

Article

# Population Bias on Tornado Reports in Europe

Răzvan Pîrloagă<sup>1,2,\*</sup> , Dragoș Ene<sup>1</sup>  and Bogdan Antonescu<sup>1</sup> 

<sup>1</sup> Remote Sensing Department, National Institute of Research and Development for Optoelectronics INOE 2000, Str. Atomîștilor 409, 077125 Măgurele, Romania; dragos.ene@inoe.ro (D.E.); bogdan.antonescu@inoe.ro (B.A.)  
<sup>2</sup> Faculty of Physics, University of Bucharest, 077125 Măgurele, Romania  
\* Correspondence: razvan.pirloaga@inoe.ro

**Abstract:** Tornadoes are associated with damages, injuries, and even fatalities in Europe. Knowing the spatial distribution of tornadoes is essential for developing disaster risk reduction strategies. Unfortunately, there is a population bias on tornado reporting in Europe. To account for this bias, a Bayesian modeling approach was used based on tornado observations and population density for relatively small regions of Europe. The results indicated that the number of tornadoes could be 53% higher than are currently reported. The largest adjustments produced by the model are for Northern Europe and parts of the Mediterranean regions.

**Keywords:** tornadoes; climatology; Bayesian



**Citation:** Pîrloagă, R.; Ene, D.; Antonescu, B. Population Bias on Tornado Reports in Europe. *Appl. Sci.* **2021**, *11*, 11485. <https://doi.org/10.3390/app112311485>

Academic Editor: Ben J. Anthony

Received: 2 November 2021

Accepted: 2 December 2021

Published: 3 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Tornadoes in Europe can be associated with severe damage and can result in injuries and fatalities. Despite this, until recently their threat has been underestimated. Antonescu et al. [1] showed that European tornadoes reported between 1995–2015 resulted in 4462 injuries and 316 fatalities and damages estimated at more than €1 billion. They also indicated that the density of tornado reports was the highest over Belgium, Germany, the Netherlands, and the southeastern United Kingdom. For these countries, there is a reporting bias, as central and western Europe have a high population density compared with other regions in Europe (e.g., Eastern Europe). Coastal areas were also hot-spots for tornado reports (e.g., western Italy, eastern Spain), as these areas tend to have higher population density compared with inland areas and also because of the waterspouts that move from the Mediterranean Sea inland [2].

This difference in the population density (i.e., different regions of Europe, coastal area versus inland area) introduces a bias in the reporting of tornadoes. This is because tornadoes (and also other types of severe weather events like hail or extreme winds) are “targets of opportunity” [3]. Thus, an observer needs to witness the event and then to report it and systems need to exist for collecting and verifying the reports (i.e., tornado database). Very few countries in Europe have developed and maintained such databases, which resulted in a lack of information about tornadoes [2]. One reason for not developing tornado databases is that they do not seem justified. Compared with the United States, the impact of tornadoes in European countries (given their relative small area) is relatively low, and thus there is no need to develop tornado databases for individual countries [4]. Only when considered from a pan-European perspective does the impact of tornadoes in Europe start to emerge. The collection of tornado reports and other types of severe weather reports at the pan-European level started in 2006 with the development of the European Severe Weather Database (ESWD) by the European Severe Storms Laboratory [5]. Currently, ESWD contains more than 16,500 tornado reports collected between 1800–2020.

Given the differences in population density across Europe, the real number of tornadoes is an unknown quantity. Even in relatively high populated areas, tornadoes might not be reported because of their small spatial extent and short life time, obstruction of the

observer view point (e.g., forest, hills, buildings) or if the tornado occurred during the nighttime. As indicated in previous studies for the United States [6] and Canada [7], the population density is the key factor, besides the meteorological factors, in determining the bias in tornado reports.

Several studies have addressed the issue of population bias on tornado reports. Anderson et al. [6] used a hierarchical Bayesian model to account for the population bias on tornado occurrence using historical tornado reports for the United States from the Storm Prediction Center between 1953–2001. Their results for the central and eastern United States indicated that F0–F1 tornado reports vary less with population density compared with F2–F5 tornadoes. Starting from the hypothesis that the number of tornado reports in Canada is significantly lower than the actual number of tornadoes, Cheng et al. [7] also used a Bayesian modeling approach that considered the population bias on tornado reports. Their model also included the occurrence of cloud-to-ground lightning, as cloud-to-ground lightning can be used to quantify convective storms activity and thus can be used as a precursor for tornadoes. Their results showed that in areas with low population density, the probability of tornado occurrence is significantly higher compared with the observed tornado climatology for Canada. More recently, Potvin et al. [8] developed a Bayesian hierarchical modeling framework for correcting the reporting bias in the United States tornado database. Compared with other covariates (e.g., distance from the nearest city, terrain ruggedness index, road density) population density explained more of the variance in the number of reported tornadoes. Their model indicated that approximately 45% of the tornadoes that occurred in the study domain were reported.

The aim of this article is to analyze the effects of population bias on tornado reports in Europe using a Bayesian modeling approach. The expected tornado counts over Europe can be used to better understand the societal and economic impact of tornadoes and to be included in national disaster risk reduction strategies. This article is structured as follows. Section 2 details the tornado and the population density datasets. Section 3 describes the Bayesian modeling approach. The results and discussions are presented in Sections 4 and 5, respectively. Finally, Section 6 summarizes the results.

## 2. Tornado and Population Datasets

Tornado reports were obtained from the European Severe Weather Database. The ESWD collects information on severe storms over Europe (i.e., tornadoes, severe wind, large hail, heavy rain, heavy snowfall, damaging lightning) using a citizen-science approach [9] and through collaborations with national weather services and volunteer severe weather spotter networks. Before the inclusion in the ESWD, each report is verified and receives a quality control level. In this article, tornado reports with a quality control level Q0+ (i.e., validated with meteorological data such as radar and/or satellite imagery) have been used. Unlike the United States tornado database that contains only reports for tornadoes, the ESWD also contains reports for waterspouts [2]. The ESWD include information about the surface type (e.g., land, forest, sea, lake) over which tornadoes have been observed and the surface types crossed during the event. Thus, all the waterspouts that moved inland were included in the analyses presented in this article.

Data on population density in Europe were obtained from Eurostat [10]. The data were extracted for NUTS3 regions. NUTS (Nomenclature of territorial units for statistics) classification is a system for dividing the economic territory of European Union and the United Kingdom. The NUTS 2021 [11] classification valid from 1 January 2021 contains 1166 regions at NUTS3 level. NUTS3 level represent small regions for specific diagnoses. For example, for Romania there are 42 NUTS3 regions (41 counties and Bucharest, the capital city). In previous studies, for the United States, it was argued that the rural population density at the county level is a more appropriate measure for tornado reporting compared with total population density [12]. Here, we follow [6] and use the total population density as population tends to be distributed over much of the counties and not concentrated in isolated towns. Based on the data availability and overlap with the tornado dataset, for

each NUTS3 region the population density was average for the period 2006–2019 and the area for the period 2006–2015. These data together with the number of tornadoes reported between 2006–2020 at NUTS3 level were included in the Bayesian model.

### 3. Hierarchical Bayesian Model

#### 3.1. Model

The Bayesian model used in this article starts with the hypothesis that the population density is the main influence on tornado reporting and thus, that tornadoes are underreported in Europe. As indicated by [6,7], the occurrence of tornadoes can be described as a series of conditional models linked using the Bayes’ rule. Considering the  $t_n$  as number of observed tornadoes and  $T_n$  as the true number of tornadoes ( $T_n \geq t_n$ ), a binomial model can be specified in which

$$t_n | T_n, p_n(\beta) \sim \text{Binomial}[T_n, p_n(\beta)] \tag{1}$$

where  $p_n(\beta)$  represents the probability to observe a tornado and  $n$  indicates the NUTS3 region. The probability to observe a tornado is a function of  $\beta$ , which is related to population density ( $x_n$ ). Anderson et al. (2007) used an exponential model for  $p_n$  assuming that the probability of detection increases with population density

$$p_n(\beta) = \exp(-\beta/x_n) \tag{2}$$

The true number of tornadoes ( $T_n$ ) in the  $n$  NUTS3 region is modeled as Poisson process, which is conditioned on the climatological frequency  $\lambda$

$$T_n | \lambda \sim \text{Poisson}(\lambda a_n) \tag{3}$$

where  $a_n$  is the area of the NUTS3 region and  $\lambda$  (i.e., Poisson intensity) is a measure the tornado frequency per unit area [6].

#### 3.2. Estimation

For the Bayesian approach, the prior distribution for  $\beta$  and  $\lambda$  need to be specified. These prior distributions are non-informative (i.e., large variance) because there is no prior knowledge that can inform the distributions. Thus, the population parameter  $\beta$  is specified as

$$\exp(\beta) \sim N(\mu_\beta, \sigma_\beta^2) \tag{4}$$

The distribution of  $\exp(\beta)$  is a normal distribution characterized by the mean  $\mu_\beta$  and variance  $\sigma_\beta^2$ . In the model developed in [6],  $\mu_\beta$  was set to 0.5 and  $\sigma_\beta^2$  to 10,000. For the climatological frequency parameter  $\lambda$ , a prior gamma distribution was used

$$\lambda \sim \text{gamma}(q, r) \tag{5}$$

Following [6], the shape parameter  $q$  was set to 0.001 and the scale parameter  $r$  to 0.001 corresponding to a prior mean of 1 and a prior variance of 1000 (non-informative). Using the Bayes’ rule

$$\underbrace{p(\beta, \lambda, T_1, \dots, T_N | t_1, \dots, t_N)}_{\text{Posterior model}} \propto \prod_{n=1}^N \underbrace{p(t_n | T_n, \beta)}_{\text{Data model (Equation (1))}} \times \underbrace{p(T_n | \lambda)}_{\text{Explanatory model (Equation (3))}} \times \underbrace{p(\beta) p(\lambda)}_{\text{Parameter model (Equations (4) and (5))}} \tag{6}$$

A Markov Chain Monte Carlo analysis of the Bayesian model was applied to obtain a sequence of realizations from the posterior model ([6,7]) using the WinBUGS software [13]

(Available online at <https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/>, accessed on 23 October 2021).

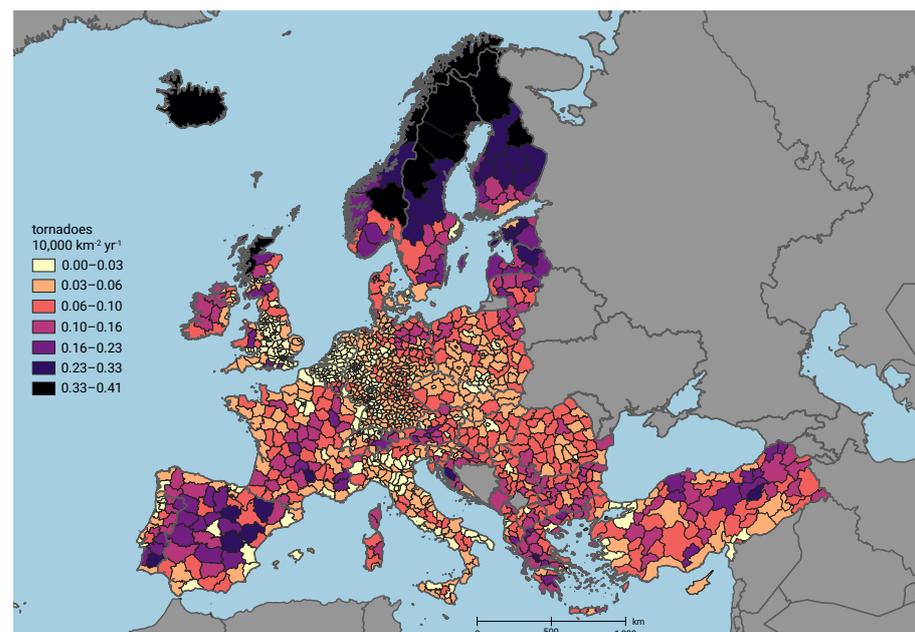
#### 4. Results

Based on the data from ESWD a total number of 2319 tornadoes were observed in Europe between 2006–2020 over the NUTS3 regions considered in this article. The predicted number of tornadoes by the model considering the population bias is 3563 tornadoes (Table 1). Thus, 65% of tornadoes predicted by the model were reported.

**Table 1.** Posterior parameters and model predictions.

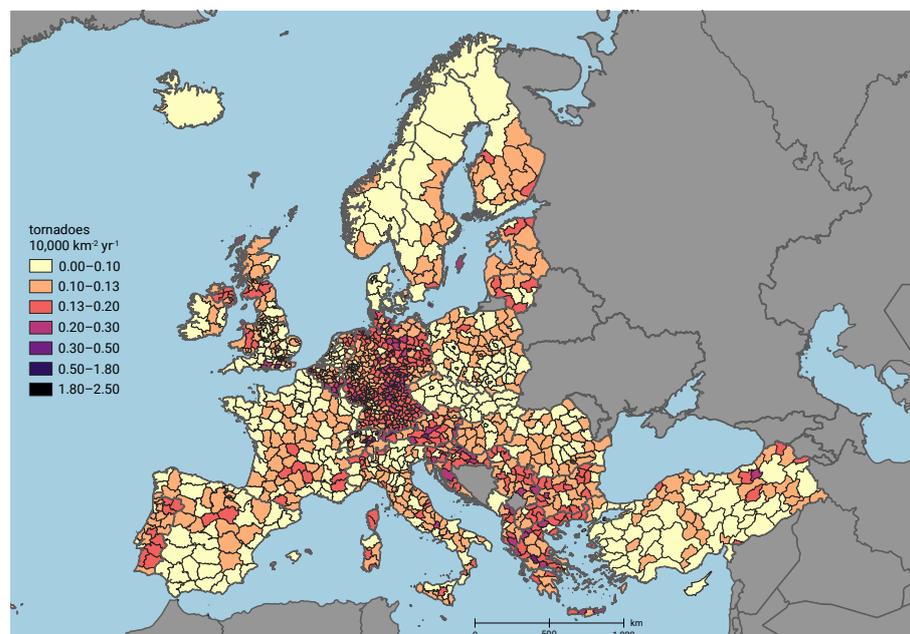
	Mean	Standard Deviation
$\beta$	14.91	1.050
$\lambda$	0.0006008	0.00001669
	Total (15 yr)	Mean (yr <sup>-1</sup> )
$\sum T_n$	3563	237.5
$\sum t_n$	2319	154.6

The difference between predicted and observed number of tornadoes is low ( $<0.027$  tornadoes  $10,000 \text{ km}^{-2} \text{ yr}^{-1}$ ) over parts of the United Kingdom, the Netherlands, Belgium, Germany, and Italy (Figure 1). As indicated previously [1,14], these are regions characterized by both a high number of tornado reports and high population density. For some NUTS3 regions there were no tornadoes reported during the study period, for example, parts of north-western and southeastern France, central Romania, central Italy, Albania, Northern Macedonia, southern Bulgaria, Finland. For these regions the predicting values of tornado counts are less than 2 tornadoes over the 15 years study period. The highest difference between predicted and observed tornadoes (between  $0.33\text{--}0.41$  tornadoes  $10,000 \text{ km}^{-2} \text{ yr}^{-1}$ ) is over Iceland and northern parts of the United Kingdom, Norway, Sweden, and Finland (Figure 1). This is not surprising given that the population density of these regions is low compared with other regions of Europe and also given that the number of observed tornadoes by NUTS3 regions in this area is less than 10 tornadoes over the entire study period.



**Figure 1.** Difference between posterior adjusted number of tornadoes and the number of observed tornadoes normalized by the area of NUTS3 regions and year (shaded according to the scale,  $10,000 \text{ km}^{-2} \text{ yr}^{-1}$ ).

Low values for standard deviation ( $<0.13$  tornadoes  $10,000 \text{ km}^{-2} \text{ yr}^{-1}$ ) of the tornado occurrence from the posterior distribution are found over most of Europe (Figure 2), with the exception of an area stretching from eastern Germany over Austria, Croatia, Bulgaria, and Greece characterized by values between  $0.13$ – $0.30$  tornadoes  $10,000 \text{ km}^{-2} \text{ yr}^{-1}$ . In these regions the NUTS3 areas are characterized by relatively small area and low population density and also a low number of observed tornadoes during the study period.



**Figure 2.** As in Figure 1, but for the standard deviation of tornado occurrence from posterior distribution.

## 5. Discussion

The Bayesian model used in this article only included the population effects, but previous studies using a similar approach have also included meteorological data [7]. These are large areas where the data collection is unreliable and thus can reduce the predictive capacity of the model. In their study of the probability of tornado occurrence across Canada, Cheng et al. [7] included the cloud-to-ground lightning climatology to account for the spatial variability of tornadoes. For Europe, an indication regarding the predictive capacity of the model can be obtained by comparing the results from Figure 1 with lightning density over Europe between 2008–2012 developed by Anderson and Klugmann [15] using data from the Arrival Time Differing NETWORK. Over northern Norway, Sweden, and Finland, where NUTS3 have a large area and the collection of data is not as reliable as for other regions of Europe, the Bayesian model is introducing large adjustments. For these regions, the lightning density (Figure 4 from [15]) is lower ( $<0.4$  flashes  $\text{km}^{-2} \text{ yr}^{-1}$ ) compared with almost any other regions in Europe. Thus, the adjustments in the number of tornadoes are less realistic for the regions from a meteorological point of view. For the Baltic states (i.e., Estonia, Latvia, Lithuania) the corrections from the model are realistic due to the relatively large lightning density (i.e.,  $0.4$ – $2.5$  flashes  $\text{km}^{-2} \text{ yr}^{-1}$ ) in this region. The predictive capacity of the model in these areas can be improved by improving data collection (e.g., ref. [16] used satellite data to obtain information on unreported tornadoes that occurred in forested regions) or by considering as covariate meteorological factors related to tornado occurrence [17].

Future research will develop the model by considering meteorological covariates such as lightning density (e.g., from Arrival Time Difference long-range lightning detection network [18]) and tornadic environments (e.g., ERA5 reanalysis data [19]).

## 6. Conclusions

In this article, a Bayesian model was applied to adjust the number of observed tornadoes over Europe. The hypothesis was that the number of observed tornadoes in Europe is lower than the real number due to the differences in population density. The model indicated that the average annual number of tornadoes during the study period (2006–2020) was 237.5 compared with an annual average of 154.6 observed tornadoes. The largest adjustments occur over northern Europe (e.g., Iceland, Norway, Sweden, Finland), but also parts of the Mediterranean region (e.g., Spain, Greece).

The corrected distribution of tornadoes in Europe can be used to better understand the risk posed by European tornadoes. Compared with the previous distribution of tornadoes in Europe, the distribution obtained in this article is more relevant for risk reduction strategies, as it is including a correction for the population bias on tornado reporting. Furthermore, the current results can be used by decision-makers and emergency managers to develop disaster risk reduction strategies for tornadoes. Very few countries in Europe have developed tornado preparedness and response programs.

**Author Contributions:** Conceptualization, B.A.; methodology, R.P.; formal analysis, R.P., D.E. and B.A.; original draft preparation, R.P., D.E. and B.A.; visualization, R.P.; funding acquisition, B.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Romanian Ministry of Education and Research, CNCS-UEFISCDI (Project No. PN-III-P1-1.1-TE-2019-0649) within PNCDI III.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data used in this study can be obtained by request to the authors.

**Acknowledgments:** This work was supported by the Romanian National Core Program (Contract No. 18N/2019) and by a grant of the Romanian Ministry of Education and Research, CNCS-UEFISCDI (Project No. PN-III-P1-1.1-TE-2019-0649) within PNCDI III. Further support was provided by the European Regional Development Fund through the Competitiveness Operational Programme 2014–2020, Action 1.1.3 Creating synergies with H2020 Programme, project H2020 Support Centre for European project management and European promotion, MYSMIS code 107874, project Strengthen the participation of the ACTRIS-RO consortium in the pan-European research infrastructure ACTRIS, ACTRIS-ROC, MYSMIS code 107596 (contract No.337/2021). The authors thank Cristina Marin for her comments on an earlier version of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Antonescu, B.; Schultz, D.M.; Holzer, A.; Groenemeijer, P. Tornadoes in Europe: An underestimated threat. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 713–728. [\[CrossRef\]](#)
2. Antonescu, B.; Schultz, D.M.; Lomas, F.; Kühne, T. Tornadoes in Europe: Synthesis of the Observational Datasets. *Mon. Weather Rev.* **2016**, *144*, 2445–2480. [\[CrossRef\]](#)
3. Brooks, H.E. Severe thunderstorm and climate change. *Atmos. Res.* **2013**, *123*, 129–138. [\[CrossRef\]](#)
4. Doswell, C.A. Societal impacts of severe thunderstorms and tornadoes: Lessons learned and implications for Europe. *Atmos. Res.* **2003**, *67–68*, 135–152. [\[CrossRef\]](#)
5. Groenemeijer, P.; Púčik, T.; Holzer, A.M.; Antonescu, B.; Riemann-Campe, K.; Schultz, D.M.; Kühne, T.; Feuerstein, B.; Brooks, H.E.; Doswell, C.A.; et al. Severe convective storms in Europe: Ten years of research and education at the European Severe Storms Laboratory. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 2641–2651. [\[CrossRef\]](#)
6. Anderson, C.J.; Winkle, C.K.; Zhou, Q.; Royle, J.A. Population influences on Tornado Reports in the United States. *Weather Forecast.* **2007**, *22*, 571–579. [\[CrossRef\]](#)
7. Cheng, V.Y.S.; Arhonditsis, G.B.; Sills, D.M.L.; Auld, H.; Shephard, M.W.; Gough, W.A.; Klaassen, J. Probability of tornado occurrence across Canada. *J. Clim.* **2013**, *26*, 9415–9428. [\[CrossRef\]](#)
8. Potvin, C.K.; Broyles, C.; Skinner, P.S.; Brooks, H.E.; Rasmussen, E. A Bayesian hierarchical modeling framework for correcting reporting bias in the U.S. tornado database. *Weather Forecast.* **2019**, *34*, 15–30. [\[CrossRef\]](#)
9. Muller, C.L.; Chapman, L.; Johnston, S.; Kidd, C.; Illingworth, S.; Foody, G.; Overeem, A.; Leigh, R.R. Crowdsourcing for climate and atmospheric sciences: Current status and future potential. *Int. J. Climatol.* **2015**, *35*, 3185–3203. [\[CrossRef\]](#)

10. Eurostat Database. Available online: <https://ec.europa.eu/eurostat/data/database> (accessed on 1 November 2021).
11. Nomenclature of Territorial Units for Statistics (NUTS). Available online: <https://ec.europa.eu/eurostat/documents/345175/629341/NUTS2021.xlsx> (accessed on 1 November 2021).
12. Changnon, S.A. Trends in tornado frequencies: Fact or fallacy. In Proceedings of the Preprints 12th Conference on Severe Local Storms, San Antonio, TX, USA, 12–15 January 1982; pp. 42–44.
13. Lunn, D.J.; Thomas, A.; Best, N.; Spiegelhalter, D. WinBUGS—A Bayesian modelling framework: Concepts, structure, and extensibility. *Stat. Comput.* **2000**, *10*, 325–337. [[CrossRef](#)]
14. Groenemeijer, P.; Kühne, T. A climatology of tornadoes in Europe: Results from the European Severe Weather Database. *Mon. Weather Rev.* **2014**, *142*, 4775–4790. [[CrossRef](#)]
15. Anderson, G.; Klugmann, D. A European lightning density analysis using 5 years of ATDnet data. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 815–829. [[CrossRef](#)]
16. Shikhov, A.; Chernokulsky, A. A satellite-derived climatology of unreported tornadoes in forested regions of northeast Europe. *Remote Sens. Environ.* **2018**, *204*, 553–567. [[CrossRef](#)]
17. Taszarek, M.; Allen, J.T.; Púčik, T.; Hoogewind, K.A.; Brooks, H.E. Severe convective storms across Europe and the United States. Part II: ERA5 environments associated with lightning, large hail, severe wind, and tornadoes. *J. Clim.* **2020**, *33*, 10263–10286. [[CrossRef](#)]
18. Enno, S.E.; Sugier, J.; Alber, R.; Seltzer, M. Lightning flash density in Europe based on 10 years of ATDnet data. *Atmos. Res.* **2020**, *235*, 104769. [[CrossRef](#)]
19. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [[CrossRef](#)]