2022 36$^{\text{th}}$ International Conference on Lightning Protection (ICLP)

Cape Town, South Africa

2 - 7 October 2022

IEEE Catalog Number: CFP2209T-ART

ISBN: 978-1-6654-9024-5

2022 36th International Conference on Lightning Protection (ICLP) | 978-1-6654-9024-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICLP56858.2022.9942575

Probability distributions for holdover time of lightning-caused wildfires

Jose V Moris
Department of Agricultural, Forest and Food Sciences
University of Turin
Grugliasco, Italy
moris.josev@gmail.com

Davide Ascoli
Department of Agricultural, Forest and Food Sciences
University of Turin
Grugliasco, Italy
d.ascoli@unito.it

Hugh GP Hunt
The Johannesburg Lightning Research Laboratory, School of Electrical and Information Engineering
University of the Witwatersrand
Johannesburg, South Africa
hugh.hunt@wits.ac.za

***Abstract*—The holdover phase (i.e., the time between lightning-induced ignition and fire detection) is a phenomenon characteristic of lightning-caused wildfires. In this paper, we analyzed multiple holdover time datasets to determine the best statistical distribution for this phenomenon. We found that the gamma distribution seems a suitable candidate to model holdover times. We propose applications of the gamma distribution to improve the investigation of lightning-caused wildfires.**

Keywords—lightning, holdover fires, statistical modelling

I. Introduction

Some wildfires ignited by lightning present an initial latent phase characterized by smoldering (i.e., burning slowly with smoke but no flame) of the soil organic matter before turning into detectable surface fires [1]. Since the presence and duration of this latent (holdover) phase is usually unknown, holdover duration is commonly defined, for practical reasons, as the time between lightning-induced fire ignition and fire detection [2, 3]. The goal of this study is to find a suitable probability distribution that characterizes holdover times of lightning-caused wildfires. For that purpose, we address the following specific objectives:

- What probability distribution fits best empirical estimations of holdover time?
- How variable are the probability distributions of holdover fitted with data from different studies?
- Does size interval of frequency distributions obtained from empirical estimations of holdover times affect the probability distributions?

II. Background

Our current knowledge on the holdover phenomenon is still limited. Holdover durations can range from a few minutes [4], to occasionally some weeks and even months [5]. It is commonly accepted that the majority of lightning fires have short holdover times (less than 24 hours), and that holdover distributions present a right-skewed distribution with an exponential decay [2, 3, 6]. While some lightning fires may start propagating right after ignition (i.e., no smouldering) [7], other fires may extinguish during the latent phase before being reported [2, 3].

The holdover phenomenon is one of the major challenges to study lightning-caused wildfires [8, 9], and the reasons are manifold. First, holdover time is supposed to vary across regions. Lightning fires have only been studied in a few regions and it is uncertain how findings from a few studies may be valid in other parts of the globe [10]. Second, the drivers of holdover fires are not yet well understood [4]. The durations of holdover fires are rarely examined in detail, although recent studies focused on the holdover phenomenon and its causes [11]. Third, approaches to estimate the holdover duration are variable. Until lightning data from Lightning Location Systems (LLS) became widely available, holdover durations were estimated from the elapsed time between fire discovery and the most recent lightning storm over the area of the wildfire [12]. Nowadays, the methods to estimate holdover times are mainly based on identifying the most likely igniting lightning (from a lightning dataset) [6]. Several methodologies are used, although a spatio-temporal index of proximity has become the most common method to find igniting lightning, and consequently derive holdover times [4, 6, 11, 13]. However, current methods do not take into consideration the temporal distribution of holdover fires, and holdover time estimations can vary depending on the method and parameters applied [6].

III. Data

We used a global database of holdover times of lightning fires [10]. On May 1 2022, the database included 34 frequency distributions of holdover times (Table I). The holdover datasets contained 2051 records, collected from 26 publications, with more than 102,552 lightning fires distributed across 12 countries in four continents from 1921 to 2020. Each holdover frequency distribution contains several records with the number or proportion of lightning fires in a specific time interval. Therefore, the database contains interval-censored data exclusively. For example, the amount of fires with a holdover time shorter than 24 hours (i.e., the lightning fires detected during the first day), the number of fires with a holdover time between 24 and 48 hours, and so on (Fig. 1). The duration of the time intervals is variable among studies, although most of the holdover frequency distributions report the number of lightning fires in days or hours.

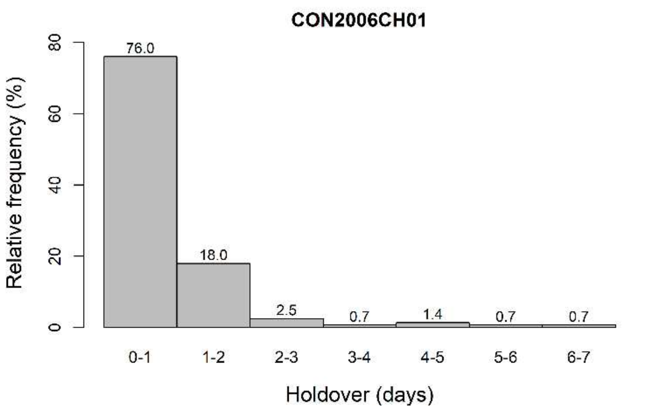

Fig. 1. Example of a frequency distribution from the global database of holdover times of lightning fires [10].

2022 36th International Conference on Lightning Protection (ICLP) | 978-1-6654-9024-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICLP56858.2022.9942474

TABLE I. DATASETS OF HOLDOVER TIME ANALYZED IN THIS PAPER.

Study id	Study area	Period	n fires	n records	Rel.
CON2006CH01	Ticino (CH)	1981-2004	156	7	2
CON2006IT01	Aosta Valley (IT)	2003-2003	29	6	1
KOU1967CA01	Canada	1960-1963	3615	16	2
NAS1996CA01	Alberta and Saskatchewan (CA)	1988-1993	2551	15	3
WOT2005CA01	Ontario (CA)	1992-2001	5169	28	3
SCH2019US01	United States	2012-2015	797	15	3
MAC2019US01	Western United States	2017	95	11	1
DOW2009AU01	Victoria (AU)	2000-2009	1797	4	3
PIN2014ES01	Catalonia (ES)	2004-2009	464	24	3
PIN2017ES01	Catalonia (ES)	2009-2014	357	19	3
GIS1926US01	Northern Rocky Mountains (US)	1924-1925	1933	11	2
GIS1931US01	Northern Rocky Mountains (US)	1924-1928	4149	11	2
BAR1951US01	Northern Rocky Mountains (US)	1931-1945	16368	13	2
TAY1969US01	Northern Rocky Mountains (US)	1950-1965	14489	4	2
SHO1923US01	California (US)	1921		6	1
SHO1930US01	California (US)	1921-1922	443	6	2
BAR1978US01	Arizona and New Mexico (US)	1960-1974	28377	8	2
MOR1948US01	Oregon and Washington (US)	1940-1944	5357	28	2
DUN2010US01	Florida (US)	1986-2003	230	2	1
LAR2005FI01	Finland	1996-2002	106	5	1
MUL2021AT01	Austria	2013-2020	303	10	3
MOR2020IT01	Aosta Valley (IT)	2012-2018	32	150	1
MOR2020CH01	Switzerland	2001-2018	263	238	3
PER2021GR01	Greece	2017-2019	914	95	1
PER2021FR01	Mediterranean France	2012-2015	36	242	1
PER2021PT01	Portugal	2009-2015	309	93	3
PER2021ES01	Spain	2009-2015	2702	336	3
HES2022US01	Alaska (US)	2001-2012	402	5	1
HES2022US02	Alaska (US)	2012-2018	287	5	1
HES2022CA01	Northwest Territories (CA)	2001-2018	550	5	1
MEN2022BR01	Pantanal (BR)	2012-2017	265	65	1
PER2022US01	Arizona and New Mexico (US)	2009-2013	6301	168	3
PER2022US02	Florida (US)	2009-2013	2693	167	3
PIN2022ES01	Catalonia (ES)	2003-2020	1013	233	3

n = number; Rel. = reliability class.

IV. METHODS

Before fitting probability distribution to the datasets, we assigned a "reliability class" to each empirical holdover distribution. The lowest reliability (class 1) was given to studies (i) with < 150 fires, (ii) or < 4 records, (iii) or those with a frequency of holdover fires in the second day similar or higher than the frequency in the first day (in total 12 studies; Table I). For the rest of the studies, the highest reliability (class 3) was given to those that estimated holdover times using lightning data from LLS (13 studies), and class 2 to the remaining ones (9 studies).

We explored the holdover time datasets to search for distributions that could fit the histograms. We selected eight candidate distributions (exponential, chi-squared, log-normal, log-logistic, F, gamma, Weibull, and Pareto) that satisfied the following requirements: (i) to have one or two parameters; (ii) to be a positive continuous distribution; and (iii) to have positive skewness and long tail. We used the function "fitdistcens" from the R package "fitdistrplus" to fit, by maximum likelihood, the eight probability distributions to each of the 34 datasets [14]. We applied Akaike and Bayesian Information Criteria (AIC and BIC) to select the best distribution among those fitted to each dataset, and goodness-of-fit plots to compare empirical data and fitted distributions (Fig. 2).

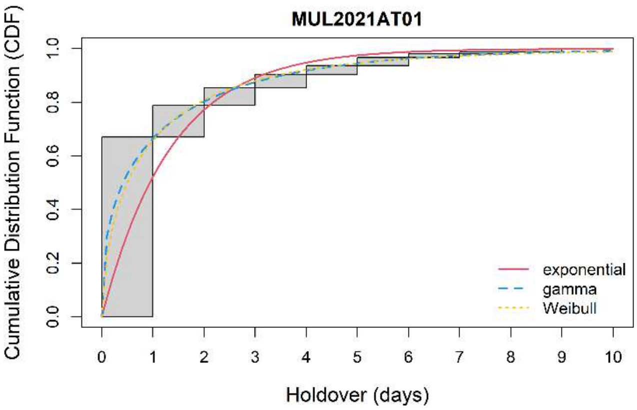

Fig. 2. Example of CDF of three distributions plotted against the empirical cumulative frequency (grey rectangles).

Given the good fit of gamma distributions, we decided to use this distribution in the rest of the analyses. First, we chose several variables, such as the two parameters that define the gamma distribution (shape and rate), to calculate basic statistics for the fitted gamma distributions. Second, we compared gamma distributions fitted with continuous holdover times (i.e., non-censored data) from nine different studies, against gamma distributions fitted with interval-censored data (i.e., hourly and daily histograms of holdover).

V. RESULTS

Overall, the gamma distribution had lower AIC and BIC values than the rest of the distributions, and the gamma distributions were selected as the best fit more frequently with increasing reliability of the holdover time datasets (Table II). In addition, the 34 fitted gamma distributions (Table III) showed a high variability in their parameters (shape and rate), mean and median holdover times, as well as the percent of fires detected within the first 24 hours after ignition (CDF day 1; Table IV).

We observed that the intervals used to censor the holdover data seem to influence the fit of the gamma distributions (Table V), although this influence is relatively limited on the probabilities (Fig. 3). In general, in gamma distributions with shape and rate < 1 (i.e., mode = 0), increasing the time interval seems to concentrate more mass of the distributions on the left of the figure, although this trend did not occur in all the nine datasets. On the other hand, in gamma distributions with shape and rate > 1 (i.e., mode > 0), the trend was the opposite (more mass on the right of the figure).

TABLE II. NUMBER AND PERCENT (IN PARENTHESES) OF DISTRIBUTIONS SELECTED AS THE BEST FIT ACCORDING TO AIC.

Distribution	Reliability 1-2-3	Reliability 2-3	Reliability 3
Gamma	14 (41.2)	12 (54.5)	9 (69.2)
Weibull	6 (17.6)	2 (9.1)	1 (7.7)
Log-logistic	3 (8.8)	3 (13.6)	0 (0.0)
Chi-squared	3 (8.8)	2 (9.1)	2 (15.4)
Exponential	3 (8.8)	0 (0.0)	0 (0.0)
Pareto	2 (5.9)	1 (4.5)	1 (7.7)
Log-normal	2 (5.9)	1 (4.5)	0 (0.0)
F	1 (2.9)	1 (4.5)	0 (0.0)
TOTAL	34 (100)	22 (100)	13 (100)

TABLE III. FITTED GAMMA DISTRIBUTIONS.

Study id	Shape	Rate	r	mean (h)	median (h)	CDF d 1 (%)
CON2006CH01	0.570	0.821	0.989	16.7	8.4	76.7
CON2006IT01	0.804	0.719	0.991	26.8	16.8	61.1
KOU1967CA01	0.489	0.361	0.994	32.5	14.5	61.3
NAS1996CA01	0.507	0.206	0.994	59.1	27.2	47.3
WOT2005CA01	0.631	0.179	0.996	84.5	46.0	35.2
SCH2019US01	0.272	0.193	0.998	33.8	7.0	68.1
MAC2019US01	1.032	0.345	0.992	71.8	50.4	27.7
DOW2009AU01	0.145	0.030	0.996	115.3	4.2	64.2
PIN2014ES01	0.341	0.971	0.984	8.4	2.5	89.7
PIN2017ES01	0.315	0.690	0.991	10.9	2.9	85.7
GIS1926US01	0.596	1.206	0.989	11.9	6.2	84.8
GIS1931US01	0.512	0.994	0.983	12.4	5.7	83.7
BAR1951US01	0.295	0.352	0.997	20.1	4.8	75.8
TAY1969US01	0.341	0.360	0.980	22.7	6.7	72.6
SHO1923US01	0.398	0.322	0.977	29.7	10.7	65.8
SHO1930US01	0.345	0.277	0.989	29.8	9.0	67.3
BAR1978US01	0.386	0.920	0.972	10.1	3.5	87.2
MOR1948US01	0.269	0.240	0.986	26.9	5.4	71.8
DUN2010US01	1.047	0.766	1.000	32.8	23.1	51.3
LAR2005FI01	0.945	0.478	0.977	47.5	32.1	40.7
MUL2021AT01	0.376	0.306	0.995	29.5	9.9	66.5
MOR2020IT01	0.533	0.479	0.972	26.7	12.8	65.1
MOR2020CH01	0.383	0.299	0.997	30.7	10.6	65.5
PER2021GR01	1.172	0.831	0.988	33.8	24.8	48.7
PER2021FR01	0.276	0.178	0.956	37.2	7.9	66.4
PER2021PT01	0.560	0.585	0.987	23.0	11.5	68.5
PER2021ES01	0.214	0.136	0.990	37.8	4.7	69.7
HES2022US01	2.374	1.277	0.996	44.6	38.5	26.1
HES2022US02	2.442	1.404	0.998	41.7	36.2	28.5
HES2022CA01	2.398	1.112	0.992	51.8	44.8	20.4
MEN2022BR01	1.645	1.688	0.990	23.4	18.9	61.6
PER2022US01	0.456	0.528	0.992	20.7	8.7	72.4
PER2022US02	0.333	0.373	0.983	21.5	6.1	73.8
PIN2022ES01	0.226	0.385	0.989	14.1	2.0	82.7

r = Pearson correlation coefficient between empirical and theoretical CDF values; CDF 1 d = cumulative probability of day 1.

TABLE IV. SUMMARY OF FITTED GAMMA DISTRIBUTIONS.

Variable	Mean	Median	SD	CV (%)	Min	Max
Shape	0.695	0.473	0.626	90.1	0.145	2.442
Rate	0.589	0.431	0.410	69.6	0.030	1.688
Mean (h)	33.5	29.6	22.3	66.5	8.4	115.3
Median (h)	15.4	9.4	13.9	90.1	2.0	50.4
CDF 1 d (%)	62.8	66.4	18.6	29.7	20.4	89.7

CDF 1 d = cumulative probability of day 1; SD = standard deviation; CV = coefficient of variation; Min = minimum; Max = maximum.

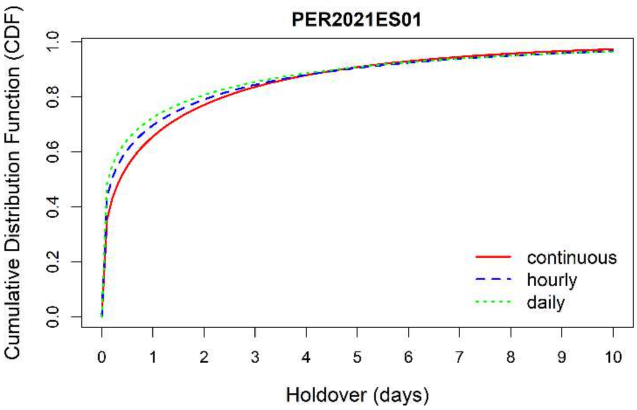

Fig. 3. Example of CDF of gamma distributions fitted with interval-censored (daily and hourly) and non-censored (continuous) holdover data.

VI. DISCUSSION

Our results show that gamma is a suitable probability distribution to describe holdover times. This is not surprising since gamma distributions are used to predict waiting times, based on the Poisson process, and that the exponential distribution is just a special case of the gamma distribution. The large variability among holdover datasets resulted in important differences between fitted gamma distributions. Although part of the variability may be due to ecological factors, (e.g., climate and dominant vegetation), we suspect that methodological aspects, such as the method used to derive holdover times and the maximum holdover considered, may influence the distributions of holdover data.

The gamma distribution has several strengths. (i) It is a simple distribution, with only two parameters, that allows for different shapes and rates of change. (ii) The probabilities follow a non-linear model, which fits the empirical distribution of holdover time (Fig. 4). (iii) Setting a maximum holdover time is not strictly necessary when using a gamma distribution, while current methods to study lightning fires require a temporal threshold to avoid wrong matches between lightning and wildfires by pure chance [2, 6].

TABLE V. GAMMA DISTRIBUTIONS FITTED WITH CONTINUOUS, HOURLY AND DAILY HOLDOVER DATA.

Holdover data	Variable	MOR2020IT01	MOR2020CH01	PER2021FR01	PER2021PT01	PER2021ES01	MEN2022BR01	PER2022US01	PER2022US02	PIN2022ES01
Continuous	Shape	0.575	0.454	0.363	0.609	0.285	1.547	0.524	0.489	0.334
	Rate	0.517	0.355	0.234	0.635	0.181	1.590	0.605	0.546	0.565
	Mean (h)	26.7	30.7	37.3	23.0	37.8	23.4	20.8	21.5	14.2
	Median (h)	13.6	12.7	12.0	12.2	8.5	18.6	9.8	9.6	4.1
	CDF 1 d (%)	64.3	63.5	62.4	67.9	65.6	62.0	71.5	71.1	81.3
Hourly	Shape	0.533	0.383	0.276	0.560	0.214	1.645	0.456	0.333	0.226
	Rate	0.479	0.299	0.178	0.585	0.136	1.688	0.528	0.373	0.385
	Mean (h)	26.7	30.7	37.2	23.0	37.8	23.4	20.7	21.5	14.1
	Median (h)	12.8	10.6	7.9	11.5	4.7	18.9	8.7	6.1	2.0
	CDF 1 d (%)	65.1	65.5	66.4	68.5	69.7	61.6	72.4	73.8	82.7
Daily	Shape	0.303	0.469	0.160	1.075	0.181	3.145	0.373	0.396	0.210
	Rate	0.285	0.364	0.109	1.085	0.120	3.251	0.481	0.479	0.401
	Mean (h)	25.4	30.9	35.2	23.8	36.3	23.2	18.6	19.9	12.6
	Median (h)	6.3	13.2	1.9	17.0	2.8	20.8	6.2	7.1	1.5
	CDF 1 d (%)	71.6	63.0	74.3	63.1	72.4	59.8	75.8	74.2	84.5

CDF 1 d = cumulative probability of day 1.

What gamma distribution to use to model probabilities of reaching holdover times? The answer is not trivial given the multiple options to consolidate multiple datasets and distributions [15]. Table VI shows four gamma distributions obtained by combining holdover datasets (Table I) and gamma distributions (Table III). In the 1st consolidation method, all 34 datasets were pooled before fitting a gamma distribution, while in the 2nd method, the holdover data were pooled using relative frequencies to avoid giving more weight to studies with more fires. The 3rd method averaged the probability density functions (PDF) of all 34 gamma distributions. Similarly to the 3rd method, the 4th method used the reliability class to obtain a weighted average density. The four gamma distributions from Table VI must be considered as preliminary before the consolidation methods are explored further.

Gamma distributions can be used to model the holdover phenomenon, although further examinations are needed to test if these distributions improve the current methods that link lightning and wildfire data [6, 13]. Likewise, gamma distributions may be combined with more sophisticated methods, based on spatial error data from LLS, to add a temporal dimension to the calculation of probabilities of lightning striking the reported ignition areas [16].

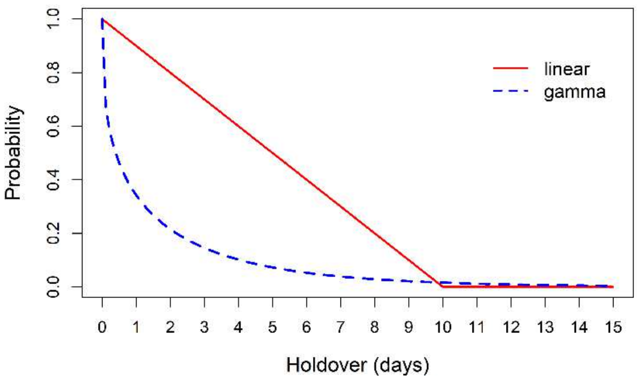

Fig. 4. Probability of a lightning fire reaching a certain holdover time according to a gamma distribution against the linear approach used in the temporal component of the index of proximity according to [13].

TABLE VI. FITTED GAMMA DISTRIBUTIONS AFTER DIFFERENT METHODS OF CONSOLIDATION.

Variable	Method 1	Method 2	Method 3	Method 4
Shape	0.313	0.392	0.379	0.331
Rate	0.302	0.297	0.271	0.238
Mean (h)	24.9	31.7	33.5	33.3
Median (h)	6.5	11.2	11.4	9.4
CDF 1 d (%)	71.7	64.6	63.9	65.8

CDF 1 d = cumulative probability of day 1; consolidations methods: 1 - pool data, 2 - pool data equally, 3 - average density, 4 - weighted average density.

VII. CONCLUSIONS

The holdover phenomenon is one of the most challenging aspects of lightning fires. Using a global database of holdover time of lightning fires, this study shows that the gamma distribution is a likely candidate to model holdover durations (i.e., the time from lightning-induced ignition to fire detection). Gamma distributions could be applied to improve the calculation of probabilities used to investigate igniting lightning and lightning-caused fires.

ACKNOWLEDGMENT

This work was supported by a postdoctoral fellowship funded by the Government of Asturias (Spain) through FICYT (AYUD/2021/58534). We would like to thank M. M. Müller, F. J. Pérez-Invernón, N. Pineda, T. D. Hessilt and L. S. Menezes for providing data to the holdover database.

REFERENCES

[1] K. Anderson, “A model to predict lightning-caused fire occurrences”, Int. J. Wildland Fire, vol. 11, pp. 163–174, 2002, http://doi.org/10.1071/WF02001

[2] B. M. Wotton and D. L. Martell, “A lightning fire occurrence model for Ontario”, Can. J. For. Res., vol. 35, pp. 1389–1401, 2005, http://doi.org/10.1139/x05-071

[3] A. J. Dowdy and G. A. Mills, “Atmospheric states associated with the ignition of lightning-attributed fires”, CAWCR Technical Report No. 019, 42 pp., 2009.

[4] N. Pineda and T. Rigo, “The rainfall factor in lightning-ignited wildfires in Catalonia”, Agr. For. Meteorol., vol. 239, pp. 249–263, 2017, https://doi.org/10.1016/j.agrformet.2017.03.016

[5] P. E. Frost, L. G. Kleyn, R. van den Dool, M. Burgess, L. Vhengani, K. Steenkamp and K. Wessels, “The Elandskraal Fire, Knysna: A data driven analysis”, CSIR Report number 271960-1, 71 pp., 2018.

[6] J. V. Moris, M. Conedera, L. Nisi, M. Bernardi, G. Cesti, and G. B. Pezzatti, “Lightning-caused fires in the Alps: Identifying the igniting strokes”, Agr. For. Meteorol., vol. 290, art. 107990, 2020, https://doi.org/10.1016/j.agrformet.2020.107990

[7] G. Cesti, M. Conedera and F. Spinedi, “Considerazioni sugli incendi boschivi causati da fulmini”, Schweiz. Z. Forstwes., vol. 156, pp. 353–361, 2005, https://doi.org/10.3188/szf.2005.0353.

[8] A. R. Taylor, “Lightning effects on the forest complex”, 9th Tall Timbers Fire Ecology Conference, pp. 127–150, 1969.

[9] M. D. Flannigan and B. M. Wotton, “Lightning-ignited forest fires in northwestern Ontario”, Can. J. For. Res., vol. 21, pp. 277–287, 1991, https://doi.org/10.1139/x91-035

[10] J.V. Moris, H. Hunt D. and D. Ascoli, “A global database of lightning-caused holdover fires”, 13th SISEF National Congress, 2022.

[11] N. Pineda, P. Altube, F. J. Alcasena, E. Casellas, H. San Segundo and J. Montanyà, “Characterizing the holdover phase of lightning-ignited wildfires in Catalonia”, in press, 2022.

[12] H. T. Gisbone, “Lightning and forest fires in the northern Rocky Mountain region”, Mon. Weather Rev., vol. 54, pp. 281–286, 1926.

[13] M. Larjavaara, J. Pennanen, and T. J. Tuomi, “Lightning that ignites forest fires in Finland”, Agr. For. Meteorol., vol. 132, pp. 171–180, 2005, https://doi.org/10.1016/j.agrformet.2005.07.005

[14] M. L. Delignette-Muller and C. Dutang, “fitdistrplus: An R package for fitting distributions”, J. Stat. Soft., vol. 64, art. 4, 2015, http://doi.org/10.18637/jss.v064.i04

[15] T. P. Hill and J. Miller, “How to combine independent data sets for the same quantity”, Chaos, vol. 21, art. 033102, 2011, http://doi.org/10.1063/1.3593373.

[16] H. G. P. Hunt, K. J. Nixon, and J. A. Naudé, “Using lightning location system stroke reports to evaluate the probability that an area of interest was struck by lightning”, Electr. Pow. Syst. Res., vol. 153, pp. 32–37, 2017, http://doi.org/10.1016/j.epsr.2016.12.010