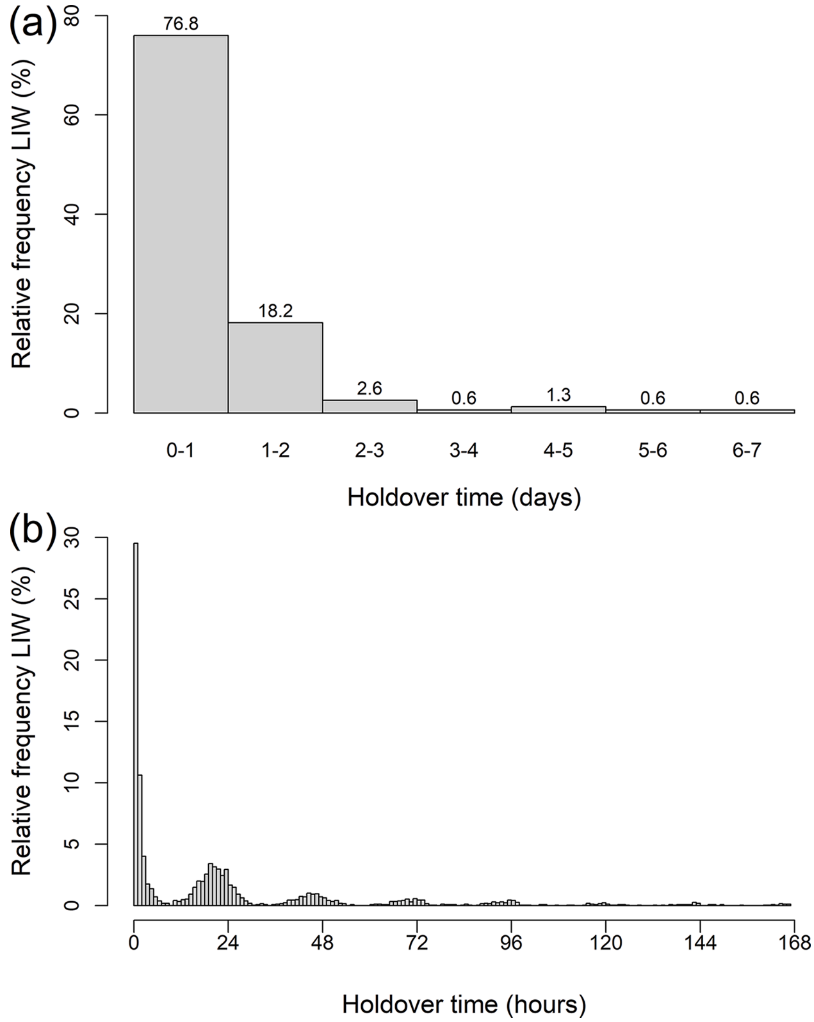

Fig. 1. Examples of frequency distributions (censored datasets) of holdover time from the database. (a) Daily frequency distribution obtained from 154 LIWs in Ticino (Switzerland) between 1981 and 2000. (b) Hourly frequency distribution obtained from 2693 LIWs in Florida (United States) between 2009 and 2013.

gamma, Weibull, Pareto, and Gompertz) among common ones used in survival analysis [42]. These distributions satisfy the following requirements: (i) are positive continuous distributions, (ii) have one or two parameters, and (iii) allow positive skewness and long tails. We applied the function "fitdistcens" from the R package "fitdistrplus" to fit, by maximum likelihood estimation, the nine probability distributions to each of the 42 censored datasets [43]. We used Akaike and Bayesian Information Criteria (AIC and BIC), as metrics measuring the agreement between the empirical datasets and the theoretical distributions, to select the best fitted parametric distribution for each dataset [35,42,43]. In addition, we generated goodness-of-fit CDF plots to compare graphically the fits of the nine candidate distributions by applying the function "cdfcompcens" from the same package "fitdistrplus".

Given the good fit of gamma distributions to holdover time data (Table S1), we used this distribution in the rest of the subsequent analyses. First, we explored whether a three-parameter distribution, such as the generalized gamma, was a better option to estimate survival functions of holdover time than the 2-parameter gamma distribution. To do so, we fitted both distributions to each of the 42 censored datasets and inspected the differences. Second, we used the nine non-censored datasets to explore whether the length of time intervals in which censored data are divided could influence the fits. To that end, we compared gamma distributions fitted to continuous values of holdover time (i.e., non-censored data) against gamma distributions fitted to hourly and daily censored data. We then examined differences in plots and statistics, such as the median holdover time and the cumulative probability of day one (CDF d1), which indicates the proportion of LIWs with holdover times shorter than 24 h.

Additionally, we examined the existence of daily local peaks of holdover time. With that purpose, we initially selected non-censored datasets with $>$ 2000 LIWs to avoid potential problems derived from probability density estimation with small sample sizes. We then applied, to each selected non-censored dataset, the "density" function (default parameters) from the base "stats" R package to compute kernel density estimates from the continuous values of holdover time [44]. Next, the function "peaks" (default parameters) from the R package "splus2R" was applied to the kernel density estimates in order to detect local peaks of holdover time [45].

Finally, we built several "consolidated" survival functions of holdover time by exploring four different approaches that combine single censored datasets and gamma distributions (Table 1) [39]. Each approach can be summarized as follows:

- In the first consolidation approach (approach 1), we pooled all the censored data from the 42 datasets into a single dataset, and a gamma distribution was fitted to this first combined dataset of holdover time censored data. Thus, in the first approach censored datasets with more LIWs have more weight in the combined dataset.
- On the other hand, in the second approach we used relative frequencies (instead of number of LIWs) to provide the same number of

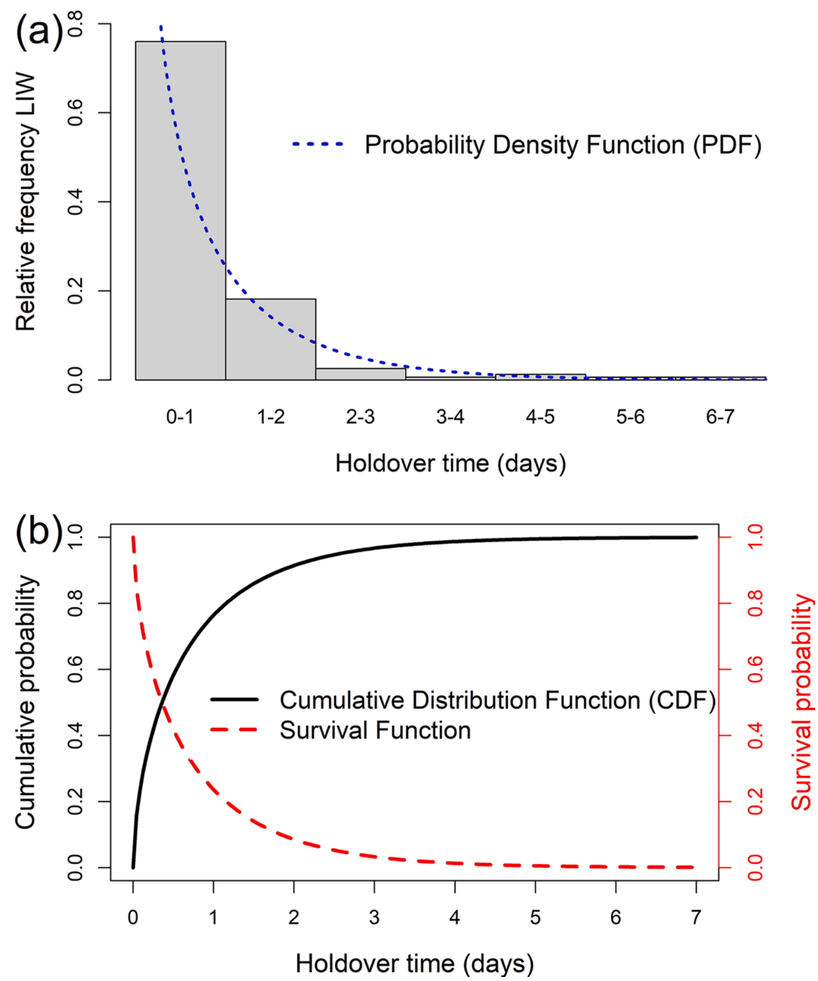

Fig. 2. Example of a probability distribution fitted to a censored dataset of holdover time obtained from 154 LIWs in Ticino (Switzerland) between 1981 and 2000. (a) Probability density function. (b) Cumulative distribution and survival functions.

Table 1
Four approaches used to combine holdover time data.

Approach	Data pooled	Holdover time	Weight criteria	Size consolidated dataset
1	42 observed censored datasets	Censored	Number LIWs	152,475
2	42 observed censored datasets	Censored	Equal	420,000
3	42 theoretical gamma distributions	Non-censored	Equal	4,200,000
4	42 theoretical gamma distributions	Non-censored	Reliability class	9,100,000

observations (10,000) from each single censored dataset to the second equally-weighted combined dataset. We then fitted a gamma distribution to this second combined dataset.

- The third approach was based on generating 100,000 random values of holdover time from each of the 42 gamma distributions fitted to the 42 single censored datasets (Table S3). Consequently, we created a third combined dataset made of 4,200,000 theoretical continuous values of holdover time. A gamma distribution was fitted to the third combined dataset.
- The fourth approach was similar, but the number of random values per gamma distribution was derived from the reliability class assigned to the datasets (300,000, 200,000 and 100,000 random values for reliability classes 3, 2 and 1 respectively). Therefore, each single gamma distribution was not equally weighted in the fourth approach. In total, the fourth combined dataset was made of 9,100,000 theoretical continuous values of holdover time.

In summary, in the first and second approaches gamma distributions were fitted to pooled observed censored data on holdover time, while in the third and fourth approaches we pooled theoretical non-censored data on holdover time in order to combine the 42 gamma distributions derived from single datasets. These four approaches were repeated by dividing the holdover time database into three different biomes: boreal forests (13 datasets), temperate coniferous forests (19 datasets), and Mediterranean forests (seven datasets), to obtain additional specific survival functions for these biomes. All the analyses were carried out using the R software environment [44]. Furthermore, all the holdover time data as well as the R scripts to perform the analysis can be found in Appendix 2.

3. Results

The gamma distribution has lower AIC values more often than the rest of distributions (Table 2). In fact, the gamma distribution is selected as the best fit more frequently with increasing reliability class (Table 2). Table S1 shows the best distribution fitted to each single censored dataset of holdover time. Analogous results are found by applying BIC (Table S2). Thus, overall the gamma distribution yields the best fits to

Table 2
Number and percent (in parenthesis) of distributions selected as the best fit according to AIC.

Distribution	Reliability 1-2-3	Reliability 2-3	Reliability 3
Gamma	20 (47.6 %)	17 (58.6 %)	14 (70.0 %)
Weibull	4 (9.5 %)	3 (10.3 %)	2 (10.0 %)
Gompertz	4 (9.5 %)	0	0
Chi-squared	3 (7.1 %)	2 (6.9 %)	2 (10.0 %)
Exponential	3 (7.1 %)	0	0
Log-logistic	3 (7.1 %)	3 (10.3 %)	0
Log-normal	2 (4.8 %)	1 (3.4 %)	0
Pareto	2 (4.8 %)	2 (6.9 %)	2 (10.0 %)
F	1 (2.4 %)	1 (3.4 %)	0
TOTAL	42	29	20

holdover time data, although graphically gamma distributions can be similar to other distributions such as Weibull (Fig. 3). The gamma distributions fitted to single censored datasets are characterized by PDFs with right skewness, mean holdover times > medians, modes = 0, and both shape and rate parameters < 1 (Table S3), with the exception of 8 out of the 42 gamma distributions, that present shape parameters > 1 (i.e., mode > 0; Fig. S1).

The generalized gamma distribution, when compared with the gamma distribution, is selected more often as the best fit (in 74 % and 60 % of the 42 censored datasets according to the minimum AIC and BIC, respectively). Nonetheless, goodness-of-fit CDF plots show that both distributions offer comparable fits for many datasets, which derive in similar survival functions (Table S3; Fig. 4a). For example, the mean absolute difference between observed data and fitted distributions in terms of median holdover time is equal to 3.5 h and 3.6 h for the gamma and generalized gamma distributions, respectively (Table S3). On the other hand, we detected some issues with generalized gamma distributions. We often observed computing problems when fitting generalized gamma distributions to holdover time data, especially with maximum likelihood estimation. In some cases, we were not able to estimate the parameters of generalized gamma distributions, and for some datasets we ended up with anomalous shapes in the survival distributions (e.g., Fig. 4b).

In general, distributions fitted to censored daily data show more differences from continuous data-based distributions than distributions fitted to hourly data (Table 3). Overall, we observed that the distributions tend to have shorter holdover times with increasing time interval, which derives in slightly lower median values and higher cumulative probabilities within the first day (Table 3), as well as slightly lower survival probabilities during the first hours and days (Fig. 5). This is more evident in gamma distributions with values of the shape parameter < 1 (i.e., PDFs with mode = 0). However, gamma distributions fitted to holdover time data from Brazil (Study id = MEN2022BR01) have values of the shape parameter > 1 (i.e., mode > 0), and in this particular case the tendency is the opposite: the distributions tend to have longer holdover times with increasing time interval from continuous to hourly and daily. Despite these minor differences, the fitted distributions may be considered reasonably similar, and therefore the time interval did not affect the general shape of the survival functions (e.g., Fig. 5).

We examined the presence of local peaks of holdover time in three datasets of non-censored data (one from Spain and two from the United States), given that only these three non-censored datasets contain > 2000 LIWs. The period between two consecutive local peaks (local maxima) in the smooth Kernel curves was within a range of 17–27 h, and on average locals peaks occurred approximately every 23–24 h in the three datasets (Table 4). The size of the local peaks tended to decrease with increasing holdover time (Fig. 6).

The four approaches of data consolidation result in gamma distributions with similar values of means, medians and cumulative probabilities of holdover time (Table 5). Likewise, the survival functions derived from these approaches are graphically analogous (Fig. 7). On the contrary, we found clear differences between biomes. Gamma distributions from boreal data present longer holdover times (Table 5) and survival probabilities (Fig. 8) than distributions derived from temperate and Mediterranean biomes.

4. Discussion

4.1. Gamma distributions

Our results indicate that gamma distributions are suitable to estimate survival functions of holdover time. Gamma distributions fit well to the data (Fig. 3), and are selected as the best fits to single censored datasets of holdover time more frequently than the rest of the eight distributions considered in this study, especially for datasets that, we suspect, may represent more faithfully the real frequency distributions of holdover time (Table 2). The large variability among censored datasets results in clear differences between individual fitted gamma distributions (Table S3). As expected, the PDFs of gamma distributions fitted to single censored datasets show right skewness and a decreasing frequency of LIWs with increasing holdover time (Fig. S1). Only eight PDFs were not strictly decreasing (i.e., mode > 0). Certain issues with the original

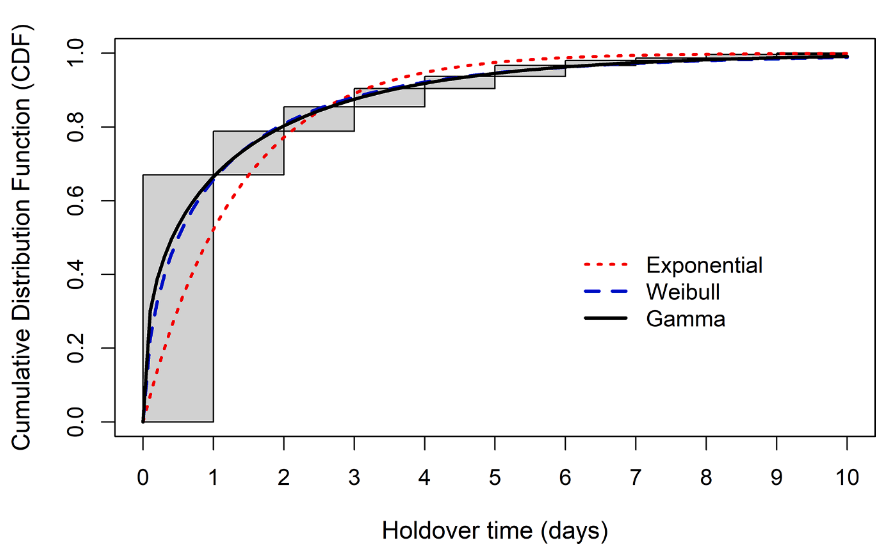

Fig. 3. Example of the empirical cumulative distribution of a censored dataset of holdover time obtained from 303 LIWs in Austria between 2013 and 2020 (grey rectangles) plotted against the CDFs (lines) of three fitted distributions.

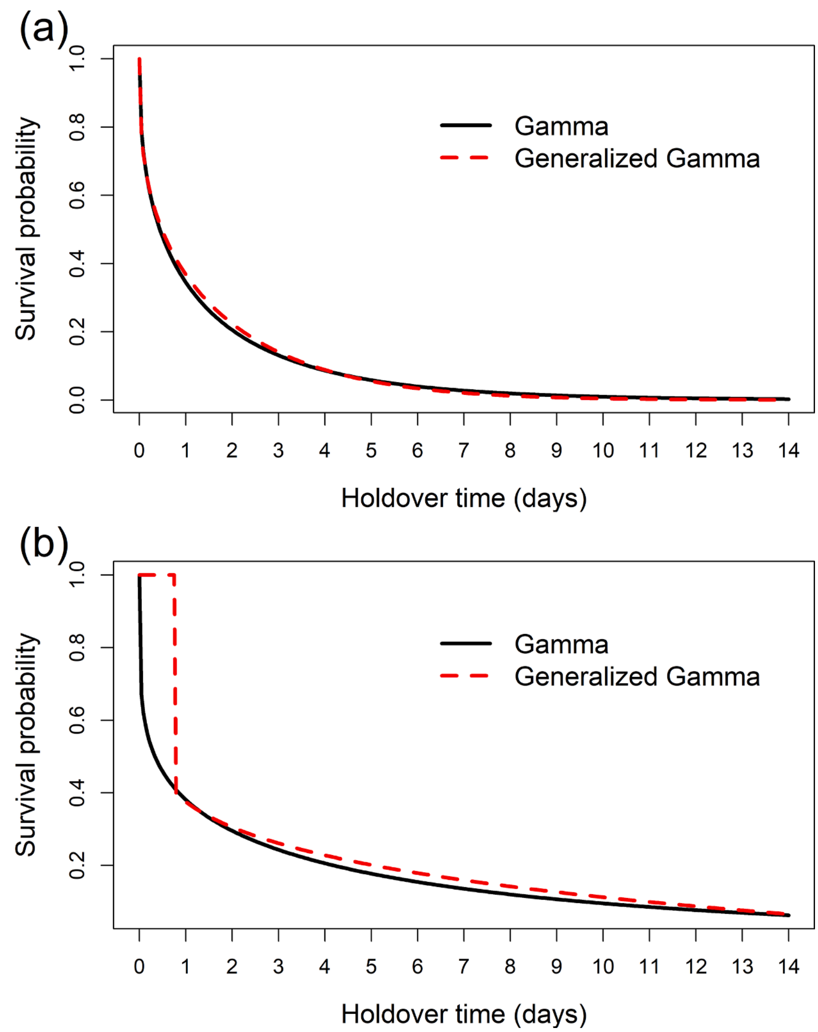

Fig. 4. Examples of survival functions derived from gamma and generalized gamma distributions fitted to two different censored datasets of holdover time. (a) Dataset obtained from 263 LIWs in Switzerland between 2001 and 2018. (b) Dataset obtained from 16,940 LIWs in southern and central British Columbia (Canada) between 2000 and 2020.

Table 3
Gamma distributions fitted to continuous, hourly and daily holdover time data.

Study id	Median (*h*) Continuous	Hourly	Daily	CDF d1 (%) Continuous	Hourly	Daily
MOR2020IT01	13.6	12.8	6.3	64.3	65.1	71.6
MOR2020CH01	12.7	10.6	13.2	63.5	65.5	63.0
PER2021FR01	12.0	7.9	1.9	62.4	66.4	74.3
PER2021PT01	12.2	11.5	17.0	67.9	68.5	63.1
PER2021ES01	6.5	4.7	2.8	69.3	69.7	72.4
MEN2022BR01	18.0	18.2	20.4	63.1	62.8	61.3
PER2022US01	9.8	8.7	6.2	71.5	72.4	75.8
PER2022US02	7.6	6.1	7.1	70.5	73.8	74.2
PIN2022ES01	4.1	2.0	1.5	81.3	82.7	84.5

CDF d1: cumulative probability of day one. See Table S1 for a short description of the nine datasets that appear in this table.

censored datasets may help explain this deviation. For example, five datasets used satellite images as a source of fire data, which could likely delay the discovery time of LIWs, overestimating holdover durations in comparison with other datasets [32]. The remaining three datasets belong to the lowest reliability class for reasons such as a low number of LIWs and few time intervals. In conclusion, the gamma distribution shows some advantages. It is simple and versatile; with only two parameters, the gamma distribution is flexible enough to change its shape greatly. Accordingly, gamma distributions provide a good fit to right-skewed holdover time data, and the estimation of parameters is easy to compute generally.

Despite generalized gamma distributions may fit slightly better to holdover time data (because a third parameter provides more flexibility), we preferred gamma distributions for several reasons. The survival functions from both distributions tend to be so similar, that differences in survival probabilities may be irrelevant for most practical applications (Fig. 4a). On the other hand, we found several disadvantages when fitting generalized gamma distributions. First, we needed an

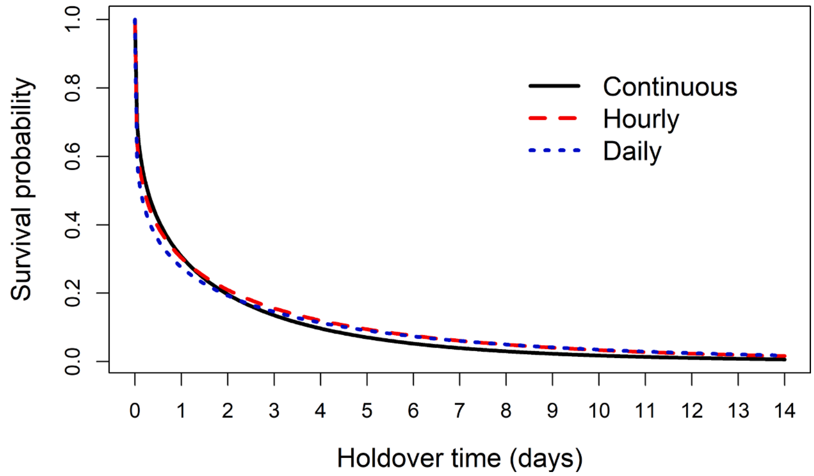

Fig. 5. Example of survival functions derived from gamma distributions fitted to censored (daily and hourly) and non-censored (continuous) holdover time data obtained from 2702 LIWs in Spain between 2009 and 2015.

Table 4
Times (in hours) at which local peaks of holdover time were detected in three non-censored datasets.

Local peak	PER2021ES01	PER2022US01	PER2022US02
1	0.9	1.6	1.1
2	18.4	19.2	20.5
3	45.7	45.4	45.3
4	70.9	69.4	70.1
5	93.3	94.1	94.2
6	118.5	117.8	118.6
7	143.0	142.9	142.3
8	168.2	164.4	164.2

See Table S1 for a short description of the three datasets that appear in this table.

additional R package to load the generalized gamma distribution [46]. More importantly, we detected computational problems [42]. For example, we were not able to fit generalized gamma distribution to some datasets (with maximum likelihood estimation and other fitting methods), and survival functions for other censored datasets showed anomalous, non-smooth shapes (Fig. 4b). In summary, we found that gamma distributions were easier to fit and overall the results were equivalent to those provided by generalized gamma distributions.

4.2. Censoring and daily peaks

We do not know the true holdover times. Holdover time is usually estimated by using the time of the specific lightning event selected as the most likely ignition source [26,40]. The values from the holdover time database were collected from different studies that used different methods and no uncertainty measures were included [32]. Holdover time uncertainty is complicated by the spatial inaccuracy of lightning location networks and the holdover phenomenon itself, that may result in a large number of lightning events (e.g., hundreds) classified as potential igniting strikes for a single LIW [40,47]. In addition, different methodologies produce different values of holdover time for some LIWs [40,48]. In this study, we explored time uncertainty by comparing gamma distributions fitted to different interval-censored data (hourly and daily) containing the estimated holdover times (non-censored values), keeping in mind that true holdover times are unknown.

We found that the length of the time interval used to censor data seems to have a limited influence on the fit of gamma distributions (Table 3). The use of non-censored data (i.e., continuous values of holdover times) results in small differences between survival functions (Fig. 5). This is relevant to us because the current version of the holdover time database relies mainly on daily and hourly censored data [32]. Thus, our results indicate that even survival functions derived from daily holdover time data should be reasonably suitable for many applications.

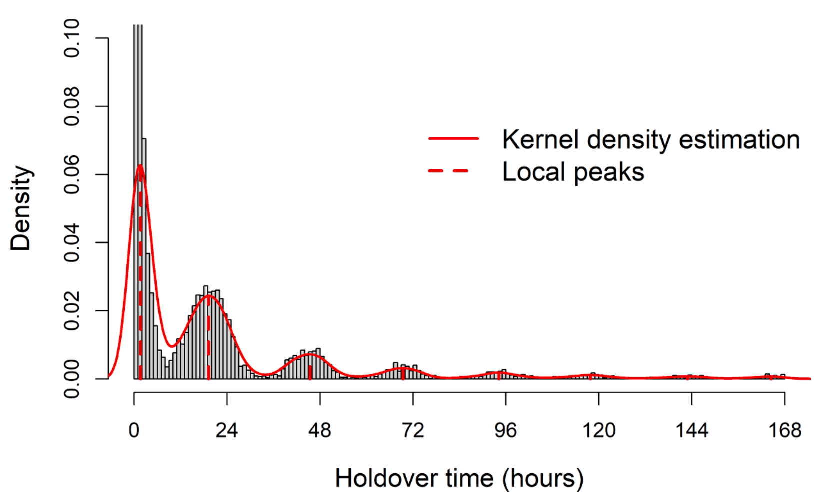

Fig. 6. Example of histogram of a non-censored dataset obtained from 6301 LIWs in Arizona and New Mexico (United States) between 2009 and 2013 with the Kernel density curve and the position of the local peaks of holdover time.

Table 5
Fitted gamma distributions according to four different approaches of data consolidation.

Datasets	Approach	Gamma Shape	Rate	Holdover time Mean (*h*)	Median (*h*)	CDF d1 (%)
All (42)	1	0.277	0.175	38.0	8.1	66.0
	2	0.369	0.235	37.6	12.4	62.0
	3	0.362	0.225	38.7	12.4	61.8
	4	0.318	0.190	40.1	10.7	63.1
Boreal (13)	1	0.486	0.185	63.0	27.9	46.9
	2	0.650	0.268	58.2	32.3	42.6
	3	0.558	0.231	58.1	28.9	45.8
	4	0.463	0.176	63.0	26.7	47.9
Temperate (19)	1	0.267	0.214	29.9	5.9	70.3
	2	0.352	0.281	30.1	9.3	66.9
	3	0.370	0.305	29.1	9.6	67.0
	4	0.341	0.282	29.0	8.6	68.0
Mediterranean (7)	1	0.278	0.236	28.3	6.1	70.7
	2	0.317	0.319	23.8	6.3	72.4
	3	0.304	0.307	23.8	5.9	72.7
	4	0.286	0.327	21.0	4.7	75.4

CDF d1: cumulative probability of day one. The parameters shape and rate were estimated using holdover times in days.

Shorter time intervals and exact times should be preferred to reduce data uncertainty though.

Temporal and spatial differences in holdover time appear to follow the expected effects of weather on fire behavior at different scales. For example, Pineda et al. [26] show that lightning-induced ignitions occurring during the early afternoon tend to have shorter holdover times than the rest of ignitions; ignitions at the beginning and end of the natural fire season tend to have longer holdover times; and holdover time increases with altitude. Here, by applying Kernel density estimation, we detected daily peaks of holdover time in three non-censored datasets (Table 4). The daily peaks represent local maxima in the frequency of holdover times separated by approximately 24 h (Fig. 6) [32, 49]. These daily peaks are probably associated with two processes. First, the diurnal heating and nocturnal cooling cycle that occurs daily favors the spread and detection of LIWs in the late afternoon-early evening. [50,51]. Second, summer thunderstorms and lightning tend to occur more frequently in the late afternoon hours [52,53]. The daily local peaks may thus be explained by the confluence of hourly distributions of both lightning-caused ignition and fire discovery, which seem to follow bell shapes with maximum values only separated by a few hours [26, 54]. This means that afternoon-evening ignitions that smolder for one or more nights may also tend to be reported during the afternoon-evening hours, creating daily peaks of frequency.

The relative importance of daily peaks decreases with time (Fig. 6). It is more likely to find larger local peaks within the first 24 and 48 h because only a small proportion of LIWs present long holdover times. We were surprised by the good temporal adjustment of local peaks between datasets (Table 4) since each dataset belongs to a different region, and

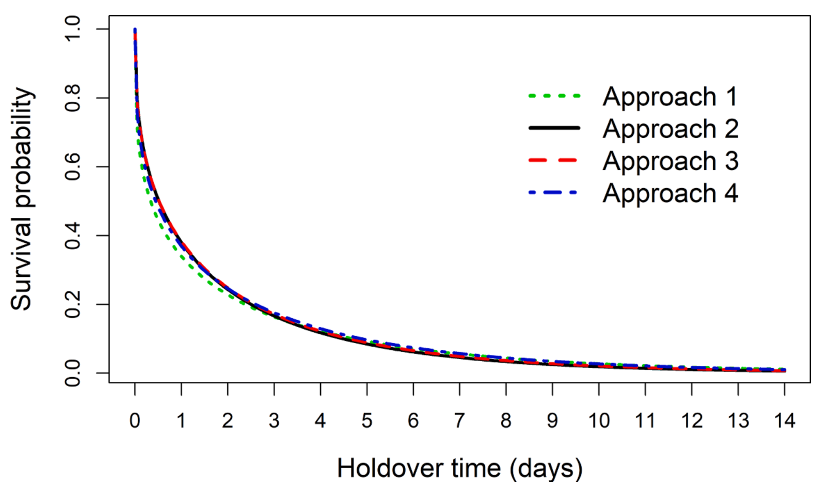

Fig. 7. Survival functions of holdover time derived from four different approaches of data consolidation.

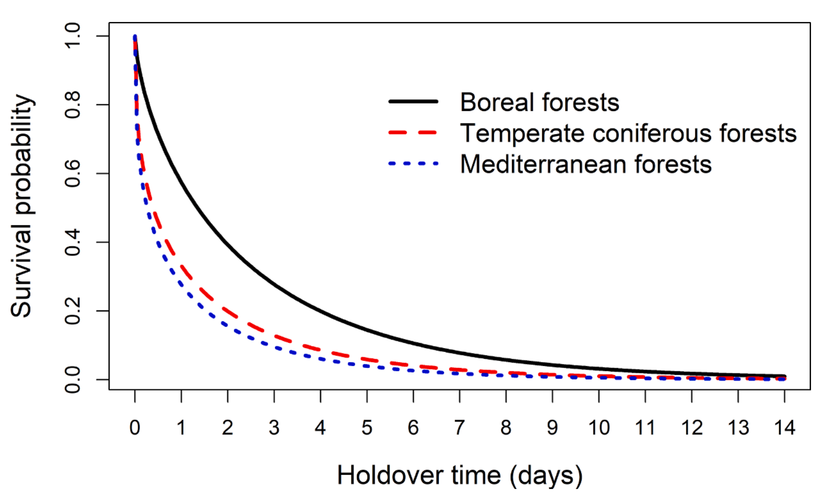

Fig. 8. Survival functions of holdover time for three biomes derived from the second approach of data consolidation.

the original lightning and fire datasets from the United States and Spain are independent. Nonetheless, caution is advised. The three analyzed datasets come from the same authors, and furthermore, we do not expect that the presence of daily peaks of holdover time are so evident in other regions. From a theoretical point of view, daily peaks should not have an implication for the survival functions, which reflect the general negative trend over time instead of cyclical components such as daily peaks.

4.3. Survival functions

Our aim was to estimate survival functions of holdover time. These functions are meant to provide the probability of a LIW reaching a certain holdover time. We used parametric models because we are interested in functions that can be applied easily (i.e., the functions represent smooth curves and are completely defined by a few known parameters). All the survival functions are based on gamma distributions. Because the different censored datasets from the holdover time database come from different sources [32], we tried four approaches of data consolidation before fitting the gamma distributions [55]. These approaches combine censored datasets or probability distributions, equally or unequally weighted. Our results show that survival functions derived from the four approaches produce similar probabilities for the same values of holdover time (Fig. 7). Consequently, the choice of a function may not have practical implications in most cases. Although we cannot recommend any specific function in favor of another, approaches 2 (combination of datasets equally) and 3 (combination of distributions equally) generated the most analogous functions. Conversely, we found strong differences between survival functions built with data from different biomes (Fig. 8). We only estimated survival functions for three biomes (boreal, Mediterranean, and temperate coniferous forests) because the database did not include enough data from other biomes. Gamma distributions from boreal forests yield longer holdover times than temperate and Mediterranean forests (Table 5), which results in higher survival probabilities for the same holdover time (Fig. 8) reflecting the role of factors such as climate and vegetation in the initial phases of LIWs. For other biomes, we recommend applying the survival functions built with all datasets as a reasonable approximation (Table 5).

The survival functions presented in this study are characterized by: (i) high probability values for short holdover times (i.e., LIWs with short holdover times occur much more frequently), and (ii) low probability values for long holdover times (i.e., long holdover times are rare in LIWs) [32]. For example, according to the survival function based on the second approach for temperate coniferous forests, holdover times of 3 h and 10 days yield probabilities of 65.8 and 1.0 %, respectively. On the other hand, studies on LIWs usually assume that holdover times cannot exceed arbitrary durations [27,40,50,51]. Survival functions do not require specifying a maximum holdover time. Instead, survival functions assume that long holdover times are unlikely but not impossible. Finally, the probabilities obtained from survival functions should be interpreted as relative (and not absolute) probabilities. Therefore, the probabilities should be compared within the same function in order to, for example, rank the likelihood of observing certain holdover times. For instance, we cannot expect that every year 1 % of LIWs in temperate coniferous forests have holdover times $>$ 10 days. As shown in Table S3, each study area and period presents a different distribution of holdover time driven by spatial (e.g., climate, vegetation, soil) and temporal (e.g., fuel moisture, fire weather) factors [26,31,32,40,51].

Why is it important to model the holdover time? The holdover phenomenon hinders our capacity to verify the ignition time of LIWs [26,40]. The lack of proper field observations forces us to model holdover times to better understand the different phases of LIWs [31]. Parametric (e.g., gamma) survival analyses may be a suitable alternative to identify and quantify the main drivers of holdover time (in general) and smoldering combustion (in particular) in different regions [38]. Even if other methods are used to model holdover times, such as generalized linear models, gamma regressions remain a reasonable alternative beforehand. Another important application is the match between lightning and fire data. Due to data uncertainties and other factors, identifying the single CG lightning event responsible for a wildfire (and so the exact moment of ignition) is currently a difficult task [40]. Many studies on LIWs rely on a proper selection of lightning igniting wildfires as a fundamental initial step to carry out the rest of analyses [26,27,40,50,51]. The spatio-temporal proximity index developed by Larjavaara et al. [56] is one of the most common methods used to distinguish lightning discharges that cause wildfires [32]. This method assumes that the temporal component of the proximity index decreases linearly with time (Fig. 9). The survival functions presented here could be used as data-driven alternatives to provide more appropriate probabilities of holdover time. In combination with more sophisticated spatial-based methods [57], the application of survival functions of holdover time could result (or not) in the selection of different lightning events as the most likely ignition sources for some wildfires, helping us to refine one of the most challenging methodological aspects in the study of LIWs.

5. Conclusions

The holdover phenomenon is a particular characteristic of LIWs that complicates the study of these wildfires. Using a global database on holdover time (i.e., time between lightning-induced ignition and fire detection) collected from more than 152,375 LIWs [41], this study shows that gamma distributions fit well to holdover time data and are selected as the best fits more frequently than the rest of the probability distributions considered. In fact, our results suggest that more complex distributions, such as generalized gamma distributions, may not be necessary to model holdover times. Moreover, interval censoring of holdover time data (e.g., hourly and daily) seems to have a limited influence on gamma distribution fits. We also show that the presence of daily peaks of holdover time (i.e., local maxima in the frequency of holdover times separated by approximately 24 h) may be a general pattern of LIWs in different regions. This study presents several survival functions of holdover time based on gamma distributions. These survival functions are meant to provide relative probabilities of observing holdover times in LIWs occurring globally as well as in boreal, Mediterranean and temperate coniferous forest biomes. Survival functions of holdover time may provide a more reliable way to assess time-based probabilities of lightning causing wildfires, helping to reduce the uncertainty about when and under what conditions fire ignition, smoldering and arrival phases take place. Accordingly, future research could explore the applications of survival functions to select igniting lightning and model drivers of holdover time. Lastly, we would like to encourage the interdisciplinary collaboration between lightning and wildfire researchers. More collaborations of this kind are needed to keep advancing our understanding of lightning-ignited fires, especially under global change and in regions in which natural fires are a major component of the fire regime and extreme wildfire events.

CRediT authorship contribution statement

Jose V. Moris: Conceptualization, Data curation, Funding acquisition, Methodology, Software, Validation, Visualization, Writing – original draft. **Davide Ascoli:** Conceptualization, Supervision, Writing – review & editing. **Hugh G.P. Hunt:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

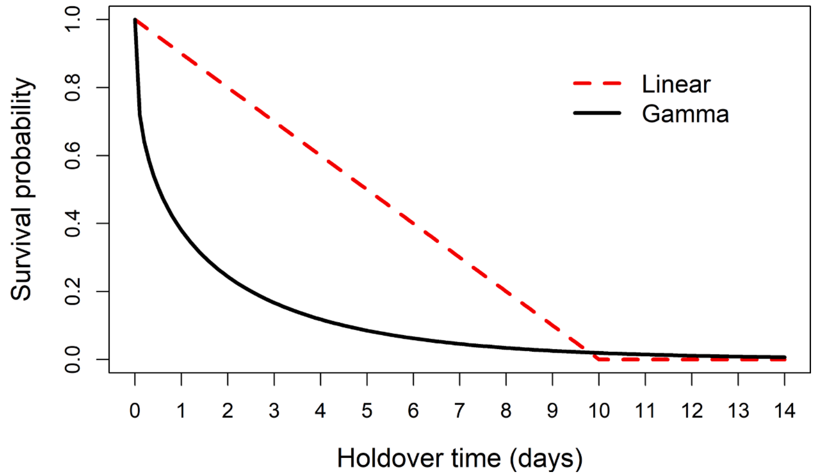

Fig. 9. A survival function according to a gamma distribution (from this study) and a linear approach applied to the spatio-temporal proximity index for matching fire and lightning data.

Data availability

Datasets related to this article can be found in Appendix 2 of the Supplementary Materials. The original data can be found at https://doi.org/10.5281/ZENODO.7352172 (version 1.0.0).

Acknowledgment

This work was supported by a postdoctoral fellowship funded by the government of Asturias (Spain) through FICYT (Fundación para el Fomento en Asturias de la Investigación Científica Aplicada y la Tecnología) [grant number AYUD/2021/58534].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.epsr.2024.110296.

References

[1] J.S. Barrows, Lightning fire research in the rocky mountains, J. Forest. 52 (1954) 845–847.

[2] S. Veraverbeke, B.M. Rogers, M.L. Goulden, R.R. Jandt, C.E. Miller, E.B. Wiggins, J. T. Randerson, Lightning as a major driver of recent large fire years in North American boreal forests, Nat. Clim. Change 7 (2017) 529–534, https://doi.org/10.1038/nclimate3329.

[3] F.J. Pérez-Invernón, F.J. Gordillo-Vázquez, H. Huntrieser, P. Jöckel, Variation of lightning-ignited wildfire patterns under climate change, Nat. Commun. 14 (2023) 739, https://doi.org/10.1038/s41467-023-36500-5.

[4] S.J. Pyne, P.L. Andrews, R.D. Laven, Introduction to Wildland Fire, 2nd ed., Wiley, New York, 1996.

[5] A.C. Scott, D.M.J.S. Bowman, W.J. Bond, S.J. Pyne, M.E. Alexander, Fire On Earth: An Introduction, 1st ed., Wiley, New Jersey, 2014.

[6] J.G. Pausas, J.E. Keeley, A burning story: the role of fire in the history of life, Bioscience 59 (2009) 593–601, https://doi.org/10.1525/bio.2009.59.7.10.

[7] J.E. Keeley, J.G. Pausas, Evolutionary ecology of fire, Annu. Rev. Ecol. Evol. Syst. 53 (2022) 203–225, https://doi.org/10.1146/annurev-ecolsys-102320-095612.

[8] V.A. Rakov, M.A. Uman, Lightning: Physics and Effects, 1st ed., Cambridge University Press, 2003 https://doi.org/10.1017/CBO9781107340886.

[9] G. Maslowski, V.A. Rakov, K. Michishita, Advances in lightning research and protection technologies, Electr. Power Syst. Res 181 (2020) 106206, https://doi.org/10.1016/j.epsr.2020.106206.

[10] B.C. Brusso, Dendrology and lightning protection, IEEE Ind. Appl. Mag. 29 (92) (2023) 6–10, https://doi.org/10.1109/MIAS.2022.3216741.

[11] P.M. Fernandes, G.M. Delogu, V. Leone, D. Ascoli, Wildfire policies contribution to foster extreme wildfires, in: F. Tedim, V. Leone, T.K. McGee (Eds.), Extreme Wildfire Events and Disasters, Elsevier, Amsterdam, 2020, pp. 187–200, https://doi.org/10.1016/B978-0-12-815721-3.00010-2.

[12] C. Tymstra, B.J. Stocks, X. Cai, M.D. Flannigan, Wildfire management in Canada: review, challenges and opportunities, Prog. Disaster Sci. 5 (2020) 100045, https://doi.org/10.1016/j.pdisas.2019.100045.

[13] F.J. Oddi, Fire Regime, in: S.L. Manzello (Ed.), Encyclopedia of Wildfires and Wildland-Urban Interface (WUI) Fires, Springer, Cham, 2018, pp. 1–12, https://doi.org/10.1007/978-3-319-51727-8_73-1.

[14] C.C. Hanes, X. Wang, P. Jain, M.-A. Parisien, J.M. Little, M.D. Flannigan, Fire-regime changes in Canada over the last half century, Can. J. For. Res. 49 (2019) 256–269, https://doi.org/10.1139/cjfr-2018-0293.

[15] J.T. Abatzoglou, C.A. Kolden, J.K. Balch, B.A. Bradley, Controls on interannual variability in lightning-caused fire activity in the western US, Environ. Res. Lett. 11 (2016) 045005, https://doi.org/10.1088/1748-9326/11/4/045005.

[16] A.J. Dowdy, G.A. Mills, Characteristics of lightning-attributed wildland fires in south-east Australia, Int. J. Wildland Fire 21 (2012) 521–524, https://doi.org/10.1071/WF10145.

[17] D. Owens, M. O'Kane, Final report of the NSW bushfire inquiry, Department of Premier and Cabinet (NSW), 2020.

[18] M. Rodrigues, À. Cunill Camprubí, R. Balaguer-Romano, C.J. Coco Megía, F. Castañares, J. Ruffault, P.M. Fernandes, V. Resco De Dios, Drivers and implications of the extreme 2022 wildfire season in Southwest Europe, Sci. Total Environ. 859 (2023) 160320, https://doi.org/10.1016/j.scitotenv.2022.160320.

[19] P.E. Frost, L.G. Kleyn, R, van den Dool, M. Burgess, L. Vhengani, K. Steenkamp, K. Wessels, The Elandskraal fire, Knysna: a data driven analysis, CSIR Report number 271960–1, 2018.

[20] T.D. Hessilt, J.T. Abatzoglou, Y. Chen, J.T. Randerson, R.C. Scholten, G. Van Der Werf, S. Veraverbeke, Future increases in lightning ignition efficiency and wildfire occurrence expected from drier fuels in boreal forest ecosystems of western North America, Environ. Res. Lett. 17 (2022) 054008, https://doi.org/10.1088/1748-9326/ac6311.

[21] J.V. Moris, M. Conedera, L. Nisi, G.B. Pezzatti, Blitzschlagbrände und Sommertrockenheit: gibt es einen Zusammenhang? Swiss Forest. J. 171 (2020) 281–287, https://doi.org/10.3188/szf.2020.0281.

[22] R.C. Scholten, D. Coumou, F. Luo, S. Veraverbeke, Early snowmelt and polar jet dynamics co-influence recent extreme Siberian fire seasons, Science 378 (2022) 1005–1009, https://doi.org/10.1126/science.abn4419.

[23] N. Pineda, O. Rodríguez, ERA5 reanalysis of environments conducive to lightning-ignited wildfires in Catalonia, Atmosphere (Basel) 14 (2023) 936, https://doi.org/10.3390/atmos14060936.

[24] B.M. Wotton, D.L. Martell, A lightning fire occurrence model for Ontario, Can. J. For. Res. 35 (2005) 1389–1401, https://doi.org/10.1139/x05-071.

[25] G. Cesti, M. Conedera, F. Spinedi, Considerazioni sugli incendi boschivi causati da fulmini, Swiss Forest. J. 156 (2005) 353–361, https://doi.org/10.3188/szf.2005.0353.

[26] N. Pineda, P. Altube, F.J. Alcasena, E. Casellas, H.S. Segundo, J. Montanyà, Characterising the holdover phase of lightning-ignited wildfires in Catalonia, Agr. For. Meteorol. 324 (2022) 109111, https://doi.org/10.1016/j.agrformet.2022.109111.

[27] F.J. Pérez-Invernón, H. Huntrieser, J.V. Moris, Meteorological conditions associated with lightning ignited fires and long-continuing-current lightning in Arizona, New Mexico and Florida, Fire 5 (2022) 96, https://doi.org/10.3390/fire5040096.

[28] M.E. Cattau, C. Wessman, A. Mahood, J.K. Balch, Anthropogenic and lightning-started fires are becoming larger and more frequent over a longer season length in the U.S.A, Global Ecol. Biogeogr. 29 (2020) 668–681, https://doi.org/10.1111/geb.13058.

[29] Y. Chen, D.M. Romps, J.T. Seeley, S. Veraverbeke, W.J. Riley, Z.A. Mekonnen, J. T. Randerson, Future increases in Arctic lightning and fire risk for permafrost carbon, Nat. Clim. Chang. 11 (2021) 404–410, https://doi.org/10.1038/s41558-021-01011-y.

[30] A. Descals, D.L.A. Gaveau, A. Verger, D. Sheil, D. Naito, J. Peñuelas, Unprecedented fire activity above the Arctic Circle linked to rising temperatures, Science 378 (2022) 532–537, https://doi.org/10.1126/science.abn9768.

[31] K. Anderson, A model to predict lightning-caused fire occurrences, Int. J. Wildland Fire 11 (2002) 163–172, https://doi.org/10.1071/WF02001.
[32] J.V. Moris, P. Álvarez-Álvarez, M. Conedera, A. Dorph, T.D. Hessilt, H.G.P. Hunt, R. Libonati, L.S. Menezes, M.M. Müller, F.J. Pérez-Invernón, G.B. Pezzatti, N. Pineda, R.C. Scholten, S. Veraverbeke, B.M. Wotton, D. Ascoli, A global database on holdover time of lightning-ignited wildfires, Earth Syst. Sci. Data 15 (2023) 1151–1163, https://doi.org/10.5194/essd-15-1151-2023.
[33] D.M. Fuquay, R.G. Baughman, A.R. Taylor, R.G. Hawe, Characteristics of seven lightning discharges that caused forest fires, J. Geophys. Res. 72 (1967) 6371–6373, https://doi.org/10.1029/JZ072i024p06371.
[34] D.M. Fuquay, A.R. Taylor, R.G. Hawe, C.W. Schmid, Lightning discharges that caused forest fires, J. Geophys. Res. 77 (1972) 2156–2158, https://doi.org/10.1029/JC077i012p02156.
[35] J.P. Klein, M.L. Moeschberger, Survival Analysis: Techniques for Censored and Truncated Data, Springer, New York, 2003, https://doi.org/10.1007/b97377.
[36] D.G. Kleinbaum, M. Klein, Survival Analysis: A Self-Learning Text, Springer, New York, 2012, https://doi.org/10.1007/978-1-4419-6646-9.
[37] P. Schober, T.R. Vetter, Survival analysis and interpretation of time-to-event data: the tortoise and the hare, Anesth. Analg. 127 (2018) 792–798, https://doi.org/10.1213/ANE.0000000000003653.
[38] A.A. Morin, A. Albert-Green, D.G. Woolford, D.L. Martell, The use of survival analysis methods to model the control time of forest fires in Ontario, Canada, Int. J. Wildland Fire 24 (2015) 964, https://doi.org/10.1071/WF14158.
[39] J.V. Moris, D. Ascoli, H.G. Hunt, Probability distributions for holdover time of lightning-caused wildfires, in: 36th International Conference on Lightning Protection (ICLP), IEEE, Cape Town, 2022, pp. 496–499, https://doi.org/10.1109/ICLP56858.2022.9942474.
[40] J.V. Moris, M. Conedera, L. Nisi, M. Bernardi, G. Cesti, G.B. Pezzatti, Lightning-caused fires in the Alps: identifying the igniting strokes, Agric. For. Meteorol. 290 (2020) 107990, https://doi.org/10.1016/j.agrformet.2020.107990.
[41] J.V. Moris, P. Álvarez-Álvarez, M. Conedera, A. Dorph, T.D. Hessilt, H.G.P. Hunt, R. Libonati, L.S. Menezes, M.M. Müller, F.J. Pérez-Invernón, G.B. Pezzatti, N. Pineda, R.C. Scholten, S. Veraverbeke, B.M. Wotton, D. Ascoli, Database on holdover time of lightning-ignited wildfires, Zenodo v1.0.0 (2022), https://doi.org/10.5281/ZENODO.7352172.
[42] N. Taketomi, K. Yamamoto, C. Chesneau, T. Emura, Parametric distributions for survival and reliability analyses, a review and historical sketch, Mathematics 10 (2022) 3907, https://doi.org/10.3390/math10203907.
[43] M.L. Delignette-Muller, C. Dutang, fitdistrplus : an R package for fitting distributions, J. Stat. Softw. 64 (2015) 4, https://doi.org/10.18637/jss.v064.i04.
[44] R. Core Team, R: a language and environment for statistical computing. R Foundation For Statistical Computing, Vienna, Austria, 2021. https://www.R-project.org/.
[45] W. Constantine, T. Hesterberg, splus2R: supplemental S-PLUS functionality in R, R package version 1.3-3, 2021. https://CRAN.R-project.org/package=splus2R.
[46] M.H.J. Saldanha, A.K. Suzuki, ggamma: generalized gamma probability distribution. R package version 1.0.1, 2019. https://CRAN.R-project.org/package=ggamma.
[47] W.J. Braun, J.E. Stafford, Multivariate density estimation for interval-censored data with application to a forest fire modelling problem, Environmetrics 27 (2016) 345–354, https://doi.org/10.1002/env.2396.
[48] C.J. Schultz, N.J. Nauslar, J.B. Wachter, C.R. Hain, J.R. Bell, Spatial, temporal and electrical characteristics of lightning in reported lightning-initiated wildfire events, Fire 2 (2019) 18, https://doi.org/10.3390/fire2020018.
[49] W.G. Morris, What is the time between ignition and discovery of lightning fires? For. Res. Notes 40 (1947).
[50] N. Pineda, J. Montanyà, O.A. Van Der Velde, Characteristics of lightning related to wildfire ignitions in Catalonia, Atmos. Res. 135–136 (2014) 380–387, https://doi.org/10.1016/j.atmosres.2012.07.011.
[51] N. Pineda, T. Rigo, The rainfall factor in lightning-ignited wildfires in Catalonia, Agric. For. Meteorol. 239 (2017) 249–263, https://doi.org/10.1016/j.agrformet.2017.03.016.
[52] D.R. Poelman, W. Schulz, G. Diendorfer, M. Bernardi, The European lightning location system EUCLID – part 2: observations, Nat. Hazard. Earth Syst. Sci. 16 (2016) 607–616, https://doi.org/10.5194/nhess-16-607-2016.
[53] T.L. Koehler, Cloud-to-ground lightning flash density and thunderstorm day distributions over the contiguous United States derived from NLDN measurements: 1993–2018, Mon. Weather Rev. 148 (2020) 313–332, https://doi.org/10.1175/MWR-D-19-0211.1.
[54] J.S. Barrows, Forest Fires in the Northern Rocky Mountains, USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, Missoula, 1951. Research Paper RM-28.
[55] T.P. Hill, J. Miller, How to combine independent data sets for the same quantity, Chaos 21 (2011) 033102, https://doi.org/10.1063/1.3593373.
[56] M. Larjavaara, J. Pennanen, T.J. Tuomi, Lightning that ignites forest fires in Finland, Agric. For. Meteorol. 132 (2005) 171–180, https://doi.org/10.1016/j.agrformet.2005.07.005.
[57] H.G.P. Hunt, K.J. Nixon, J.A. Naudé, Using lightning location system stroke reports to evaluate the probability that an area of interest was struck by lightning, Electr. Power Syst. Res. 153 (2017) 32–37, https://doi.org/10.1016/j.epsr.2016.12.010.