A Climatology of Thunderstorms across Europe from a Synthesis of Multiple Data Sources

MATEUSZ TASZAREK,a,b JOHN ALLEN,c TOMÁŠ PŮČIK,d PIETER GROENEMEIJER,d BARTOSZ CZERNECKI,a LESZEK KOLENDOWICZ,a KOSTAS LAGOUVARDOS,e VASILIKI KOTRONI,e and WOLFGANG SCHULZf

a Department of Climatology, Adam Mickiewicz University, Poznań, Poland
b Skywarn Poland, Warsaw, Poland
c Central Michigan University, Mount Pleasant, Michigan
d European Severe Storms Laboratory, Wessling, Germany
e Institute of Environmental Research and Sustainable Development, National Observatory of Athens, Athens, Greece
f Austrian Electrotechnical Association–Austrian Lightning Detection and Information System (OVE-ALDIS), Vienna, Austria

(Manuscript received 10 June 2018, in final form 3 December 2018)

ABSTRACT

The climatology of (severe) thunderstorm days is investigated on a pan-European scale for the period of 1979–2017. For this purpose, sounding measurements, surface observations, lightning data from ZEUS (a European-wide lightning detection system) and European Cooperation for Lightning Detection (EUCLID), ERA-Interim, and severe weather reports are compared and their respective strengths and weaknesses are discussed. The research focuses on the annual cycles in thunderstorm activity and their spatial variability. According to all datasets thunderstorms are the most frequent in the central Mediterranean, the Alps, the Balkan Peninsula, and the Carpathians. Proxies for severe thunderstorm environments show similar patterns, but severe weather reports instead have their highest frequency over central Europe. Annual peak thunderstorm activity is in July and August over northern, eastern, and central Europe, contrasting with peaks in May and June over western and southeastern Europe. The Mediterranean, driven by the warm waters, has predominant activity in the fall (western part) and winter (eastern part) while the nearby Iberian Peninsula and eastern Turkey have peaks in April and May. Trend analysis of the mean annual number of days with thunderstorms since 1979 indicates an increase over the Alps and central, southeastern, and eastern Europe with a decrease over the southwest. Multiannual changes refer also to changes in the pattern of the annual cycle. Comparison of different data sources revealed that although lightning data provide the most objective sampling of thunderstorm activity, short operating periods and areas devoid of sensors limit their utility. In contrast, reanalysis complements these disadvantages to provide a longer climatology, but is prone to errors related to modeling thunderstorm occurrence and the numerical simulation itself.

1. Introduction

Thunderstorms, particularly severe events accompanied by large hail, damaging wind gusts, tornadoes, or flash floods, pose a considerable risk to society (Brooks 2013; Papagiannaki et al. 2013; Terti et al. 2017; Papagiannaki et al. 2017). Therefore, knowledge of their local climatology is not only important for weather forecasting purposes, but also for risk assessment by emergency managers or the (re)insurance industry. Another pressing question is whether such phenomena are becoming more frequent as a result of changing climate (e.g., Trapp et al. 2007; Kapsc et al. 2012; Allen et al. 2014; Seeley and Romps 2015; Gensini and Mote 2015; Allen 2018). To answer this question, a reliable record of observations over a period of many years is necessary, which is challenging, particularly where direct observations over long periods are sparse. A number of approaches have been taken in the past to tackle this issue in different regions.

A straightforward way to build a thunderstorm climatology is to use direct observations from manned weather stations, some of which offer decades of observations (even up to 100-yr periods; Changnon and Changnon 2001; Bielec-Bąkowska 2003). Important disadvantages are that

Denotes content that is immediately available upon publication as open access.

Corresponding author: Mateusz Taszarek, mateusz.taszarek@amu.edu.pl

DOI: 10.1175/JCLI-D-18-0372.1
© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
human contributions to these observations introduce errors, such as inhomogeneities (Czernicki et al. 2016), and that periods for which observations are available may be intermittent. Furthermore, the spatial coverage of observing stations may be too dispersed to capture the scale of most thunderstorms. Stations also seldom offer observations of severe weather, such as tornadoes or large hail. An exception to this pattern is a Chinese dataset of hail size observations and reasonably high spatial density that has allowed for analyses of long-term trends (Xie et al. 2008; Li et al. 2016). To capture severe events occurring between manned stations, their records can be supplemented with reports from the general public. Such observations may include severe weather reports from trained spotters or from untrained individuals. Recent advances in technology have allowed the development of crowd-sourcing techniques to collect such reports (e.g., Dotzek et al. 2009; Elmore et al. 2014; Seimon et al. 2016; Holzer et al. 2017; Groenemeijer et al. 2017). While these techniques increase our ability to detect severe thunderstorms, considerable temporal and spatial inhomogeneity exists for historical severe thunderstorm observations developed in the United States, Europe, or Australia (Dotzek et al. 2009; Tippett et al. 2015; Allen and Allen 2016).

A second method for monitoring thunderstorm activity is by remote sensing, such as by radar systems, lightning detection networks, and satellite-based sensors. These sensors exhibit fewer spatial and temporal inhomogeneities than in situ observations. A number of thunderstorm climatologies have been based on lightning detection networks at national, continental, or global scales (e.g., Betz et al. 2009; Pohjola and Mäkelä 2013; Virts et al. 2013; Wu et al. 2016; Galanaki et al. 2018; Zhang et al. 2018). Satellites equipped with lightning sensors can be used for the same purpose, but are often biased toward equatorial and tropical zones or only cover certain regions for short periods (Price and Rind 1992; Christian et al. 2003; Cecil et al. 2014, 2015; Dewan et al. 2018). Geostationary satellites equipped with lightning detectors such as the GOES-16 and GOES-17 Geostationary Lightning Mapper (GLM) can continuously observe higher latitudes but no such satellite is available for Europe yet (the first European lightning imager on geostationary orbit is scheduled for launch in 2019; Dobber and Grandell 2014). Climatological aspects of (severe) thunderstorms can also be studied using national radar networks. Although several efforts have been made to study short-term climatologies of large hail and storm tracks (e.g., Davini et al. 2011; Cintineo et al. 2012; Kaltenböck and Steinheimer 2015), inconsistencies in the types of radars used by different countries limit the possibility of this approach for Europe. Furthermore, the indirect nature of remotely sensed data leads to several limitations, such as overdetection (Allen et al. 2015; Tippett et al. 2015). For example, using cloud-top temperatures to approximate hail occurrence is associated with a significant false-alarm rate over tropical areas. The changing quality and homogeneity over time with short temporal coverage can also make these data unsuitable for longer-term climatologies (Tippett et al. 2015).

A third approach is the use of environmental proxies (convective parameters), which has also become widespread, particularly where observational data are sparse or inconsistent through time (Brooks et al. 2003; Brooks 2013; Allen 2018). These proxies are based on the knowledge of the conditions favorable to the formation of thunderstorms (i.e., the ingredients-based approach of identifying instability, lower-tropospheric moisture, and a triggering mechanism; Doswell et al. 1996). These ingredients can be applied as a nonconditional proxy for the probability that a (severe) thunderstorm occurs. A number of studies have demonstrated that the likelihood of severe convection increases along with increasing instability and increasing vertical wind shear that governs the organization and longevity of updrafts (e.g., Weisman and Klemp 1982; Brooks et al. 2003; Allen et al. 2011; Püäikä et al. 2015; Westermayer et al. 2017; Taszarek et al. 2017). Using this information, the presence of (severe) thunderstorms can be inferred from large-scale numerical weather model data or from observed soundings. In this way, present and future climatologies of severe thunderstorms have been developed for the United States, Europe, Australia, and the globe using reanalysis datasets and climate models (e.g., Marsh et al. 2007, 2009; Trapp et al. 2011; Allen and Karoly 2014; Gensini and Mote 2014; Püäikä et al. 2017; Taszarek et al. 2018). The primary limitation of this method lies in the environmental proxies being only an imperfect approximation of (severe) thunderstorm activity, as not every potentially favorable environment produces a severe thunderstorm or a thunderstorm at all. This issue may be overcome using high-resolution, convection-allowing models that can directly simulate convective storms. However, these approaches are also imperfect representations, and the high demand for computational power in such simulations precludes both spatial and temporal extent of this approach (Allen 2018).

Although thunderstorm climatologies based on a variety of sources have been constructed for multiple countries, a multidataset study on a pan-European scale is lacking. Therefore, the main aim of this work is to combine these data sources (i.e., manned observations of thunderstorms, severe weather reports, lightning detection networks, sounding observations, and reanalysis data) to construct a comprehensive European climatology of (severe) thunderstorms. We compare the
results obtained from different datasets, highlight their strengths and weaknesses across Europe, and show whether a simple, threshold-based approach of environmental proxies is able to reasonably model the observed occurrence of (severe) thunderstorms. This will be accomplished by considering the spatial distribution of events, their annual cycles and regional variability of days with (severe) thunderstorms, and trends of thunderstorm environments over the last 39 years. First, the methodology and datasets are described, followed by the resulting spatiotemporal distributions of (severe) thunderstorm days, and a comparison of these results with earlier studies and outlining future directions.

2. Dataset and methodology

a. The concept of a thunderstorm day and severe thunderstorm day

To investigate climatological aspects of thunderstorms in Europe, the concept of a thunderstorm day (TD) and severe thunderstorm day (SevTD) was adopted. This approach has been used in many previous studies (e.g., Sakamoto 1973; Falconer 1984; Jacovides and Yonetani 1990; Watson and Holle 1996; Novák and Kyzarová 2011; Pinto et al. 2013) and has multiple advantages. First, it allows us to compare databases with different spatial and temporal resolutions (e.g., surface observations, reanalysis). Second, it is less susceptible to variations of data quality and detection/reporting efficiency in time and space (e.g., lightning networks, severe weather reports). Last, it provides empirical values, which can be easily compared with other studies and interpreted for societal applications. The definition of a TD and SevTD for each database is discussed in the following sections and presented in Table 1.

2. Dataset and methodology

a. The concept of a thunderstorm day and severe thunderstorm day

To investigate climatological aspects of thunderstorms in Europe, the concept of a thunderstorm day (TD) and severe thunderstorm day (SevTD) was adopted. This approach has been used in many previous studies (e.g., Sakamoto 1973; Falconer 1984; Jacovides and Yonetani 1990; Watson and Holle 1996; Novák and Kyzarová 2011; Pinto et al. 2013) and has multiple advantages. First, it allows us to compare databases with different spatial and temporal resolutions (e.g., surface observations, reanalysis). Second, it is less susceptible to variations of data quality and detection/reporting efficiency in time and space (e.g., lightning networks, severe weather reports). Last, it provides empirical values, which can be easily compared with other studies and interpreted for societal applications. The definition of a TD and SevTD for each database is discussed in the following sections and presented in Table 1.

2. Dataset and methodology

a. The concept of a thunderstorm day and severe thunderstorm day

To investigate climatological aspects of thunderstorms in Europe, the concept of a thunderstorm day (TD) and severe thunderstorm day (SevTD) was adopted. This approach has been used in many previous studies (e.g., Sakamoto 1973; Falconer 1984; Jacovides and Yonetani 1990; Watson and Holle 1996; Novák and Kyzarová 2011; Pinto et al. 2013) and has multiple advantages. First, it allows us to compare databases with different spatial and temporal resolutions (e.g., surface observations, reanalysis). Second, it is less susceptible to variations of data quality and detection/reporting efficiency in time and space (e.g., lightning networks, severe weather reports). Last, it provides empirical values, which can be easily compared with other studies and interpreted for societal applications. The definition of a TD and SevTD for each database is discussed in the following sections and presented in Table 1.

b. Reanalysis data

The European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERA-Interim; Dee et al. 2011) was used in the analysis. The dataset has 0.75° horizontal grid spacing with 29 pressure levels extending from 1000 to 50 hPa. The grid-based model

### Table 1. Description of databases used in the analysis and the definition of a thunderstorm day (TD) and severe thunderstorm day (SevTD).

<table>
<thead>
<tr>
<th>Database</th>
<th>Short name</th>
<th>Resolution</th>
<th>Coverage</th>
<th>Definition of a TD</th>
<th>Definition of a SevTD</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUCLID and ZEUS</td>
<td>Lightning</td>
<td>Detections gridded to 0.5°×0.5°</td>
<td>2008–17</td>
<td>&gt;2 flashes</td>
<td>—</td>
</tr>
<tr>
<td>ZEUS</td>
<td>Synoptics</td>
<td>Observational range of ~15–18 km</td>
<td>1979–2017</td>
<td>Report of an audible thunder</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Soundings</td>
<td>Representative air mass within the range of ~50–200 km</td>
<td>1979–2017</td>
<td>ML CAPE &gt;150 J kg⁻¹, ML CIN &gt;−75 J kg⁻¹</td>
<td>ML CAPE &gt;150 J kg⁻¹, convective precipitation &gt;0.075 mm h⁻¹</td>
</tr>
<tr>
<td></td>
<td>ERA-Interim</td>
<td>Gridded to 0.75°×0.75°</td>
<td>1979–2017</td>
<td>ML CAPE &gt;150 J kg⁻¹, convective precipitation &gt;0.075 mm h⁻¹</td>
<td>ML CAPE &gt;150 J kg⁻¹, convective precipitation &gt;0.075 mm h⁻¹, ML WMAXSHEAR &gt;400 m² s⁻²</td>
</tr>
<tr>
<td></td>
<td>ESWD</td>
<td>Gridded to 1.5°×1.5°</td>
<td>2011–17</td>
<td>—</td>
<td>Severe convection weather report</td>
</tr>
</tbody>
</table>

**footnotes:**

a. Enno (2015); Czernecki et al. (2016).

b. Brooks et al. (1994); Potvin et al. (2010).
profiles are extended to a thirtieth level by including the surface with temperature and dewpoint for 2 m above ground level (AGL) and $U$ and $V$ wind vectors for 10 m AGL. The research domain covers almost the entire European continent (Fig. 1a), containing 59 latitudinal (from $31.5^\circ$ to $75^\circ$N) and 108 meridional (from $27^\circ$W to $53.25^\circ$E) grid points (i.e., a total of 6372 points). The temporal resolution of the resulting dataset includes 56,980 time steps (1979–2017) at 6-hourly intervals (0000, 0600, 1200, and 1800 UTC). For each time step, the mixed-layer (0–500 m AGL) convective available potential energy (ML CAPE), the 10 m to 6 km AGL deep-layer shear (DLS), and a combined parameter (ML WMAXSHEAR; $\sqrt{2 \times (MLCAPE)} \times DLS$; Brooks 2013; Taszarek et al. 2017, 2018) are computed.

c. Soundings

Radiosonde measurements were acquired from the University of Wyoming sounding database (http://weather.uwyo.edu/upperair/sounding.html). For the years 1979–2017 all available measurements for 1200 UTC were downloaded from 119 stations covering the research domain (Fig. 1a). Although it was also possible to use 0000, 0600, and 1800 UTC data, the majority of stations had an unequal ratio of these measurements (e.g., only a few stations had all time steps available, some had 1200 and 0000 UTC, while others only 1200 or 0000 UTC). Thus, we only consider observations at 1200 UTC, because it best represents the typical preconvective storm environment as demonstrated by cloud-to-ground lightning activity peaking between 1400 and 1600 UTC (e.g., Kotroni and Lagouvardos 2008; Wapler 2013; Mäkelä et al. 2014; Taszarek et al. 2015; Galanaki et al. 2015; Poelman et al. 2016). The highest frequency of severe thunderstorm phenomena such as tornadoes, large hail, or severe wind gusts similarly peaks between 1500 and 1800 UTC (e.g., Groenemeijer and Kühne 2014; Taszarek and Brooks 2015; Punge and Kunz 2016; Celiński-Mysław and Palarz 2017; Groenemeijer et al. 2017). Similar to the procedures for ERA-Interim, ML CAPE, DLS, and ML WMAXSHEAR were also computed. To filter soundings containing errors in measurements (erroneous or unrealistic values), we applied quality-control assumptions following the procedures of Taszarek et al. (2018).

d. Lightning data

A 10-yr (2008–2017) record of lightning data that combines both the European Cooperation for Lightning Detection (EUCLID; Poelman et al. 2016; Schulz et al. 2016) and ZEUS (Kotroni and Lagouvardos 2008, 2016) European lightning location systems is used in the analysis. The EUCLID network consists of 149 surface sensors mostly located in a triangle between southern Portugal, northern Norway, and southern Italy with a spatial accuracy of lightning detection between 100 and 500 m, and detection efficiency exceeding 90% (Schulz et al. 2016). Although the number of sensors slightly increased in recent years, it can be assumed that, from 2008 onward, the improvements to the network in terms of flash detection efficiency and location accuracy should have the minimal influence on the results. Even though stroke detection efficiency has improved, the flash detection efficiency remained rather stable, which should not influence estimates of days with detection (Schulz et al. 2016).

ZEUS is a long-range detection network operated by the National Observatory of Athens. Sensors are located at six sites: Chilbolton (United Kingdom), Roskilde (Denmark), Iaşi (Romania), Larnaka (Cyprus), Mazagon (Spain), and Alexandria (Egypt). ZEUS has a lower location accuracy and detection efficiency than EUCLID, of approximately 6.5 km and 25% respectively (Lagouvardos et al. 2009); however, it supplements lightning data over southern and southeastern parts of Europe where EUCLID is devoid of sensors. Even though ZEUS has lower lightning detection efficiency, this negative aspect is strongly minimized when focusing on TDs instead of absolute values such as lightning density (Lagouvardos et al. 2009; Price et al. 2011). It should also be noted that out of the 10-yr period, data from December 2015 to April 2016 are missing due to software failure. Since the fraction of nonavailable days is 4%, and it coincides with a low annual regime for lightning activity and overlaps with EUCLID (except for the southern and southeastern edges of the domain), this does not considerably alter the robustness of results.

The numbers of lightning detections (flashes) from both networks are summed and gridded in 0.5° boxes at 1-day intervals over the domain shown in Fig. 1a. The size of the grid box is a compromise between locational accuracy of both networks and ease of comparison with reanalysis data sounding and surface observational data. We consider a TD to occur if more than two flashes are detected in a grid box (Table 1). This value was chosen to avoid the possibility of false thunderstorm detections.

e. Surface observations

Human observations of thunderstorms were taken from surface synoptic observations (SYNOP) reports acquired from the NOAA National Climatic Data Center (NCDC; now the National Centers for Environmental Information) for the equivalent period to the reanalysis (1979–2017). The database contains 11 million daily summaries from 821 stations located throughout the research domain (Fig. 1b). The stations chosen had at least 15 years of observations, and the
FIG. 1. (a) Location of sounding stations (white points) and severe convective weather reports (red points) used in the analysis. The dark red polygon denotes domain used for the proximity sounding analysis associated with severe and nonsevere thunderstorms. The dark orange polygon indicates the lightning domain (EUCLID and ZEUS networks) and area used for the proximity sounding analysis associated with lightning and nonlightning events. (b) Location of SYNOP stations (black points) and sites for which annual cycles are being analyzed in the latter part of the study (red points). ERA-Interim orography (m MSL) is indicated by the shaded color scale.
remaining stations from the initial 1738 sites were removed. At some locations thunderstorm observations were discontinued due to station relocation or automation (e.g., Germany). In such cases only years where these observations were available were taken into account. Detection of a TD is considered if there is at least one report of audible thunder on a day.

**f. Severe weather reports**

Reports of tornadoes, large hail (at least 2 cm in diameter), severe convective wind gusts, and heavy convective precipitation were taken from the European Severe Weather Database (ESWD; Dotzek et al. 2009) for the years 2011–17 (Fig. 1a). According to ESWD definitions, a severe wind gust is considered “to have a speed of at least 25 m s$^{-1}$ or one doing such damage that a wind speed of 25 m s$^{-1}$ or higher is likely to have occurred.” Heavy rain is defined as “rain falling in such large amounts, that significant damage is caused, or no damage is known, but exceptionally high precipitation amounts have been observed within a period of at most 24 hours.” A tornado “is a vortex extending between a convective cloud and the Earth’s surface, in which the wind is strong enough to cause damage to objects” (further details on ESWD reporting criteria are available at the ESSL web page, www.essl.org). Only cases with a credibility status of QC0+ (plausibility check passed), QC1 (report confirmed), and QC2 (event fully verified) were included in the analysis. Reports that did not contain information whether an event had originated from a convective source were cross-referenced with lightning data from EUCLID and ZEUS networks. If the event was associated with a lightning detection at the same grid point, it was assigned as convective; otherwise, it was excluded from the analysis. This filtering was applied in order to exclude severe wind and heavy rain events not associated with a deep moist convection.

In the final phase all reports were gridded to 1.5° grid boxes at daily intervals. The size of the grid box was chosen as a compromise between sample size and consistency of the obtained results. A SevTD is defined if at least one severe weather report is available within a given day and grid box (Table 1).

**g. Proximity sounding analysis**

To estimate the frequency of thunderstorm and severe thunderstorm environments, we performed a proximity analysis [for further details on this methodology, see Potvin et al. (2010)] associated with lightning and severe weather events. Previous studies have used a variety of distance thresholds ranging from 40 to 400 km (e.g., McCaul 1991; Brooks et al. 1994; Rasmussen and Blanchard 1998; Thompson et al. 2003; Cohen et al. 2007; Púčik et al. 2015). In this study we used a threshold of 125 km, similar to Taszarek et al. (2017), which for central Europe is a compromise between the representativeness of soundings and the limiting of sample size that a more stringent proximity criteria would impose. Temporally, soundings were deemed proximal if the event took place up to 4 h following the sounding, focusing on an assessment of the preconvective environment. For each observed sounding site, we also used the nearest (by geographical
distance) grid point from ERA-Interim, and produced an equivalent model sounding.

All lightning cases that occurred in the years 2008–17 within the EUCLID and ZEUS domains (Fig. 1a) and met the proximity criteria were included. As the occurrence of lightning is strongly dependent on the availability of thermodynamic instability (Craven and Brooks 2004; Kaltenbäck et al. 2009; Westermayer et al. 2017; Kolendowicz et al. 2017; Taszarek et al. 2017) we use ML CAPE as the discriminator between lightning and nonlightning cases. To focus exclusively on the environments that are unstable, all zero CAPE cases were excluded from the analysis. Proximity analysis indicates that the 75th percentile of nonlightning cases have ML CAPE below 150 J kg$^{-1}$ for both reanalysis and soundings (Fig. 2). Lightning cases are associated with a considerably higher ML CAPE with the 25th percentile for reanalysis and soundings having a value of 175 and 125 J kg$^{-1}$, respectively. Based on these distributions, a threshold of 150 J kg$^{-1}$ is defined as a proxy for a TD.

ML CIN for soundings (proxy for convective initiation; Gensini and Ashley 2011; Diffenbaugh et al. 2013) is used as an additional criteria to eliminate nonlightning cases. For ERA-Interim, convective precipitation was used instead of ML CIN as the reanalysis is able to estimate areas for convective initiation. A similar approach was also used by Trapp et al. (2009), Půčík et al. (2017), and Groenemeijer et al. (2017). For both categories, the 60th percentile (arbitrary decision) of nonlightning cases is used to eliminate cases not associated with thunderstorms (ML CIN $< -75$ J kg$^{-1}$ and convective precipitation $< 0.075$ mm h$^{-1}$ for sounding and reanalysis, respectively).

The criteria for discriminating between severe and nonsevere thunderstorm environments are defined on the basis of proximity soundings associated with ESWD reports. Since tornadoes occurring over water surface (waterspouts) and heavy convective rain events can occur in a variety of environmental conditions (Půčík et al. 2015; Taszarek et al. 2017), these are excluded from
Mean annual number of days with thunderstorm within months (lightning data)

January  
February  
March 

April  
May  
June 

July  
August  
September 

October  
November  
December 

0 5 10 15

FIG. 4. Mean annual number of days with thunderstorm in each month, based on EUCLID and ZEUS lightning data. The definition of a TD is provided in Table 1.
proximity analysis. A domain covering the highest density of reports (Fig. 1a) and years 2011–17 is chosen for the proximity analysis in order to minimize the effects of the spatial inhomogeneity of the reports, which have a bias toward central Europe (Groenemeijer and Kühne 2014; Groenemeijer et al. 2017). As Brooks et al. (2003) indicated, the probability of convective hazards is predominantly a function of thermodynamic instability and vertical wind shear, though different weights of the respective terms apply for individual severe convective storm phenomena (Brooks 2013). ML WMAXSHEAR (and its modifications) provides a combination of the two and has been shown to discriminate well between severe and nonsevere thunderstorms (Craven and Brooks 2004; Brooks 2009, 2013; Allen et al. 2011; Allen and Karoly 2014; Púčik et al. 2015; Taszarek et al. 2017). Therefore, this covariate is used for all lightning environments to separate between nonsevere and severe thunderstorm events (Fig. 2). Despite these categories partially overlapping, the 75th percentile of nonsevere \(450 \text{ m}^2 \text{s}^{-2}\) and 25th percentile of severe category \(350 \text{ m}^2 \text{s}^{-2}\) allow setting a discriminator of \(400 \text{ m}^2 \text{s}^{-2}\). However, we are aware that no threshold will perfectly distinguish between these categories (Doswell and Schultz 2006). A summary of TD and SevTD criteria is provided in Table 1.

3. Dataset limitations
Each database used in this study features disadvantages and advantages over the others. Thunderstorm observations performed by humans at meteorological stations are usually limited to detection of storms within a radius of 15–18 km (Reap and Orville 1990; Enno 2015; Czernecki et al. 2016), which entails an areal coverage of approximately 1000 km\(^2\). Since a human factor is involved, storm reporting efficiency strongly depends on the perception of individuals and thus

![Day of the year with peak thunderstorm activity](image-url)
nonmeteorological differences among various locations can be observed. Lightning data are gridded to 0.5\degree resolution such that the box areas range from 1300 km\(^2\) at 65\degree latitude to 2500 km\(^2\) at 35\degree latitude, and thus are larger than those that can be inferred using SYNOP. Although lightning networks are devoid of human observation errors, they provide shorter measurement periods, and detection is heavily dependent on the location of sensors (e.g., the lack of data for far eastern Europe). Contrasting these direct observations, sounding and reanalysis data denote physical properties of the air mass and thus it is difficult to simply determine the size of the area for which these estimates are representative, as evidenced by the variety of proximity criteria used in the past (Potvin et al. 2010). In addition, sounding data are also prone to measurement errors and are not homogeneous in time and space since the sensors have changed over time and vary between different countries. Reanalysis, on the other hand, offers data that are continuous in time and space but unfortunately are only an approximation of atmospheric conditions at a given snapshot in time (Thorne and Vose 2010) and are burdened with errors in estimating thermodynamic instability or vertical wind shear (Gensini et al. 2014; Allen and Karoly 2014; Taszarek et al. 2018). In contrast to lightning or SYNOP data where the thunderstorm can be directly identified, the use of sounding and reanalysis data requires using an approximation of environmental conditions supportive of thunderstorm development. This estimation is also prone to errors, as not all conditions supportive to thunderstorm development necessarily produce a thunderstorm, and they can vary among locations. However, direct observations also have limitations and relatively short temporal records, which makes the continuity of sounding and soundings more attractive. Some locations feature full measurement periods of SYNOP and have a good lightning detection efficiency, while in other locations issues with the data are involved. The time difference across Europe is also important, with 1200 UTC soundings representing a favorable part of the diurnal cycle for afternoon thunderstorms in eastern Europe (e.g., 1500 LT in Moscow) but an early preconvective environment for western Europe (3 h earlier in Lisbon). In addition, ESWD has a strong spatial and temporal inhomogeneity and different types of storms can produce heavy convective precipitation, large hail, severe convective wind gusts, and tornadoes. All these factors should be taken to indicate that local results should be interpreted with caution in any single dataset, and that the differences among the various sources may be the consequence of varying structure and quality of the data, rather than the underlying meteorology.

4. Results

a. Thunderstorm days

The distribution of TDs over Europe for all four datasets (surface observations, lightning data, soundings, and reanalysis) shows a high frequency over coastal zones of
the Mediterranean and mountainous regions (Fig. 3). The highest frequencies can be found predominantly along the Italian peninsula, the eastern shores of the Adriatic Sea, and the southern slopes of the Alps. Other regions with more than 35 TDs yr\(^{-1}\) are typically in the foothills of mountainous regions, such as the Pyrenees, the Atlas mountains in northern Africa, the Alps through central Europe, and the Carpathians. The annual frequency of TDs based on lightning detection networks decreases northward, from 25–35 days over southern Europe to fewer than 10 days over Scandinavia. These frequencies are spatially consistent with SYNOP measurements, which show a similar distribution of frequency but slightly lower peak values, perhaps reflecting an underestimation by human observers. Soundings demonstrate a similar pattern, but with some differences. Favorable TD environments are less likely over the Pyrenees and Greece compared to both SYNOP and observed lightning frequencies. In contrast, over Turkey, Ukraine, southwest Russia, Estonia, and Romania, the frequency of favorable thunderstorm environments is considerably higher than observed, reflecting limitations in capturing of convective initiation in soundings and thus leading to overestimates. This difference may also result from the choice of ML parcel, which in some situations may be ineffective in reflecting thunderstorms that arise from elevated convection. Also significant is the time difference, which in the case of eastern Europe is 1500 LT for 1200 UTC soundings, which compared to western Europe (3 h earlier in Portugal) favors soundings with more representative preconvective environment. The outlier in terms of spatial distribution of the four datasets is ERA-Interim. Over regions with peak lightning activity the reanalysis shows similar frequencies. However, the frequency is considerably overestimated over Scandinavia (particularly Norway), the North Sea, the British Isles, northwestern Spain, and over the Atlantic. This bias may be indicative of problems associated with the convective parameterization scheme in the reanalysis in providing an accurate rendition of the initiation of convection over these regions (de Leeuw et al. 2015) and/or the metrics used to derive TDs. It is also possible that simulated convection is too shallow to produce lightning.

Fig. 7. Mean annual number of days with (top left) 2-cm+ hail, (top right) heavy convective precipitation, (bottom left) severe convective wind gust, and (bottom right) tornado reports, based on the records from ESWD.
Nonetheless, the reanalysis does provide data across the entire domain and allows estimates over areas devoid of observational and remotely sensed data.

Focusing on the lightning data as ground truth for TDs over Europe (Fig. 4), we can also explore the monthly spatial distribution of TDs. Over the majority of Europe, October–March is relatively quiescent for thunderstorms, with generally less than 6–7 TDs month\(^{-1}\) over the warmer waters of the Mediterranean and its surrounding coastline. TD activity begins to ramp up into April over the continent, particularly over Italy, the Pyrenees, and Turkey with frequency above 5 TDs yr\(^{-1}\). May delineates the beginning of widespread convective activity, with the region of \(\geq 5\) TDs month\(^{-1}\) extending northward through much of central Europe, and peak frequencies over mountainous areas of the southern Alps and Balkans exceeding 10 TDs month\(^{-1}\). The later parts of May and June –August correspond to the peak of Scandinavian TDs, with frequency generally less than 5 TDs month\(^{-1}\), except during July, when Finland, Sweden, and eastern Norway all exhibit peaks between 5 and 7 days month\(^{-1}\). The highest frequencies over coastal margins tend to peak in June (>10 TDs; e.g., Italy, Turkey), while over the higher mountains this peak is delayed until July (>15 TDs). During August, the Atlas Mountains feature the highest frequency in thunderstorm activity while the frequencies across Europe begin to wane, and are generally below 10 TDs month\(^{-1}\) except over the high peaks of the Alps. As the season winds down, frequencies >5 TDs become confined to the Mediterranean area where peak activity is in September–November.

A 30-day moving average of TDs is used to estimate a day of the year (DOY) with peak thunderstorm activity (a day with the highest value) or alternatively the highest probability for a thunderstorm occurring. Such estimate is computed for each grid point of each dataset (Fig. 5). Results within this parameter point to a good agreement among all datasets. Peak thunderstorm activity occurs during July and August over northern, eastern, and central Europe. In western and southeastern Europe the highest chances for thunderstorms are in May and June. Although the western Iberian Peninsula and eastern Turkey feature peak thunderstorm activity in April and May, over coastal regions of Spain and Portugal this threat shifts to October (some of these locations actually have bimodal distributions, which will be shown in section 4c). The highest probabilities for thunderstorms over western and the central Mediterranean also occur during October and November, and shift toward December and January over the eastern Mediterranean. Although only a few TDs per year occur in the northern Atlantic, the highest chances for thunderstorms over this area occur during wintertime.

\[ b. \textbf{Severe thunderstorm days} \]

The spatial pattern of SevTDs (Fig. 6) in ERA-Interim and sounding data resembles the pattern in TDs (Fig. 3) but with a lower occurrence. The highest activity is found across the Apennine and Balkan Peninsulas, the Alps, and the Carpathians with an average of 15–20 SevTDs yr\(^{-1}\). On the other hand, the highest density of severe weather
reports is found in central Europe (Fig. 6). Although the southwestern part of this area partially overlaps with the maxima in ERA-Interim and sounding data, observed peaks in eastern Germany, the Czech Republic, and Poland do not correspond with environmental estimates. Several factors may explain this difference. The most important is a strong bias of ESWD reports toward central Europe that results from the unequal reporting efficiency and population density (Groenemeijer and Kühne 2014; Groenemeijer et al. 2017). Second, soundings or reanalysis profiles are not always representative, and provide an imperfect estimate of environmental conditions conducive to the occurrence of severe thunderstorms. In many cases thunderstorms appearing in the apparently marginal conditions (low CAPE and low DLS) produced local severe weather events (e.g., heavy precipitation and severe wind gusts), and thus were assigned as SevTD. Although large hail, severe convective wind gusts, and tornadoes are usually associated with enhanced ML WMAXSHEAR (Taszarek et al. 2017), heavy convective precipitation events may occur in a variety of conditions (Púčik et al. 2015), making their occurrence more difficult to estimate. Thus, direct comparisons should be made with caution. It is also worth noting that ERA-Interim (and/or the metrics used to derive TDs) seems to significantly overestimate SevTDs over the northwestern Iberian Peninsula and Norway, similar to how ERA-Interim overestimates of TDs.

Severe weather reports of individual types of events indicate that heavy convective precipitation is the most common type of severe weather across Europe (Fig. 7). This is most frequently reported in Austria and the Czech Republic (>10 days yr⁻¹). Severe convective wind gusts peak in eastern and western Germany (6–10 days yr⁻¹). Large hail reports are most common near the Alpine range (6–8 days yr⁻¹). Tornadoes are characterized by a larger spatial diversity. More than 10 days yr⁻¹ with tornadoes are observed over an eastern part of the Black Sea in the area of the Sochi tourist resort whereas approximately 4–8 days are observed in parts of Italy, Croatia, Greece, and southwestern Turkey.
although the majority of them are waterspouts. ESWD data provide an interesting overview on the observed SevTDs, but since they lack spatial homogeneity in reporting, the obtained results are uncertain.

Similar as for TDs, a 30-day moving average is used to compute the DOY with the highest probability of severe thunderstorm according to ESWD, sounding, and reanalysis datasets (Fig. 8). Although some spatial diversity is observed, peak severe thunderstorm activity over the majority of Europe is biased toward June, July, and August. This threat shifts to October–December near the Mediterranean and British Isles and April/May over far southeastern Europe. ERA-Interim tends to indicate peak activity approximately 1–2 months later compared to ESWD whereas sounding data have better agreement with observations.

c. Annual cycles over individual sites

As an additional validation of the annual cycle characteristics, we explore the distributions at selected locations across Europe. For each site we assign the closest SYNOP and sounding station grid box from reanalysis and lightning data and compute a 30-day moving mean of (severe) TDs. In a few individual cases, if a grid box was within a sharp geographical boundary (e.g., coastal zone, mountain range), we have chosen the adjacent grid, which was considered as more representative (based on an arbitrary decision). This type of analysis has never been performed on a pan-European scale and provides an interesting insight into how the shape of the annual cycle is captured by the respective datasets.

Stations in northern Europe (Sundsvall, Goteborg, Helsinki) are characterized by the weakest (severe) thunderstorm activity beginning in May and ending in October (Fig. 9). The warm waters of Gulf Stream favor thunderstorms in Stavanger in autumn and winter (mostly due to the increase in vertical temperature lapse rates and access to enhanced boundary layer moisture). The number of TDs for all four northern Europe stations is the highest for ERA-Interim (Stavanger is located close to the area where ERA-Interim vastly overestimates the activity; Figs. 3 and 6) and the lowest when SYNOP data are used. The biggest differences among various datasets are evident in the fourth quarter of the year in Stavanger. The mean annual
numbers of TDs (SevTDs) for these four analyzed stations range from 9 to 17 (1 to 4).

In northwestern Europe (London, Schleswig, Paris, De Bilt) the thunderstorm season begins in March or April and ends in November or as late as December (Fig. 10). London features less variability in thunderstorm activity during the year compared to more continental stations of Schleswig, Paris, and De Bilt). The mean annual number of TDs (SevTDs) for analyzed stations ranges from 15 to 22 (4 to 8). There is relatively good agreement in the annual cycle among various datasets in London, Paris, and De Bilt. In Schleswig estimates of TDs by ERA-Interim are slightly higher than other datasets. Estimates based on soundings are noticeably lower than other datasets in both Schleswig and De Bilt.

In central Europe (Fig. 11) thunderstorm activity peaks during summertime with a rapid increase in April and a decrease in October (Stuttgart, Warsaw, Budapest). The mean annual number of TDs (SevTDs) for these stations ranges from 27 to 31 (8 to 13). In Udine, the thunderstorm season lasts until December with a mean of 51 (20) TDs (SevTDs) yr\(^{-1}\)—the highest in Europe. This is due to the collocation of favorable abundant lower-tropospheric moisture supply from the Adriatic Sea and the orographic lift from the nearby Alpine range. Although stations in this domain feature a relatively good agreement in the annual cycle between the datasets, estimates based on the lightning data for Udine are higher than the mean while the reanalysis is slightly lower. Since large horizontal gradients in (severe) TDs are observed in this region, these results may be influenced by the resolution of the reanalysis especially in the areas with strong horizontal gradients (e.g., coastal zones, mountain ranges).

The biggest differences among the databases are evident over eastern Europe (Fig. 12). The highest values of SevTDs are found in soundings and reanalysis while lightning-based estimates are the lowest. The result of the latter can be affected by the decreasing detection efficiency of the lightning network (Kharkov and Moscow are on the edge of the ZEUS domain and outside the EUCLID domain). Some of the differences may be also due to the time-zone difference effect with the 1200 UTC sounding time sampling a favorable convective environment (e.g., 1500 LT in Moscow). All datasets indicate a well-defined peak in

**Fig. 11.** As in Fig. 9, but for the central Europe subdomain.
thunderstorm activity in summertime with a rapid increase in May and decrease in September/October, typical for a continental climate (Kazan, Kharkov, Moscow). The mean annual number of TDs (SevTDs) for this domain ranges from 17 to 29 (4 to 8).

Stations in southeastern Europe share similar features to those in central and eastern Europe. Belgrade, Bucharest, and Sofia have a well-defined peaks in thunderstorm activity in summertime with a rapid increase in April/May and decrease in October (Fig. 13). The mean annual number of TDs (SevTDs) for these stations ranges from 32 to 44 (8 to 13), suggesting that this area has among the highest thunderstorm activity in Europe. A quite different pattern in the annual cycle is represented by the neighboring Athens with a rather constant thunderstorm activity throughout the year. With high ambient moist conditions throughout the year, Athens is characterized by a weak annual variability of small to moderate instability compared to Bucharest, featuring a well-defined peak in July and minimum in January (Taszarek et al. 2018). This indicates a different climate regime for Athens compared to other stations located farther north.

Stations in south-central Europe feature a similar pattern to Athens (thunderstorms occur all year round) but with greater activity, driven by the warm waters of the Mediterranean Sea and more frequent convective initiation (Fig. 14). The mean annual number of TDs (SevTDs) for these stations ranges from 34 to 37 (10 to 14), and the annual cycle is generally consistent between the datasets. The peak in thunderstorm activity is observed in the late spring. The annual minimum is found between December and January except for Trapani, which has a minimum in July. A periodic decline in thunderstorm activity between July and August is also observed over Ajaccio, Zadar, and Brindisi. This may be related to the frequent presence of a high pressure system (the ridge of the Azores high) during this period inducing large-scale subsidence. However, convection in Zadar driven by the warm waters of Adriatic Sea and complex orography displays a less pronounced minimum in summertime, and is less influenced by the ridge.

A much greater influence of the Azores high inhibiting summertime thunderstorm activity is evident in Lisbon, Gibraltar, and, to a lesser extent, Madrid (Fig. 15).

Fig. 12. As in Fig. 9, but for the eastern Europe subdomain.
This leads to a bimodal structure of the annual cycle in both spring and autumn. The largest difference between the datasets in representing this pattern is found over Lisbon, where reanalysis is considerably higher. The mean annual number of TDs (SevTDs) for stations in this domain ranges from 18 to 24 (5 to 9). Although estimates for Madrid are not consistent among datasets, the mean distribution indicates a peak in May and June (mostly driven by the summertime thermal convection) and reduced activity in July and August (most likely due to the influence of the Azores high). The annual cycle for Palma de Mallorca, on the other hand, resembles a cycle typical for the central Mediterranean, with a peak in autumn and decreased activity throughout the remainder of the year.

d. Multiannual changes in the mean cycle

Since ERA-Interim provides data continuous in time and space, changes in frequency over the past 39 years for (severe) TDs are explored. For this purpose, a mean computed for 1998–2017 is compared with a mean from 1979 to 1997 (Fig. 16). We do not apply this method to the other datasets, since none of them ensures sufficient data continuity or homogeneity during the considered period. However, we are also aware that the reanalysis is less than ideal for this application, as its purpose is to produce the best analysis of any given day or time step, rather than to necessarily provide consistent trends (Thorne and Vose 2010). Since ERA-Interim (and/or the metrics used to derive TDs) has been shown to considerably overestimate TDs over Scandinavia (particularly Norway), the North Sea, the British Isles, northwestern Spain, and over the Atlantic (Fig. 3), multiannual changes obtained for these areas may also be not reliable and should be interpreted with caution (hatched area in Fig. 16).

The analysis suggests that an increase of more than 5–10 TDs yr\(^{-1}\) has taken place over central, southeastern, and eastern Europe. The highest increase of more than 15 TDs was recorded over the Balkan Peninsula. Conversely, a small decrease is observed over parts of southwestern, south-central, and far southeastern Europe. Although the distribution of the changes in the mean annual number of SevTD shows a similar spatial pattern, the highest increase in severe thunderstorm potential is observed over southeastern Europe.
northern Italy, the Alps, and the Balkan Peninsula, mainly western Greece and Bulgaria. These results partially confirm earlier studies (Pistotnik et al. 2016; Püćik et al. 2017; Rädler et al. 2018).

Changes in the mean annual cycle pattern over station locations used in section 4c were also explored (Fig. 17). The biggest increase exceeding 10 TDs is observed in De Bilt, Athens, Udine, and Goteborg (11.1, 11.6, 13.3, and 14.4, respectively). However, Goteborg is located in the area where ERA-Interim (and/or the metrics used to derive TDs) overestimates TDs, and thus the obtained value is likely to be overestimated as well. In Bucharest an increase of 6.8 TDs consists of a significant majority of SevTDs (6.6). For Stuttgart the mean number of TDs is decreasing with a simultaneous increase in the percentage of SevTDs. The most considerable decrease in the mean TDs is found in Ajaccio (4.1), although the majority of locations are characterized by an increase of TDs. Although most of the changes concern an increase in the absolute values of (SevTDs) TDs, some stations experience changes in the pattern of the cycle. Clear examples are De Bilt (peak activity shifted from July to August), Moscow (an increase exclusively in July), Madrid (a decrease in the summertime), and Athens (a sharp increase in Autumn and peak activity shifted from December to September). However, since the studied period is relatively short, these changes can be also related to decadal variations and cycles in thunderstorm activity rather than long-term trends.

5. Comparison with previous studies

Many of our results expand on what has been obtained in previous analyses based on a single dataset. Studies using SYNOP reports indicate that the average annual number of TDs increases from 15–20 in the Baltic countries up to 30–35 toward southern Europe and the Carpathians (Bielec-Bąkowska 2003; Wapler 2013; Enno et al. 2013; Kolendowicz et al. 2017). Analyses considering TDs within lightning detection systems demonstrate a similar pattern but with slightly higher values (Holt et al. 2001; Novák and Kyznarová 2011; Mäkelä et al. 2014; Taszarek et al. 2015). More than 30 TDs yr⁻¹ occur also in northeastern Spain, Italy, and the western Balkan Peninsula.
according to the estimates by Galanaki et al. (2018). Estimates based on environmental conditions supporting thunderstorms provide comparable patterns. Püčik et al. (2017) using EURO-CORDEX data (Coordinated Downscaling Experiment–European Domain; Jacob et al. 2014) found local maxima of severe thunderstorm environments over southern France and northeastern Spain. A peak frequency of TDs in ERA-Interim was found in Italy, the Alps, the Carpathians, and the Balkan Peninsula (Groenemeijer et al. 2017; Taszarek et al. 2018; Rädler et al. 2018). Multiple studies also pointed out that the highest thunderstorm activity is over the Alps during summer and over the central Mediterranean during winter (Holt et al. 2001; Anderson and Klugmann 2014; Galanaki et al. 2015; Kotroni and Lagouvardos 2016; Taszarek et al. 2018; Galanaki et al. 2018). An increase in (severe) thunderstorm environments in the Alps and the Balkan Peninsula and a decrease in the Iberian Peninsula within the last decades was also found by Rädler et al. (2018). Our findings are consistent with the aforementioned studies, but provide a previously undocumented insight by combining increased sample size and the strength of multiple datasets with which to cross-validate the results.

6. Summary and concluding remarks

In this paper sounding measurements, surface observations, ERA-Interim, EUCLID and ZEUS lightning data, and severe weather reports were compared. A large sample size from several complementary datasets allowed improved insight into the spatial and temporal distributions of (severe) thunderstorms across Europe. In addition, the changes in the frequency of thunderstorm environments over the last 39 years were also studied. This type of analysis has never been performed on a European scale and provides an interesting comparison of how the annual cycles of thunderstorms are captured by the respective datasets. Below, the most important findings are listed:

- Proximity analysis performed with lightning and severe weather events for both reanalysis and sounding data suggests that thunderstorms are likely to occur if ML CAPE exceeds 150 J kg\(^{-1}\), and are more susceptible of becoming severe if ML WMAXSHEAR exceeds 400 m\(^2\) s\(^{-2}\).
- The distribution of TDs over Europe for all four datasets shows large frequencies of thunderstorm days over coastal zones of the Mediterranean and mountainous

![Southwestern Europe](image_url)
regions. Peak observed frequencies can be found predominantly along the Italian peninsula, the eastern shores of the Adriatic Sea, and the southern slopes of the Alps. A similar pattern is found for SevTDs except for severe weather reports, which indicate maxima over central Europe, primarily due to spatial inhomogeneities in reporting.

- Compared with other datasets the frequency of TDs in ERA-Interim is considerably overestimated over Scandinavia (particularly Norway), the North Sea, the British Isles, northwestern Spain, and over the Atlantic. This bias may be indicative of problems associated with the convective parameterization scheme in the reanalysis in providing an accurate rendition of the initiation of convection over these regions and/or the metrics used to derive TDs. It is also possible that simulated convection over these areas is too shallow to produce lightning.

- Annual peak thunderstorm activity occurs in July and August over northern, eastern, and central Europe. Over western and southeastern Europe thunderstorms are the most frequent in May and June whereas over the western Iberian Peninsula and eastern Turkey they are most frequent in April and May. Western and central parts of the Mediterranean have predominant thunderstorm activity in October and November whereas over the eastern part the highest threat shifts to December and January.

- According to estimates within ERA-Interim, an increase of more than 5–10 TDs yr\(^{-1}\) has taken place over the Alps and central, southeastern, and eastern Europe. Conversely, a small decrease was observed over parts of southwestern, south central, and far southeastern Europe. Although the distribution for SevTDs has a similar spatial pattern, the highest increase in severe thunderstorm potential is observed over northern Italy with the Alps and over the Balkan Peninsula.

- Multiannual changes reflect mostly an increase in the absolute values of (severe) TDs, but some stations also show changes in the pattern of the cycle. The most clear examples are De Bilt (peak activity shifted from July to August), Moscow (an increase exclusively in July), Madrid (a decrease in the summertime), and Athens (a sharp increase in autumn and peak activity shifted from December to September). However, since the studied period is relatively short, these changes can be also related to decadal variations and cycles in thunderstorm activity rather than long-term trends.

Although most of our conclusions are not qualitatively surprising, we have developed quantitative assessments of these characteristics, leading to a more precise estimates of thunderstorm frequencies in Europe. Analyses of 30-day moving means revealed that annual cycles of (severe) TDs may significantly differ even within small distances (e.g., Athens, Brindisi, Sofia). Obtained results may be used as a background for future studies on thunderstorm occurrence in Europe, and also can be a valuable source of information for various groups such as weather forecasters and insurance companies. Comparison of different data sources revealed that although lightning detection data seem to
FIG. 17. Multiannual changes in the mean annual cycles of days with thunderstorms (black and gray) and severe thunderstorms (red and light red), computed as a difference between mean annual values in 1979–97 (gray and light red) and 1998–2017 (black and red), based on ERA-Interim. The definitions of TDs and SevTDs are provided in Table 1.

Comparison of mean annual cycles between 1979–1997 and 1998–2017 (ERA-Interim)
sample thunderstorm activity the most objectively, short
operating periods and areas devoid of sensors remain a
challenge that limits their use. In contrast, reanalysis com-
plements these disadvantages but is prone to errors related
to modeling thunderstorm occurrence and numerical sim-
ulation itself. Compared with the other datasets, reanalysis
has the biggest problems in sampling thunderstorm activity
over coastal zones of Atlantic Ocean. This suggests that an
important step in any climatological analysis over data-
sparse regions should ideally cross-validate results between
the different sources, or combine them to leverage their
respective strengths (e.g., Rädler et al. 2018). Future studies
will continue to address this topic and improve our un-
derstanding of the differences between reanalysis and ob-
servational data. These steps are critical in order to better
model (severe) thunderstorm environments and estimate
their climatological aspects on the areas devoid of or with
limited observational data. This is particularly important in
the face of changing thunderstorm frequency as a conse-
quence of warming climate (Brooks 2013; Allen 2018;
Allen, J. T., 2018: Climate change and severe thunderstorms.
In Oxford Research Encyclopedia of Climate Science, Oxford University
——, and D. J. Karoly, 2014: A climatology of Australian severe
thunderstorm environments 1979–2011: Inter-annual vari-
doi.org/10.1002/joc.3667.
——, and E. R. Allen, 2016: A review of severe thunderstorms in
10.1016/j.atmosres.2016.03.011.
——, D. J. Karoly, and G. A. Mills, 2011: A severe thunderstorm
climatology for Australia and associated thunderstorm envi-
10.22499/2.6103.001.
——, ——, and K. J. Walsh, 2014: Future Australian severe thun-
derstorm environments. Part II: The influence of a strongly
warming climate on convective environments. J. Climate, 27,
——, M. Tippett, and A. Sobel, 2015: An empirical model relating
United States monthly hail occurrence to large-scale meteor-
https://doi.org/10.1002/2014MS000397.
Anderson, G., and D. Klugmann, 2014: A European lightning den-
sity analysis using 5 years of ATDnet data. Nat. Hazards Earth.
Betz, H. D., K. Schmidt, P. Laroche, P. Blanchet, W. P. Oertinger,
E. Defer, Z. Dziewit, and J. Konarski, 2009: LINET—An
international lightning detection network in Europe. Atmos. Res.,
Bielec-Bąkowska, Z., 2003: Long-term variability of thunderstorm
Brooks, H. E., 2009: Proximity soundings for severe convection
for Europe and the United States from reanalysis data. Atmos. Res.,
——, 2013: Severe thunderstorms and climate change. Atmos. Res.,
——, C. A. Doswell III, and J. Cooper, 1994: On the environments
of tornadic and nontornadic mesocyclones. Wea. Forecasting, 9,
606–618, https://doi.org/10.1175/1520-0434(1994)009<0606:
OTEOTA>2.0.CO;2.
——, J. W. Lee, and J. P. Craven, 2003: The spatial distribution of
severe thunderstorms and tornado environments from global
10.1016/S0169-8095(03)00045-0.
Cecil, D. J., D. E. Buechler, and R. J. Blakeslee, 2014: Gridded
lightning climatology from TRMM-LIS and OTD: Dataset
10.1016/j.atmosres.2012.06.028.
——, ——, and ——, 2015: TRMM LIS climatology of thunderstorm
occurrence and conditional lightning flash rates. J. Climate,
Celinski-Myslaw, D., and A. Palzar, 2017: The occurrence of convective
systems with a bow echo in warm season in Poland. Atmos. Res.,
Changnon, S. A., and D. Changnon, 2001: Long-term fluctuations in
thunderstorm activity in the United States. Climatic Change,
Christian, H. J., and Coauthors, 2003: Global frequency and distri-
bution of lightning as observed from space by the Optical
10.1029/2002JD002347.
Cintineo, J. L., T. M. Smith, V. Lakshmanan, H. E. Brooks, and
K. L. Ortega, 2012: An objective high-resolution hail clima-
tology for Australia and associated thunderstorm envi-
10.22499/2.6103.001.
Cohen, A. E., M. C. Coniglio, S. F. Corfidi, and S. J. Corfidi, 2007:
Discrimination of mesoscale convective system environments
using sounding observations. Wea. Forecasting, 22, 1045–1062,
https://doi.org/10.1175/WAF1040.1.
Craven, J. P., and H. E. Brooks, 2004: Baseline climatology of
lightning-derived parameters associated with deep moist

Acknowledgments. This research was supported by a grant
from the Polish National Science Centre (project
number 2017/27/B/ST10/00297). The reanalysis and
sounding computations were performed in the Poznań
Supercomputing and Networking Center (Grant 331). We
thank the EUCLID and ZEUS (http://thunderstorm24.
com/) networks for kindly providing the lightning data.
Püčik’s and Groenemeijer’s contributions were supported by
Grant 01LP1525A (ARCS) by the Ministry of Edu-
cation and Research of Germany.

REFERENCES
Allen, J. T., 2018: Climate change and severe thunderstorms. Oxford
Research Encyclopedia of Climate Science, Oxford University
——, and D. J. Karoly, 2014: A climatology of Australian severe
thunderstorm environments 1979–2011: Inter-annual vari-
doi.org/10.1002/joc.3667.
——, and E. R. Allen, 2016: A review of severe thunderstorms in
10.1016/j.atmosres.2016.03.011.


McCaul, E. W., 1991: Buoyancy and shear characteristics of Pohjola, H., and A. M.


