https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

Thunderstorm Types in Europe

Deborah Morgenstern[1, 2], Isabell Stucke[1, 2], Georg J. Mayr[1], Achim Zeileis[2], and Thorsten Simon[2, 3]

[1]Department of Atmospheric and Cryospheric Sciences (ACINN), University of Innsbruck, Innsbruck, Austria
[2]Department of Statistics, University of Innsbruck, Innsbruck, Austria
[3]Department of Mathematics, University of Innsbruck, Innsbruck, Austria

Correspondence: Deborah Morgenstern (deborah.morgenstern@uibk.ac.at)

Abstract. Lightning characteristics in all seasons are investigated across Europe because it is observed that lightning strikes to tall infrastructure have no or only a weak annual cycle whereas lightning in general has a pronounced annual cycle. Using cluster analysis on ERA5 reanalysis data and EUCLID lightning data, two major thunderstorm types are found: Wind-field thunderstorms characterized by increased wind speeds, strong updrafts, and high shear occurring mainly in winter. And mass-field thunderstorms characterized by increased mass-field variables such as large CAPE values, high dewpoint temperatures, and elevated isotherm heights, occurring mostly in summer. Several sub-types of these two main thunderstorm types exist. Using principal component analysis, four topographically distinct regions in Europe are identified that share similar thunderstorm characteristics: The mediterranean, alpine-central, continental, and coastal regions, respectively. Based on these results it is possible to differentiate lightning in different seasons without a static threshold or a seasonal criterion.

1 Introduction

Lightning may originate in various meteorological settings. Some conditions leading to lightning occur more frequently and are better understood than others. Plenty lightning climatologies are available but most of them focus on the dominant characteristics and seasons while infrequent thunderstorm types and rare lightning conditions are often neglected. This study describes thunderstorm types occurring in Europe using a balanced view on all four seasons to include also seasonally infrequent thunderstorm conditions. A comparison between different regions gives a comprehensive overview of lightning characteristics in Europe.

The general lightning pattern in Europe is well described in various climatologies (e.g., Taszarek et al., 2019; Enno et al., 2020; Poelman et al., 2016; Wapler, 2013; Taszarek et al., 2020a, b; Mäkelä et al., 2014; Vogel et al., 2016; Ukkonen and Mäkelä, 2019; Simon et al., 2017; Kotroni and Lagouvardos, 2016; Piper and Kunz, 2017; Anderson and Klugmann, 2014; Hayward et al., 2022; Holt et al., 2001; Enno et al., 2013; Poelman, 2014; Taszarek et al., 2015; Schulz et al., 2005; Coquillat et al., 2022; Manzato et al., 2022; Simon and Mayr, 2022). There is a north-south gradient of lightning frequency with a maximum in northern Italy and the Mediterranean. Lightning in central Europe follows a clear annual cycle with a maximum over land in summer (MJJA) and a secondary maximum in fall and early winter (SONDJ) in the Mediterranean (Taszarek et al., 2019; Poelman et al., 2016; Enno et al., 2020). In south-central Europe, the annual cycle is less pronounced and has sometimes two lightning maxima along with a local minimum in summer (Taszarek et al., 2019). There are differences in the annual

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

lightning cycle: Offshore and coastal areas have a lower amplitude and a later maximum compared to inland or mountainous locations (Wapler, 2013; Enno et al., 2013). In the northern Atlantic region, occasional very intense thunderstorms are possible, even though the climatological thunderstorm activity is very low (Enno et al., 2020). There, lightning in the cold season (Oct.-Apr.) occurs predominantly over the seas (North Sea, Baltic Sea, Atlantic) and less so over the land (Mäkelä et al., 2014). The question remains if there are meteorologically different thunderstorm conditions at work leading to these spatial and temporal differences in lightning characteristics in Europe.

Many processes influencing lightning occurrence are known: The diurnal lightning cycle in Europe peaks in the afternoon over land and at night over the sea (Taszarek et al., 2020a; Enno et al., 2020; Manzato et al., 2022). Nighttime offshore lightning (Bay of Biscay, the North Sea, and the Baltic Sea) is explained by convection initiated at land and advected to the sea, where it endures longer as the water surface is unaffected by nighttime cooling (Enno et al., 2020). The most pronounced diurnal cycle is found over mountainous areas. Increased lightning in mountainous regions is commonly explained by the topography, which favors more unstable environments (less CIN when the surface is close to the level of free convection), mechanical forcing (forced lifting), and thermal forcing (elevated heating leads to positive buoyancy and up-mountain flow; Manzato et al., 2022). This is particularly relevant after the snow has melted at higher elevations (Simon and Mayr, 2022). Most of continental Europe experiences $20-40$ thunderstorm days annually, but the mountain ranges in southern Europe have thunderstorm frequencies of > 60 thunderstorm days per year. High structures, such as wind turbines or radio towers increase the occurrence of lightning (March et al., 2016), especially in the cold season (Vogel et al., 2016; Pineda et al., 2018) so that lightning damage to infrastructure is evenly distributed over the year even though lightning occurrence in the surroundings has a strong annual cycle (Stucke et al., 2022). Helicopter-triggered lightning in the North Sea occurs very often in the cold season (Wilkinson et al., 2013). The sea has an effect on lightning as the number of lightning strokes and the sea-surface temperature are positively correlated in fall (Kotroni and Lagouvardos, 2016). Mallick et al. (2022) even suggests to use the sea surface temperature as proxy for seasonal lightning forecasts. Warm oceanic currents are known to increase lightning densities in each season and particularly so in winter (Iwasaki, 2014; Holle, 2016). Wintertime lightning occurs usually in mid-latitudinal cyclones (Bentley et al., 2019) and lightning bands are found in wintertime storm tracks (Zhang et al., 2018; Virts et al., 2013). Even though the European lightning patterns are in general well described (e.g., Wapler and James, 2015; Enno et al., 2014), it is less known how winter lightning differs meteorologically from lightning in summer.

Morgenstern et al. (2022a) found that thunderstorms in the cold season differ physically from thunderstorms in the warm season when studying a small observational region in northern Germany. They describe wind-field thunderstorms dominant in winter in contrast to CAPE (mass-field) thunderstorms typical for summer. This study expands their findings to all of Europe answering two questions:

- How do meteorological thunderstorm characteristics vary regionally across Europe?
- What characterizes thunderstorms in different meteorological environments and how do these thunderstorm types vary seasonally across Europe?

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

To answer these questions, Europe is first divided into 12 domains (Sect. 2) for which a large set of meteorological parameters relevant for lightning is available. Applying similar methods as in Morgenstern et al. (2022a), principal component analysis finds the answer to the first question, and k-means clustering to the second one (Sect. 3). The general thunderstorm conditions in Europe are compared in Sect. 4.1 revealing that the 12 domains can be summarized into four regions with similar lightning characteristics. Cluster analysis on each domain in Sect. 4.2 leads to two main thunderstorm types and three variations thereof. The found thunderstorm types are then analyzed seasonally (Sect. 4.3) and compared to one another (Sect. 4.4). The results are discussed in Sect. 5, and Sect. 6 summarizes the main findings.

2 Data

Two data sets are incorporated in this study (cf. Morgenstern et al., 2022a): meteorological reanalysis data (ERA5, Sect. 2.1) and lightning observations (EUCLID, Sect. 2.2).

2.1 Meteorological data: ERA5

Directly available and derived meteorological data are extracted from single-level and model-level data of the ERA5 global reanalysis of ECMWF (Hersbach et al., 2020). The distance between vertical model levels varies from 10 m near the ground to 320 m in the lower stratosphere (lowest 74 levels). The horizontal resolution is $0.25\,^\circ$ latitude–longitude and the temporal resolution is one hour. A binary land-sea mask sets ERA5 grid cells with at least 35 % land to land to capture the influence of sub-grid islands. A set of 25 variables is chosen to portray physical atmospheric processes known to influence lightning, especially charge separation. It, therefore, contains derived variables such as the height of the $-10\,^\circ$C isotherm, cloud mass between -10 and $-40\,^\circ$C (ice and snow), and the product of maximum updraft and liquid particles between -8 and $-12\,^\circ$C (*liquids updrafts*). Further derived variables are cloud size, cloud shear, wind speed at cloud base, maximum upward vertical velocity, and the temperature difference between the air mass at 1000 m a.g.l. and the surface (sea surface temperature or skin temperature). The variable set focuses on the most important drivers to reduce correlation between the variables which is required for the statistical analysis. All 25 variables are listed in Table 1 and details on them and on how the derived variables are calculated are given in the online supplement (Morgenstern et al., 2022b). To ease interpretation, we often refer to physical-based categories to which each variable belongs: *Mass-field variables* refer to temperature, pressure, and humidity. *Surface-exchange variables* include atmospheric fluxes interacting with the surface. *Wind-field variables* cover everything related to wind. *Cloud-physics variables* refer to measures directly related to clouds. *Topographic variables* refer to surface geopotential height (orography) and a binary land-sea mask.

2.2 EUCLID lightning data and geographical domains

As lightning data EUCLID (European Cooperation for Lightning Detection, Schulz et al., 2016; Poelman et al., 2016) between 2010–2020 is used as this period is most stable regarding hardware and software changes for this data set. EUCLID is a cooperation of several local lightning location systems (LLS) in Europe. Only cloud-to-ground lightning flashes are considered

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

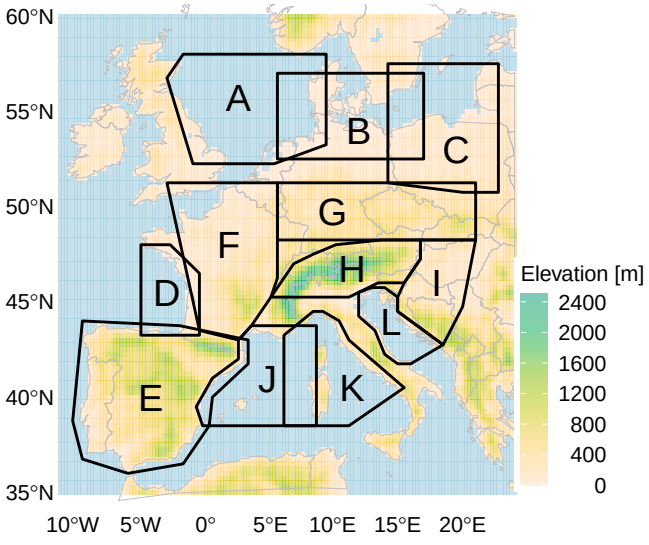

Figure 1. Overview of domains. Topographic data is based on ERA5 orography with a resolution of $0.25\,^{\circ}$ x $0.25\,^{\circ}$ lon/lat. A land-sea mask is applied where each cell containing $> 35\,\%$ land is regarded as land.

in this study. If at least one lightning flash occurred within an ERA5 cell in a given hour, the whole cell-hour is regarded as one lightning observation.

The EUCLID territory is separated into 12 domains with rather homogeneous topography and lightning detection efficiency (Fig. 1) aiming to represent typical European landscapes. Domain A covers large parts of the North Sea including the surrounding coastlines. It marks the furthest northern EUCLID domain with sufficient lightning detection efficiency and sufficient lightning observations in each season and year. Domain B covers large parts of Denmark and northern Germany as well as parts of the North Sea and the Baltic Sea. It also covers parts of the domain analyzed in Morgenstern et al. (2022a) to which the current study is an extension. Domain C is representative of the southern Baltic Sea and Poland, except for the Carpathian mountains in the South. Domain D covers the Gulf of Biscay, an Atlantic domain. Domain E covers the whole Iberian Peninsula including the Pyrenees and the surrounding coastal areas and is characterized by highlands. Domain F covers large parts of France and Belgium, being a less homogeneous but very representative domain. Domain G covers hills in Germany, Check Republic, Southern Poland, and Slovakia. Domain H covers the east-west elongated part of the European Alps. Domain I covers Hungary, Croatia, and Bosnia and Herzegovina, a basin surrounded by several mountain ranges. Domain J covers the Balearic islands in the northwestern Mediterranean Sea and surrounding coastlines. Domain K covers the Tyrrhenian Sea and the Islands of Corsica and Sardinia and Italian coastal areas. Finally, domain L covers the northern part of the Adriatic Sea including the surrounding coastlines.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

3 Methods

To investigate spatio-temporal lightning characteristics, lightning data sets for the 12 domains are constructed that have the same number of observations from each season. The lightning data sets are then combined with 25 ERA5 variables representing the atmospheric conditions at the hour of the lightning observations. Using the domain means, a *spatial* lightning analysis for Europe is performed with the help of a principal component analysis. Then, thunderstorm *types* are found individually on each domain by a cluster analysis with $k = 3$ clusters. A *seasonal* lightning analysis follows by analyzing how many observations from each season have been classified into which thunderstorm type. Finally, the thunderstorm types are compared to one another using again a principal component analysis.

3.1 Composition of data

EUCLID lightning data is aggregated to the spatio-temporal resolution of ERA5 resulting in binary cell-hours indicative of lightning. For each lightning cell-hour, we use ERA5 data at the respective cell and from the last full hour to capture the build-up of the thunderstorms. Accumulated variables such as precipitation are taken from the next full hour to capture everything from the hour in which lightning was observed. Only cell-hours with lightning are considered. To investigate seasonal differences, the available data are reduced to contain the same number of lightning cell-hours from each season (winter = DJF, spring = MAM, summer = JJA, fall = SON). Therefore a random sample without replacement is drawn from the seasons with more lightning cell-hours. Depending on the domain size and general lightning frequency, the data set in each domain consist of $5\,320 - 40\,000$ observations (Table 1). For robustness, the whole analysis is performed on 50 different samples in each domain. In the following, only one sample is discussed, as all repetitions led to qualitatively the same results.

k-means clustering requires scaled input variables that follow rather similar distributions. Therefore all ERA5 variables are square root transformed and scaled to a mean of zero and a standard deviation of one.

$$x_\mathrm{t} = \mathrm{sign}(x)\sqrt{\mathrm{abs}(x)}, \tag{1}$$

with x being the original ERA5 value and x_t its transformation.

$$x_\mathrm{s} = \frac{(x_\mathrm{t} - \mu)}{\sigma}, \tag{2}$$

μ and σ are the empirical mean and standard deviation and x_s is the scaled value. The applied algorithm is supplied in the online supplement (Morgenstern et al., 2022b). For the cluster analysis in Sect. 4.2, transformation and scaling are performed individually on each domain. For the domain comparison in Sect. 4.1 (Sect. 4.4), scaling is performed on the domain means (cluster means) of all domains together.

3.2 Statistical methods

Principal component analysis (Mardia et al., 1995) is an approach for dimension reduction that computes several linear combinations of projected input data (principal components, PC) aiming to capture as much variability from the data as possible.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

The first PC explains the most variance and each following PC is oriented perpendicular to the previous PC explaining less and less variance. Omitting the later PC results in the intended dimension reduction. We use the first two PC as axes for a so-called biplot to visualize the variance in our 25-dimensional data.

k-means cluster analysis (MacQueen, 1967) is a data-driven approach to find groups in data, aiming at maximum similarity within and minimum similarity between the groups. The similarity is measured with the squared euclidean distance between each observation and cluster means. Starting with k random cluster means, new cluster means are calculated iteratively to which the observations are assigned forming the clusters. The optimal number of clusters k for our data, derived from the sum of squared residuals, is between 2 and 4. We present the results from $k = 3$ in detail and also describe the results for $k = 2$ and ≥ 4. Cluster analysis is used to identify different thunderstorm types. To account for possible regional differences, clustering is performed separately on each of the 12 topographically homogeneous domains.

The online supplement provides the R code to replicate the cluster analysis and the principal component analysis (Morgenstern et al., 2022b).

4 Results

Thunderstorms in 12 domains in Europe are compared to one another using principal component analysis (PCA) investigating which domains have in general similar meteorological characteristics during lightning throughout the year (Sect. 4.1). Then we present the thunderstorm types found by k-means clustering and a decision tree to differentiate them (Sect. 4.2). Finally, the thunderstorm types are seasonally analyzed and compared in Sections 4.3 and 4.4.

4.1 Regional differences between thunderstorms in Europe

This section investigates how the meteorological conditions vary regionally in Europe and whether some of our 12 domains can be grouped together based on their meteorological similarities during lightning throughout the year.

Table 1 presents the meteorological mean values of 25 variables separately for each domain. These are typical values for thunderstorms throughout the year for the respective domains and are considered as 'baselines' there. A principal component analysis (Fig. 2) makes it easier to spot differences and commonalities between these domains. The first two principal components (x-axis and y-axis) explain together about 80 % of the variance within the data. The further the domains (colored symbols) are from the origin, the larger their contribution to the variance in the respective direction. Meteorological similar domains gather close to one another. The loadings (labeled arrows) indicate the direction and strength of individual meteorological variables responsible for the variation in the respective direction.

The domains A, B, C, and D (blue triangles) are all located in the top-left of Fig. 2. The labeled arrows indicate that these domains are physically characterized by increased boundary layer heights (~ 800 m) and increased wind speeds at 10 m (~ 6 m s^{-1}) and at cloud base (~ 12 m s^{-1}) relative to all other domains (Table 1). Long arrows pointing in opposite directions of domains A–D indicate decreased values. For example, decreased values in CAPE, CIN, pressure, and cloud size. The temperature difference between the ocean (or skin temperature over land) and the air at 1000 m altitude is on average 6.4 °C

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

Table 1. Meteorological mean values during lightning for each domain. Domains with similar characteristics are grouped into four regions, for which a mean value is also given. ERA5 variables are grouped by their meteorological category.

	CAPE	CIN presence *	Pressure (msl)	Water vapor (total column)	Temperature dewpoint (2 m)	-10 C isotherm height *	Temp. diff. sfc (sst or skt) - 1000 m a.g.l. *	Boundary layer height	Solar radiation (sfc, net, down)	Sensible heat flux (sfc, up)	Latent heat flux (sfc, up)	Wind speed at 10 m	Wind speed at cloud base *	Cloud shear *	Boundary layer dissipation	Vertical velocity (maximum, up) * △	Updrafts liquids -8 C -12 C *	Ice crystal mass -10 C -40 C *	Snow mass -10 C -40 C *	Supercooled liquids (total column)	Convective precipitation △	Large-scale precipitation △	Cloud size *	Land-sea mask *	Elevation	Nr. of observations
Unit	J kg^{-1}	binary	hPa	kg m^{-2}	°C	m a.g.l.	K	m	W m^{-2}	W m^{-2}	W m^{-2}	m s^{-1}	m s^{-1}	m s^{-1}	W m^{-2}	Pa s^{-1}	g Pa s^{-1}	g m^{-2}	g m^{-2}	g m^{-2}	mm	mm	m	binary	m	nr.
Category	Mass field							Surface exchange				Wind speed					Cloud physics							Topography		
Coastal region																										
A	178	0.37	1004.4	20.5	9.4	3584	6.3	820	113	10	77	7.9	13.3	14.4	7	0.678	7	66	93	40	0.052	0.012	6791	0.39	7	30792
B	228	0.43	1005.8	21.7	9.9	3706	6.3	843	131	7	82	6.6	12.6	14.3	8	0.658	7	65	93	41	0.052	0.011	6994	0.67	27	27716
C	275	0.50	1007.5	23.4	10.6	3898	6.1	820	139	3	81	5.4	11.6	13.8	8	0.630	5	58	83	41	0.051	0.011	7010	0.83	92	5320
D	250	0.43	1008.3	23.6	11.7	3965	6.9	760	114	20	96	6.5	11.5	16.4	6	0.902	9	68	96	41	0.063	0.013	7072	0.49	67	40000
Mean	233	0.43	1006.5	22.3	10.4	3788	6.4	811	124	10	84	6.6	12.3	14.7	7	0.717	7	64	91	41	0.055	0.012	6967	0.60	48	
Continental region																										
E	299	0.52	1011.9	22.1	11.4	3927	8.9	861	185	49	90	4.2	8.2	17.6	6	0.941	11	64	94	52	0.065	0.018	7678	0.81	529	40000
F	284	0.50	1008.3	23.0	11.3	3918	6.8	805	148	15	87	4.5	10.4	17.1	7	0.841	11	72	112	48	0.062	0.018	7556	0.93	254	40000
G	300	0.53	1009.9	22.7	10.4	3851	6.9	839	168	13	91	3.9	9.7	16.0	9	0.723	9	68	105	45	0.055	0.015	7464	1.00	385	27016
Mean	294	0.52	1010.0	22.6	11.0	3899	7.5	835	167	26	89	4.2	9.4	16.9	7	0.835	10	68	104	48	0.061	0.017	7566	0.91	389	
Alpine-central region																										
H	315	0.53	1011.8	19.5	9.3	3538	5.9	472	135	18	72	1.9	5.1	21.6	10	1.223	45	85	160	90	0.066	0.047	8144	1.00	891	17040
I	352	0.61	1010.1	24.0	11.7	4043	6.3	602	155	22	80	2.4	6.4	18.8	8	1.006	25	79	141	67	0.071	0.032	8179	1.00	461	39224
Mean	334	0.57	1011.0	21.8	10.5	3791	6.1	537	145	20	76	2.2	5.8	20.2	9	1.115	35	82	151	79	0.069	0.040	8162	1.00	676	
Mediterranean region																										
J	436	0.57	1010.5	25.0	13.4	4178	8.7	699	134	27	106	6.3	9.6	16.3	3	0.811	9	65	101	39	0.060	0.018	7280	0.21	55	40000
K	495	0.60	1009.7	24.2	13.3	4111	8.9	685	151	30	105	5.5	8.4	15.4	3	0.840	8	65	93	40	0.067	0.016	7396	0.34	117	40000
L	524	0.63	1009.0	24.6	13.1	4096	8.4	576	133	23	91	4.6	7.5	18.1	3	0.984	14	75	113	49	0.074	0.022	7983	0.45	124	40000
Mean	485	0.60	1009.7	24.6	13.3	4128	8.7	653	139	27	101	5.5	8.5	16.6	3	0.878	10	68	102	43	0.067	0.019	7553	0.33	99	

* Derived variables.

△ Accumulated variables considered at the next full hour after lightning observation. All other variables are considered at the last full hour.

CAPE: convective available potential energy, CIN: convective inhibition, msl: mean sea level, a.g.l.: above ground level,

Temp. diff. sfc (sst or skt) - 1000 m a.g.l.: Temperature difference between the surface and 1000 m a.g.l., where sfc is either sea surface temperature or skin temperature.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

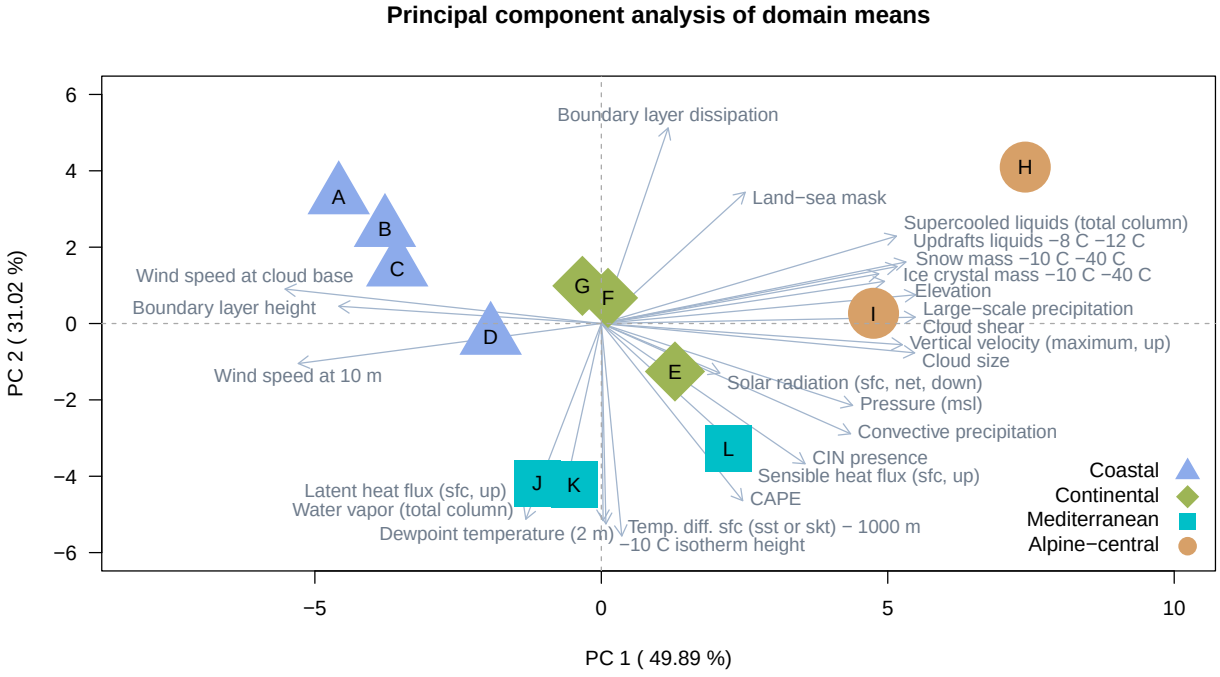

Figure 2. Spatial lightning differences in Europe. Based on Table 1, a PCA is computed and displayed as a biplot using the first two principal components (PC) as axes. Labeled arrows (loadings) indicate the contribution of each meteorological input variable to the variance in the respective direction. Domains with similar characteristics are indicated with the same color and labeled accordingly (legend).

indicating a rather cool ground (Table 1). The regional characteristics of domains A–D are their large ocean areas including coastlines, hence they are grouped together as the 'coastal' region. The domains J, K, and L (turquois squares) gather in the lower part of Fig. 2. Their common physical characteristics relative to the other domains are high 2 m dewpoint temperatures above 13 °C, elevated −10 °C isotherm heights of more than 4100 m, large CAPE values ($> 400\,\mathrm{J\,kg^{-1}}$), the presence of CIN, and high amounts of total column water vapor ($\sim 25\,\mathrm{kg\,m^{-2}}$). The temperature difference between the surface and at 1000 m altitude is more than 8 K, indicating a warm ground (Table 1). Regionally, all these domains are located in the Mediterranean, and are hence grouped together as the 'mediterranean' region. The domains H and I are located on the right or top-right of Fig. 2 (orange circles) and are physically characterized by increased cloud-physics variables and increased wind-field variables such as various increased cloud particle concentrations (ice, snow, and liquids), increased cloud shear, increased updrafts, and large amounts of large-scale precipitation. Regionally, both domains are located in Central Europe and are influenced by mountainous topography. Hence, they are grouped together as the 'alpine-central' region. The remaining domains E, F, and G (green diamonds) are all located in the center of Fig. 2 and share the physical characteristics of mostly average values but

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

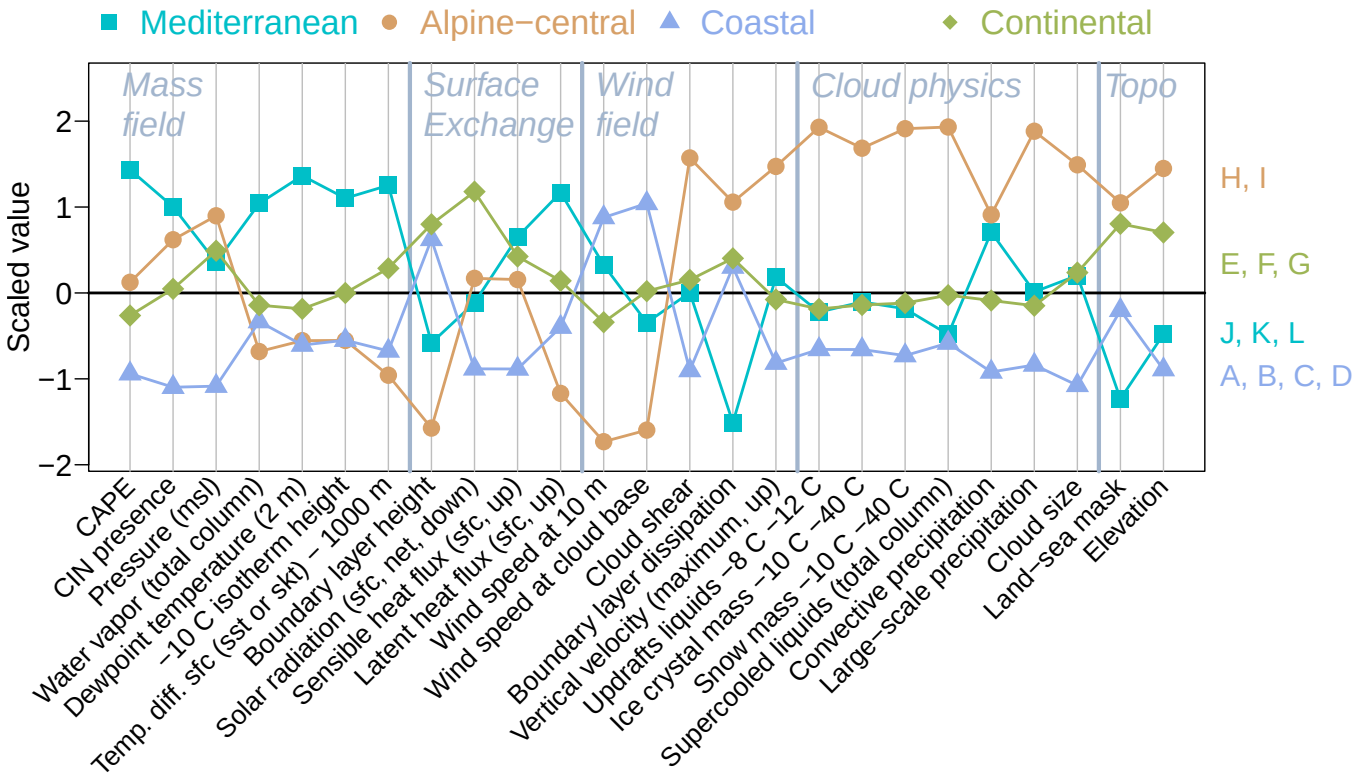

Figure 3. Parallel coordinate plot of mean meteorological values based on the means provided in Table 1. Similar domains are summarized into regions having the same color. ERA5 variables are scaled to mean zero and standard deviation one. More details are provided in Appendix A1.

increased surface exchange values. Spatially, their common characteristic is their location on the European mainland and hence they are grouped together as the 'continental' region.

Figure 3 shows how the meteorological conditions vary in detail between the four regions. It compares the mean meteorological values for each region (lines) relative to the others and shows how distinct thunderstorm conditions in Europe are. Scaled values (y-axis) close to zero indicate average values in the respective variable (x-axis) compared to the other regions. The figure shows that e.g. CAPE is in general much higher in the mediterranean region compared to the others, or that increased wind speed in the alpine-central region refers to much lower values than in all other regions. Appendix A1 dives deeper and presents the means of each domain separately.

With this the spatial different thunderstorm conditions in Europe are described and summarized into four regions, each with shared physical characteristics during lightning throughout the year: the mediterranean region, the alpine-central region, the coastal region, and the continental region.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

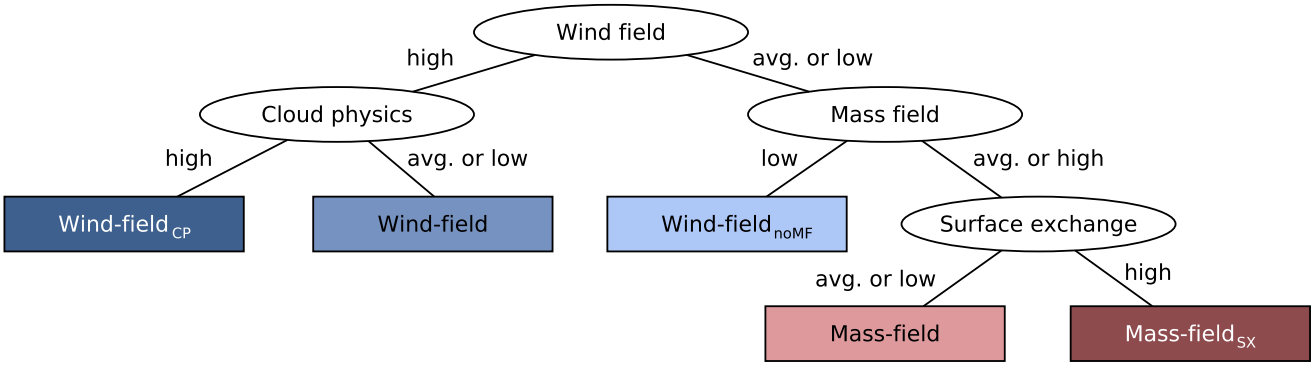

Figure 4. Decision tree to label the clusters with the corresponding two major thunderstorm types and three sub-types (colored boxes). The abbreviations in the sub-types stand for: CP = additionally increased cloud-physics variables, noMF = decreased mass-field variables, and SX = increased surface-exchange variables.

4.2 Thunderstorm types

After finding four regions where thunderstorms have similar characteristics throughout the year, the next goal is to investigate whether *individual* thunderstorms occur under similar larger-scale meteorological conditions, i.e. whether different thunderstorm types exist.

Cluster analysis with $k = 3$ is performed separately on every domain to find thunderstorm types (clusters) relative to the overall lightning characteristics in that domain. Each found cluster from each domain is then described by its driving meteorological characteristics using the average values of the 25 input variables (cluster means). Then the average values within the physically-based categories (mass field, wind field, cloud physics, surface exchange, and topography) are computed for each cluster to yield an overall characterization. Two major thunderstorm types emerge as the wind-field category and the mass-field category always deviate substantially. The decision tree in Fig. 4 distinguishes between these two types and helps to identify further sub-types. Wind-field thunderstorms are characterized by increased wind-field variables and sometimes decreased mass-field variables and are indicated by bluish colors. There are two wind-field sub-types: Wind-field$_{\mathrm{CP}}$ thunderstorms (dark blue) have additionally enhanced cloud-physics variables (CP) and wind-field$_{\mathrm{noMF}}$ thunderstorms (light blue) that have decreased mass-field variables (no MF) while wind-field variables and cloud-physics variables are at their average values. The other major thunderstorm type is a mass-field thunderstorm characterized by average or increased mass-field variables plus often decreased surface-exchange variables and is indicated by reddish colors. There is one sub-type, mass-field$_{\mathrm{SX}}$ thunderstorms (dark red), that has increased surface-exchange variables (SX) and sometimes average mass-field values.

Now we dive deeper into the characteristics of the found thunderstorm types by investigating the cluster means of three representative domains in Fig. 5. The cluster means (lines) are displayed as scaled values (y-axis) of the meteorological variables (x-axis). A value close to zero indicates a typical value for thunderstorms in that domain while large deviations indicate

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

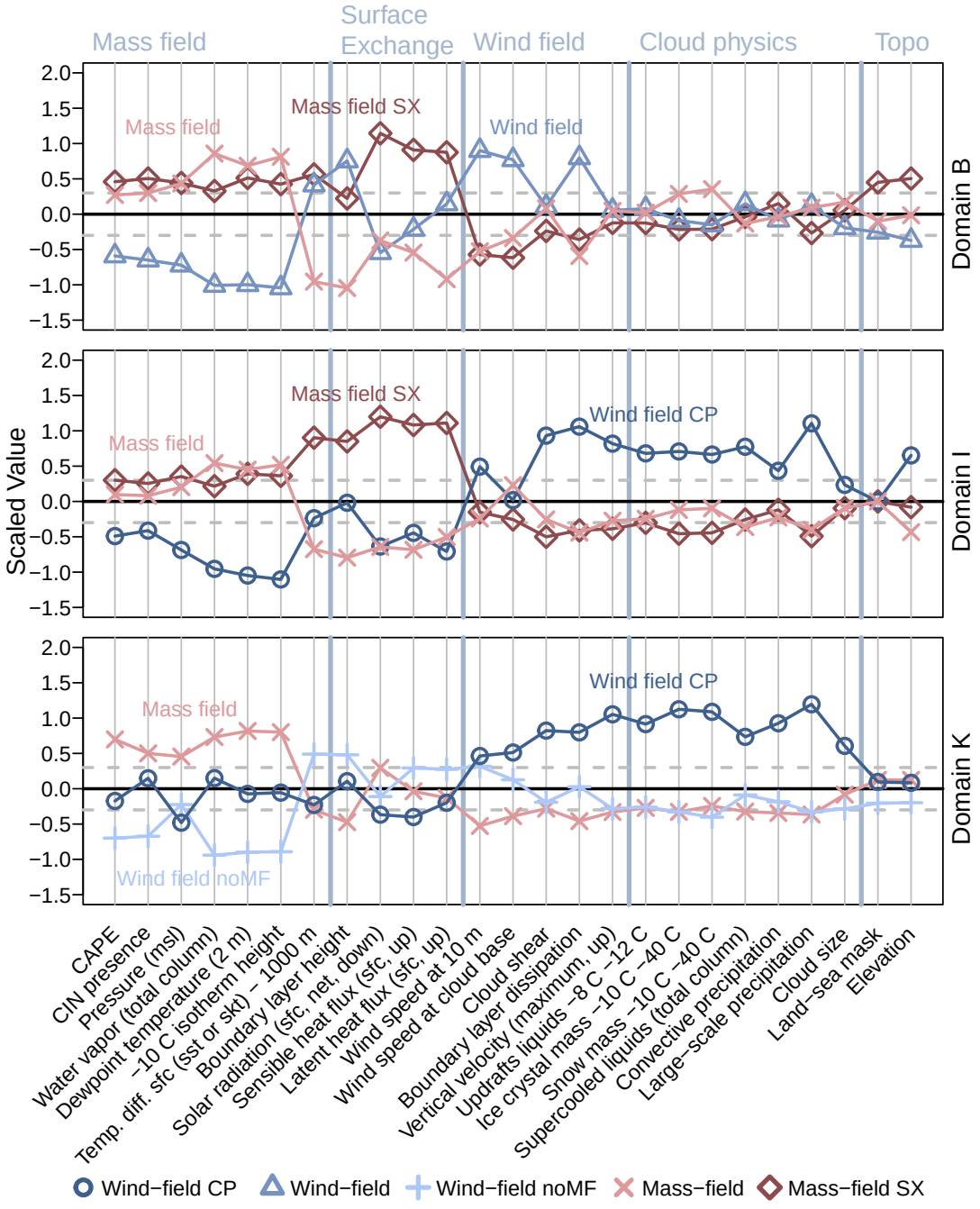

Figure 5. Parallel coordinate plot of thunderstorm types (clusters) found in three representative domains (B, I, K; panels) and expressed by the cluster means (lines) for each meteorological variable (x-axis) using scaled values (y-axis). The ERA5 variables are grouped by their meteorological category (secondary x-axis). Colors indicate the thunderstorm types (legend; blues = wind-field thunderstorms, reds = mass-field thunderstorms). The abbreviations in the sub-types stand for: CP = additionally increased cloud-physics variables, noMF = decreased mass-field variables, and SX = increased surface-exchange variables. Category means between the dashed lines (+/- 0.3) are considered as average.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

large differences in this variable for different thunderstorm types in that domain (as standard deviations). The average values or 'baselines' (y = 0) for each domain in Fig. 5 are provided in Table 1. Baselines are required to decide whether a value refers to a general high value. The unscaled cluster means are given in Table 2. The dashed lines indicate the +/- 0.3 threshold used for the decision tree. Results for all domains are supplied in Appendix B1–B4, Table 2, and in the online supplement (Morgenstern et al., 2022b). For robustness, each cluster analysis is repeated 50 times but only one representative result is shown.

Figure 5 shows, how the wind-field thunderstorms (middle-blue triangles) in domain B are driven by enhanced wind speeds, enhanced boundary layer dissipation, lower $-10\,°C$ isotherm heights, little water vapor, small CAPE, and large boundary layer heights of more than $1200\,m$ (Table 2). Different from this, the wind-field$_{CP}$ thunderstorms (dark-blue circles) in domains I and K show thick clouds ($8731\,m$ and $9756\,m$) with concentrations of cloud ice, cloud snow, and supercooled liquids that are $2-6$ times higher compared to the other two thunderstorm types in these domains as well as large precipitation amounts and strong updrafts. Wind-field$_{noMF}$ (light blue pluses) have in general very low mass-field values as seen in the lowest panel (domain K). All three bluish wind-field thunderstorm types have low CAPE values across all domains with means of $194\,J\,kg^{-1}$ in wind-field$_{CP}$ thunderstorms, $91\,J\,kg^{-1}$ in wind-field$_{noMF}$ thunderstorms, and $56\,J\,kg^{-1}$ in the remaining wind-field thunderstorms (without sub-type). This suggests to associate at least the latter two as no-CAPE thunderstorms. The reddish mass-field thunderstorms have in all domains high mass-field variables with large CAPE, high water vapor concentrations, and elevated $-10\,°C$ isotherm heights. Differences between the two reddish lines occur almost only in the surface-exchange variables. High surface-exchange values are characteristic of mass-field$_{SX}$ thunderstorms which are characterized by high downward solar radiation ($>\,320\,W\,m^{-2}$), large latent heat fluxes ($>\,140 W\,m^{-2}$), and upward oriented sensible heat fluxes. The thunderstorm sub-types wind-field$_{noMF}$ and mass-field$_{SX}$ both occur in conditions where the (sea) surface is hot relative to the air at $1000\,m$ altitude with an average temperature difference of $10.5\,K$, while in mass-field thunderstorms the air mass at $1000\,m$ is only about $3.9\,K$ colder than the surface. Regarding the topographical influences, mass-field thunderstorms occur more often over land (higher land-sea mask values), and wind-field thunderstorms more often over the sea. In each domain, at least one wind-field related thunderstorm type and one mass-field related thunderstorm type are found. The two major thunderstorm types clearly separate from one another. Varying the number of clusters k robustly finds similar results. With $k = 2$ only the two major types are found. With $k > 3$ more and more clusters are found referring to the same thunderstorm type revealing no additional meteorological insights.

In summary, there are two major thunderstorm types in Europe (wind-field thunderstorms and mass-field thunderstorms) and three sub-types thereof. Thunderstorm types are found by applying cluster analysis on 12 domains. A decision tree is developed to differentiate the found thunderstorm types using their driving meteorological categories.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

Table 2. Cluster means for each thunderstorm type and domain.

	CAPE	CIN presence	Pressure (msl)	Water vapor (total column)	Temperature dewpoint (2 m)	-10 C isotherm height	Temp. diff. sfc (sst or skt) - 1000 m a.g.l.	Boundary layer height	Solar radiation (sfc, net, down)	Sensible heat flux (sfc, up)	Latent heat flux (sfc, up)	Wind speed at 10 m	Wind speed at cloud base	Cloud shear	Boundary layer dissipation	Vertical velocity (maximum, up)	Updrafts liquids -8 C -12 C	Ice crystal mass -10 C -40 C	Snow mass -10 C -40 C	Supercooled liquids (total column)	Convective precipitation	Large-scale precipitation	Cloud size	Land-sea mask (0 = sea, 1 = land)	Elevation
Unit	$J\,kg^{-1}$	binary	hPa	$kg\,m^{-2}$	°C	m a.g.l.	K	m	$W\,m^{-2}$	$W\,m^{-2}$	$W\,m^{-2}$	$m\,s^{-1}$	$m\,s^{-1}$	$m\,s^{-1}$	$W\,m^{-2}$	$Pa\,s^{-1}$	$g\,Pa\,s^{-1}$	$g\,m^{-2}$	$g\,m^{-2}$	$g\,m^{-2}$	mm	mm	m	binary	m
Coastal region																									
A mass-field	297	0.59	1009.9	28.3	13.5	4533	3.5	440	164	4	50	5.1	9.7	13.3	2	0.587	3	56	80	29	0.049	0.005	7258	0.43	9
A wind-field$_{noMF}$	66	0.15	999.9	11.8	5.2	2582	9.3	1147	76	22	106	10.1	15.6	11.4	9	0.544	2	31	32	28	0.037	0.003	5700	0.34	5
A wind-field$_{CP}$	158	0.38	1001.6	23.8	9.8	3815	5.2	966	69	-8	71	9.8	17.4	27.9	18	1.42	36	214	335	116	0.109	0.062	8885	0.42	7
B mass-field$_{SX}$	362	0.69	1009.8	24.5	12.9	4185	9.1	925	328	54	149	4.3	8.6	11.7	3	0.615	4	45	59	37	0.066	0.004	7331	0.88	43
B mass-field	305	0.58	1009.8	30.7	13.9	4699	2	331	65	-16	22	4.5	10.2	15.1	2	0.674	7	96	153	36	0.057	0.014	7786	0.62	27
B wind-field	59	0.11	999.3	11.5	4.2	2463	7.9	1232	40	-10	83	10.2	17.8	15.6	16	0.678	8	53	66	49	0.037	0.013	6036	0.55	15
C mass-field$_{SX}$	454	0.77	1010.9	26.5	13.7	4441	9.6	994	341	59	166	3.8	7.9	10.5	3	0.586	3	39	57	35	0.065	0.003	7707	0.95	117
C mass-field	315	0.61	1010.4	30.7	14.3	4735	2.6	324	46	-15	25	4	9.3	13.3	1	0.585	4	71	113	35	0.053	0.012	7407	0.71	79
C wind-field	43	0.07	1000.4	11.3	2.9	2317	6.8	1244	41	-33	60	8.9	18.2	17.7	20	0.731	9	61	72	54	0.034	0.018	5806	0.85	81
D mass-field	410	0.63	1011.4	31.9	15.8	4902	4.1	383	138	10	57	3.7	8.7	14.5	1	0.731	3	59	89	25	0.053	0.005	7432	0.56	68
D wind-field$_{noMF}$	76	0.21	1004.2	13.9	7.4	2962	10.8	1204	103	46	156	10.2	15	14.6	9	0.771	3	48	50	29	0.064	0.004	6106	0.27	25
D wind-field$_{CP}$	115	0.29	1007.1	18.2	8.2	3148	7.5	999	60	-4	89	7.3	13.1	26.5	17	1.755	43	144	223	121	0.096	0.058	8059	0.77	155
Continental region																									
E mass-field$_{SX}$	329	0.57	1013.3	21.7	11.4	4071	13.2	1268	412	125	146	2.7	5.5	13.7	3	0.796	4	38	48	47	0.055	0.005	7725	0.99	769
E mass-field	424	0.61	1013.9	26.8	14.1	4471	5	343	59	2	35	2.8	7.2	15.9	2	0.661	3	53	80	28	0.044	0.007	7373	0.78	484
E wind-field$_{CP}$	123	0.37	1008.1	17.1	8.2	3140	8.3	993	72	17	89	7.6	12.5	24.1	14	1.428	29	106	161	85	0.102	0.047	7974	0.63	308
F mass-field$_{SX}$	408	0.68	1011.7	24.5	13.1	4261	10.5	1020	364	75	171	2.9	6.6	13	3	0.654	4	39	54	37	0.057	0.004	7574	0.99	310
F mass-field	380	0.66	1011	30.1	14.5	4656	3.3	317	33	-13	27	2.8	8.9	17.4	3	0.922	11	104	175	42	0.073	0.023	8406	0.96	276
F wind-field	58	0.16	1001.9	13.6	5.8	2763	7.3	1153	75	-11	78	7.9	15.7	20.7	17	0.926	17	68	95	66	0.054	0.026	6579	0.83	178
G mass-field$_{SX}$	451	0.73	1012.7	25.1	13	4318	10.5	1015	372	73	176	2.7	6	11.2	2	0.598	4	39	57	37	0.06	0.004	8007	1	406
G mass-field	318	0.59	1012	29.2	13.9	4612	3.2	269	29	-14	24	2.4	8.5	14.4	2	0.702	7	91	156	34	0.056	0.012	7992	1	383
G wind-field	64	0.18	1003.6	11.6	2.6	2285	6.2	1272	49	-38	50	7.3	16.3	24.8	27	0.921	19	83	113	68	0.046	0.033	6071	1	359
Alpine-central region																									
H mass-field$_{SX}$	486	0.7	1013.5	21.4	12	3932	9.9	832	342	63	171	1.7	3.9	15.4	3	0.773	5	49	77	40	0.074	0.009	8778	1	918
H mass-field	287	0.47	1013.1	21	10.2	3736	3.5	147	21	-4	13	1.4	4.3	18.1	3	0.727	4	53	80	39	0.046	0.013	7611	1	855
H wind-field$_{CP}$	122	0.37	1007.3	14.7	3.9	2687	4.1	471	24	-11	23	2.8	7.8	35.6	31	2.608	164	182	397	239	0.088	0.15	8085	1	910
I mass-field$_{SX}$	481	0.73	1012.5	25.6	13.8	4405	10.8	1036	392	79	188	2.2	5.1	11.7	2	0.647	4	37	56	37	0.061	0.004	7964	1	437
I mass-field	397	0.65	1011.5	28.9	14.2	4592	3.2	225	28	-10	25	2.1	7.5	14.7	2	0.717	5	69	119	31	0.057	0.008	7950	1	314
I wind-field$_{CP}$	139	0.4	1005.6	15.8	6	2902	4.7	554	30	-6	20	3.2	6.4	32.6	24	1.811	75	143	274	149	0.101	0.096	8731	1	676
Mediterranean region																									
J mass-field	759	0.77	1012.9	31.4	17.3	4919	6.8	457	173	17	74	4.3	7.8	14.6	1	0.607	2	42	64	22	0.035	0.005	7213	0.25	66
J wind-field$_{noMF}$	97	0.26	1008.7	16	8.7	3260	11.3	973	121	43	144	7.8	10.2	13.5	3	0.601	2	37	40	29	0.042	0.004	5905	0.14	38
J wind-field$_{CP}$	308	0.66	1008.2	26.7	13.2	4132	8.3	763	68	20	109	8.3	12.4	24.9	6	1.654	35	169	295	93	0.149	0.07	9907	0.23	59
K mass-field	930	0.84	1012.7	31	17.6	4916	7.8	501	218	31	94	3.6	6.4	12.2	1	0.598	2	38	57	24	0.04	0.004	7328	0.4	142
K wind-field$_{noMF}$	103	0.27	1008.2	15.7	8.5	3227	11	877	128	39	127	6.5	8.9	13.1	3	0.62	2	35	37	31	0.042	0.004	5992	0.24	82
K wind-field$_{CP}$	316	0.67	1006.5	25.3	12.9	4034	7.8	716	65	14	92	7.2	11.4	25.1	8	1.636	28	160	248	83	0.155	0.057	9756	0.39	129
L mass-field	933	0.9	1011	31.2	17.3	4859	7.5	448	189	23	90	3.5	6.2	14.6	1	0.691	3	52	81	29	0.058	0.006	8218	0.47	110
L wind-field$_{noMF}$	114	0.27	1008.5	16.2	8.3	3212	10.3	697	107	32	102	5.1	7.3	17	2	0.763	3	50	56	32	0.046	0.007	6645	0.32	96
L wind-field$_{CP}$	268	0.61	1005.4	23.8	11.4	3819	7	668	45	4	75	6.2	10.9	28.3	11	2.052	59	174	284	124	0.162	0.082	9720	0.61	205

CAPE: convective available potential energy, CIN: convective inhibition, msl: mean sea level, a.g.l.: above ground level,

Temp. diff. sfc (sst or skt) - 1000 m a.g.l.: Temperature difference between the surface and 1000 m a.g.l., where sfc is either sea surface temperature or skin temperature.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

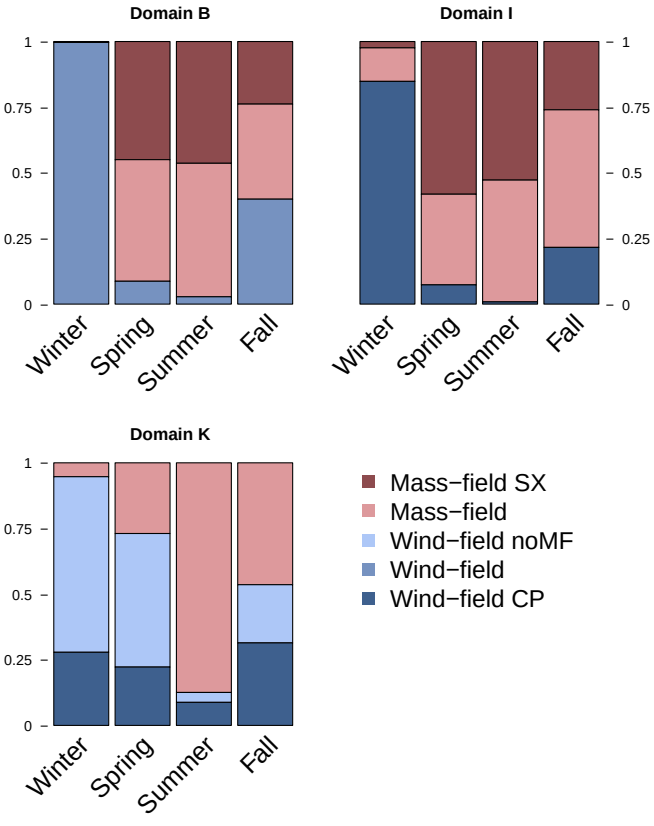

Figure 6. Seasonal variation of the thunderstorm types within three representative domains (B, I, K; panels). Bars are equally high because the same number of observations from each season is used. Results for each domain are based on cluster analysis estimated on each domain separately using local scaling values.

4.3 Seasonal differences between thunderstorm types in Europe

Now that the thunderstorm types are found, named, and described, their seasonality is investigated.

The stacked barplots in Fig. 6 show how many lightning observations from each season belong to a given thunderstorm type. As the data set is built to have the same number of observations from each season, the bars are equally high. The absolute numbers of observations per domain are given in Table 1. In all domains, winter (DJF) is dominated by wind-field thunderstorm types (blues) and summer (JJA) by mass-field thunderstorm types (reds). Spring and fall are transitional seasons with varying proportions. If a domain has two wind-field thunderstorm types (e.g., domain K), there is often a bigger one with a more pronounced annual cycle (wind-field$_{noMF}$) while the other one (wind-field$_{CP}$) is smaller and shows less seasonality.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

The map in Fig. 7 spatially compares barplots that are estimated individually on each domain using local mean and standard deviations for scaling. The polygon colors indicate which domains are similar to one another (Sect. 4.1) and hence more comparable as they are scaled with similar values (baselines, Table 1). In every domain, wind-field thunderstorms (blues) dominate in winter (first bar) and contribute to a varying fraction of thunderstorms in spring and fall, which is higher the more maritime a domain gets (domains A, D, J, K, L). Mass-field thunderstorms (reds) always dominate in summer (third bar) and over the mainland also in spring and fall. The presented cluster analysis with $k = 3$ has in all maritime domains (A, D, J, K, L) two wind-field thunderstorm types present, and in all domains at the mainland (B, C, E, F, G, H, I) two mass-field thunderstorms. This reveals the importance of wind-field thunderstorms over the seas. Higher k results in further splitting of the displayed thunderstorm types to have at least two wind-field thunderstorm types and two mass-field thunderstorm types present in every domain. In Fig. 7 wind-field thunderstorms in the southern domains (E, H, I) are accompanied by enhanced cloud-physics variables (wind-field$_{CP}$) which is remarkable in the alpine-central region (H, I) as all thunderstorms have in general very high cloud-physics variables there. As the presented thunderstorm characteristics are *relative* to other thunderstorms in each domain they are not directly comparable because some of them refer to very different baselines (Sect. 4.1). The mediterranean region (J, K, L) for example refers to much higher overall mass-field variables (Fig. 3) than the coastal domains (e.g., domain A).

In summary, we conclude that wind-field thunderstorms dominate the cold season and are more important over the sea while mass-field thunderstorms dominate the warm season and are more important over the mainland.

4.4 Comparability of the found thunderstorm types

The thunderstorm types are identified *relative* to the general meteorological conditions during lightning in each domain and the question remains how similar are thunderstorm types with the same name from different domains.

To make the thunderstorm types more comparable, a principal component analysis is estimated on all cluster means from every domain using the same scaling (Fig. 8). Again, the first two principal components are displayed on the axis explaining together about 70 % of the variance (PC 3 explains additionally 13.4 %) and the labeled arrows (loadings) indicate the contribution of each variable to the variance in the respective direction. Each domain (letters) is represented by three colored circles, the thunderstorm types found there. First of all, the figure shows that the two major thunderstorm types, wind-field thunderstorms and mass-field thunderstorms, are clearly separable from one another as the bluish and reddish circles are located in different parts of the figure. The difference between the mass-field thunderstorms is small as the reddish circles gather close to one another. Their major difference is in the surface-exchange variables that separate the light red mass-field thunderstorms from the dark red mass-field$_{SX}$ thunderstorm sub-type, which becomes more relevant in PC 3. Wind-field thunderstorms are more diverse as the bluish circles spread widely. They dominate the cold season, where climatologically little thunderstorms occur. Hence, the scarce cold-season thunderstorms in Europe originate in very diverse thunderstorm conditions while frequent summertime mass-field thunderstorms originate in similar weather patterns.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

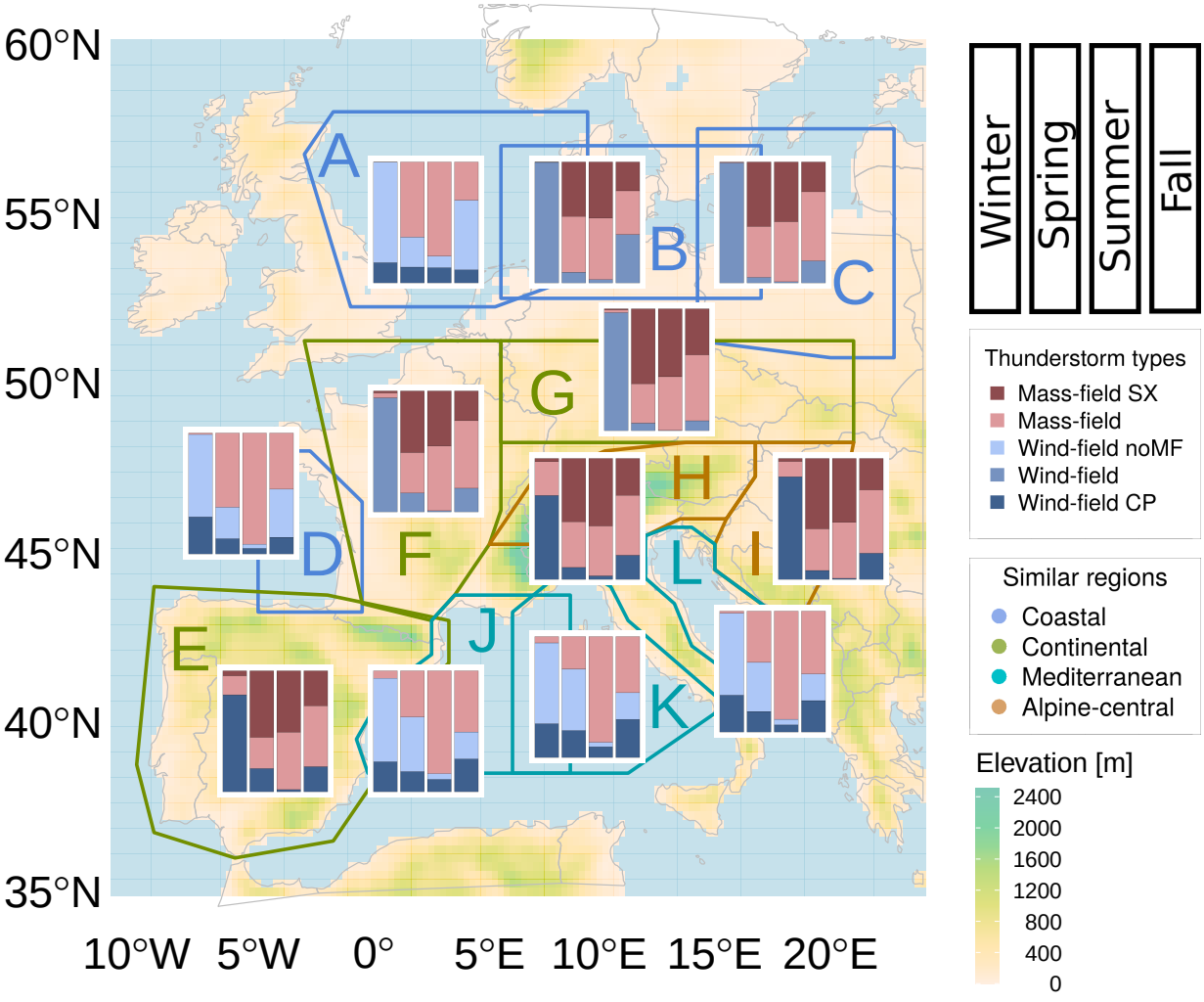

Figure 7. Spatio-temporal thunderstorm analysis in Europe. Barplots are based on cluster analyses that are individually estimated on each domain using local scaling values (as in Fig. 6). The colors of the domain borders indicate the similarity of some domains.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

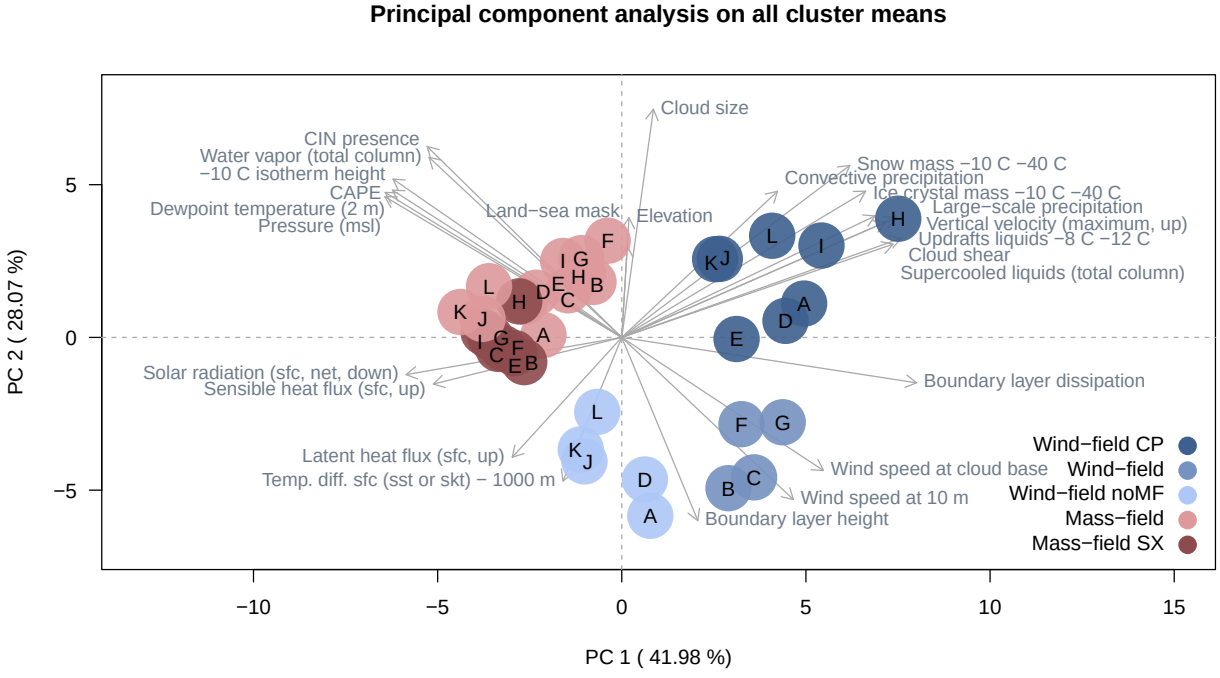

Figure 8. Comparison between all found clusters from all domains using the same scaling. The figure shows a biplot based on a PCA estimated on Table 2 and taking the first two principal components (PC) as axes. Each point represents a cluster and is colored and labeled according to its thunderstorm type and domain. The labeled arrows (loadings) indicate the contribution of each variable to the variance in the respective direction.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

5 Discussion

Regional lightning differences are described by four distinct regions: coastal, continental, mediterranean, and alpine-central. And thunderstorm characteristics in different meteorological environments are provided by the thunderstorm types (wind-field thunderstorms and mass-field thunderstorms plus sub-types).

Other authors have also investigated thunderstorm conditions. The variables important for our wind-field thunderstorms are similar to Mäkelä et al. (2013)'s investigations of winter lightning in Finland revealing the importance of vertical temperature difference between the surface and mid-troposphere ($700/500\,\mathrm{hPa}$) and shear, but not CAPE. Another classification was performed by Fujii et al. (2013) in Japan, who found that the number of winter lightning strokes and the probability of high-current lightning strokes, group into the *storm type* and *inactive type* dependent on the $-10\,^{\circ}\mathrm{C}$ isotherm height. Sherburn and Parker (2014) coined the term HSLC-thunderstorms, which are meteorological environments of high-shear ($\geq 18\,\mathrm{m\,s^{-1}}$ at $0-6\,\mathrm{km}$) and low-CAPE ($\leq 500\,\mathrm{J\,kg^{-1}}$) producing lightning in all seasons and at all times of day in the United States. Considering the overall lower CAPE values in Europe, this HSLC concept relates to our wind-field thunderstorm type, especially the sub-type with reduced mass-field values (wind-field$_{\mathrm{noMF}}$). Market et al. (2002) show in their thundersnow climatology over the contiguous United States that lightning associated with snowfall origins in seven different meteorological environments. This supports our finding of the very diverse wind-field thunderstorm conditions. Stucke et al. (2022) relate our two thunderstorm types as described in Morgenstern et al. (2022a) to upward lightning at two alpine towers and find that most upward lightning occurs in wind-field thunderstorm conditions. In general, thunderstorm frequencies under different synoptic conditions are often described (e.g., Wapler and James, 2015; Enno et al., 2014; Bielec, 2001; Kolendowicz, 2006) and regional thunderstorm differences are often subject of classical climatologies as mentioned in the introduction. For the baltic countries, Enno et al. (2013) found three distinct thunderstorm regions (continental, transitional, maritime) similar to some of our thunderstorm regions (continental, coastal). And for the UK and Ireland, Hayward et al. (2022) conduct a regional cluster analysis aiming to identify areas where the seasonal distributions of lightning densities differ. This climatology nicely complements our approach with the PCA (Sect. 4.1) as it has a better resolution and covers adjacent regions, where EUCLID data does not fulfill our quality requirements.

This study is limited by the resolution of the data used. A higher resolution of atmospheric data in space and time and longer time series in the EUCLID data would improve the analysis. Applying a cluster analysis with more clusters or including also non-lightning information would lead to the same major results but could uncover more thunderstorm sub-types.

Now that the thunderstorm types are found, several new research questions arise that are beyond the scope of this paper. How often are wind turbines or other high structures struck by which thunderstorm type? What is the relation of the thunderstorm types to lightning properties such as the lightning duration, transferred charge, polarity, or channel length? It would also be interesting to model lightning probability maps for each thunderstorm type in each season.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

6 Conclusions

This study investigates regional and seasonal thunderstorm characteristics in Europe. Very destructive lightning damages often occur in seasons and regions where lightning is climatological unlikely. Our analysis includes infrequent lightning conditions by considering the same amount of lightning observations from each season. EUCLID lightning data is combined with meteorological ERA5 data to answer two research questions: "How do meteorological thunderstorm characteristics vary regionally across Europe?" and "What characterizes thunderstorms in different meteorological environments and how do these thunderstorm types vary seasonally across Europe?". Using principal component analysis, the European territory can be separated into four regions in which the atmospheric conditions for thunderstorms are similar throughout the year: The alpine-central region with thick clouds, large cloud particle masses, and strong updrafts relative to the other regions; the mediterranean region with increased mass-field variables; the coastal region with increased wind speeds; and the continental region with in general average conditions and increased solar radiation relative to the other regions. Cluster analysis is performed individually on 12 domains in Europe to find different thunderstorm types and a decision tree is developed to easily differentiate them (Fig. 4). There are two major thunderstorm types, wind-field thunderstorms and mass-field thunderstorms, and three sub-types thereof. Mass-field thunderstorms are characterized by increased CAPE values, the presence of CIN, large 2 m dewpoint temperatures, high $-10\,°C$ isotherm heights, and high mean sea level pressure relative to other thunderstorms in that domain. They occur mostly in the warmer seasons and always in similar weather conditions and are more important over the European mainland. The mass-field$_{SX}$ sub-type is associated with enhanced surface-exchange (SX) variables such as solar radiation and sensible heat flux and accounts for about half of the mass-field thunderstorms on the European mainland. The other major thunderstorm type is wind-field thunderstorms, which are more diverse but share the characteristics of average or reduced values in mass-field variables and elevated or average values in wind-field variables (high wind speeds at different heights, strong updrafts, large cloud shear, increased boundary layer dissipation) relative to other thunderstorms in that domain. They dominate the cold season, especially winter, and are more important over the sea. Sometimes the cloud-physics (CP) variables are additionally increased leading to the wind-field$_{CP}$ thunderstorm sub-type with large cloud sizes, increased concentrations of cloud particles (snow, ice, supercooled liquids), and large precipitation amounts. Another sub-type, wind-field$_{noMF}$ is characterized by decreased mass-field variables (no MF) and occurs often over the sea. In summary, this study shows that lightning in Europe origins in different meteorological environments, that winter lightning is not just a rarer sibling of summer lightning, and provides a decision tree to easily differentiate thunderstorm types in Europe independent of a seasonal criterion or static thresholds.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

Appendix A: Additional details to Sect. 4.1 "Regional differences between thunderstorms in Europe"

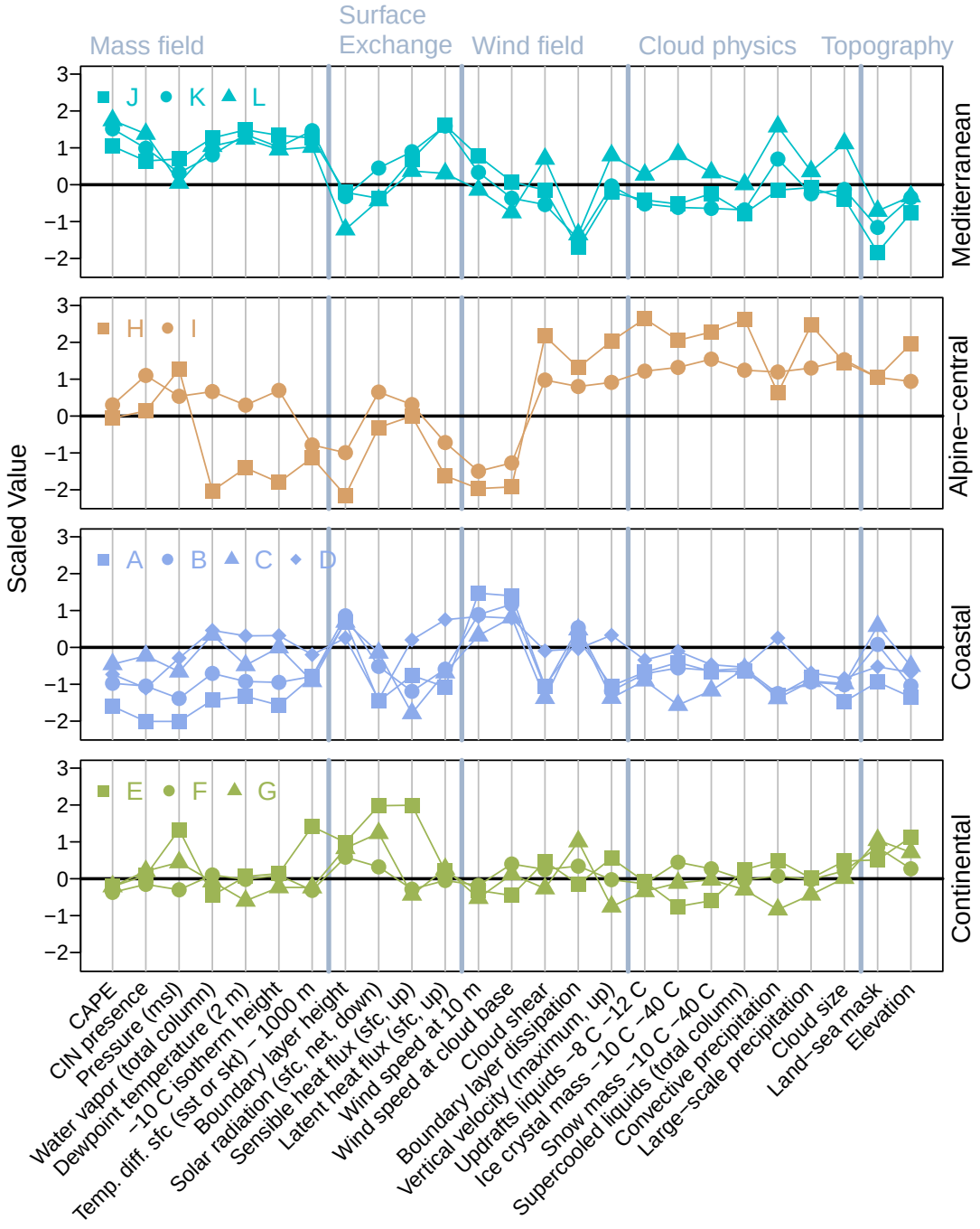

Figure A1. Additional details to the parallel coordinate plot in Fig. 3 based on Table 1. The panels are the basis for the means in Fig. 3.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

Appendix B: Additional details to Sect. 4.2 "Thunderstorm types"

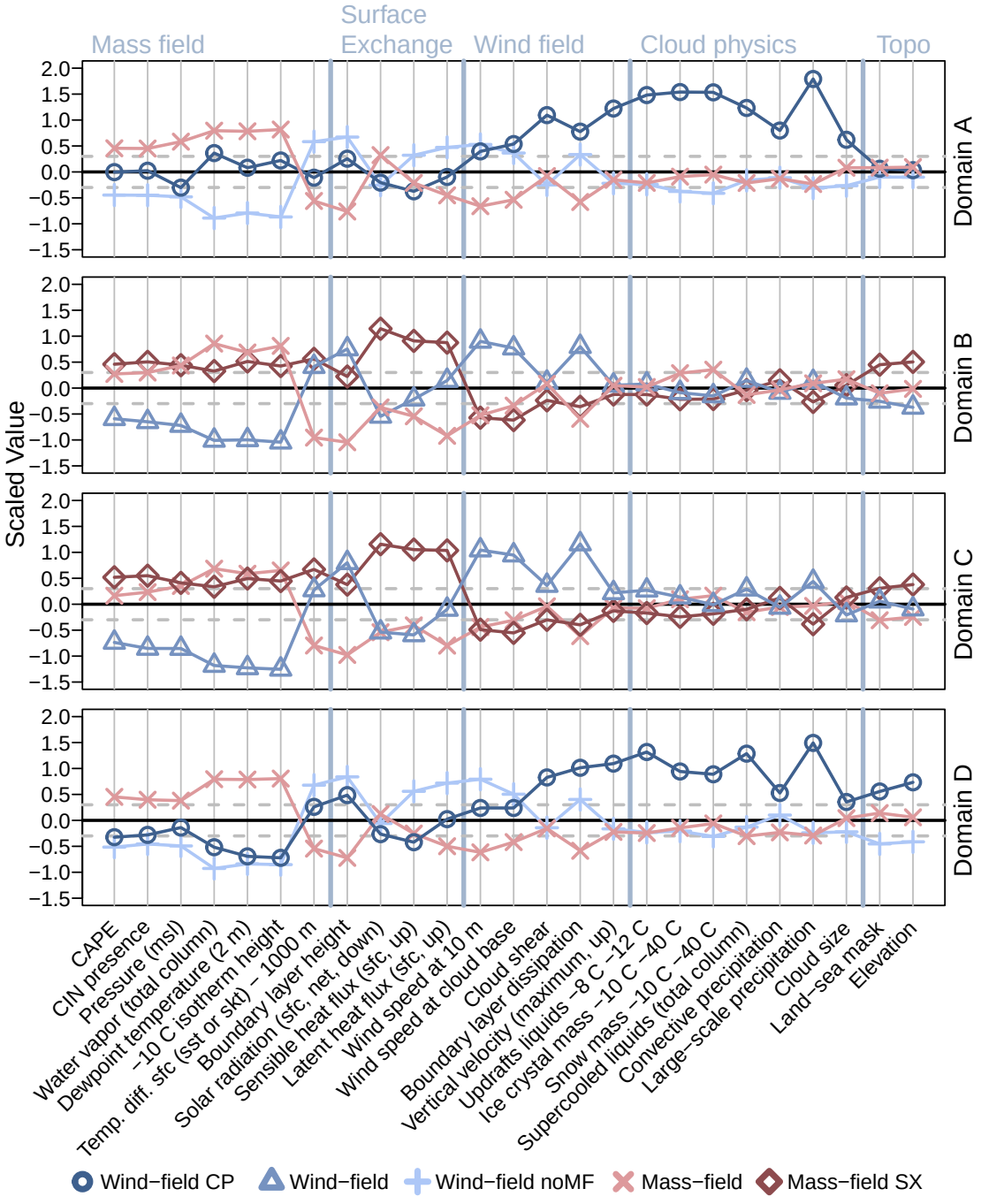

Figure B1. Additional details to Fig. 5. Here: Cluster means for the coastal domains. Numbers are given in Table 2.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

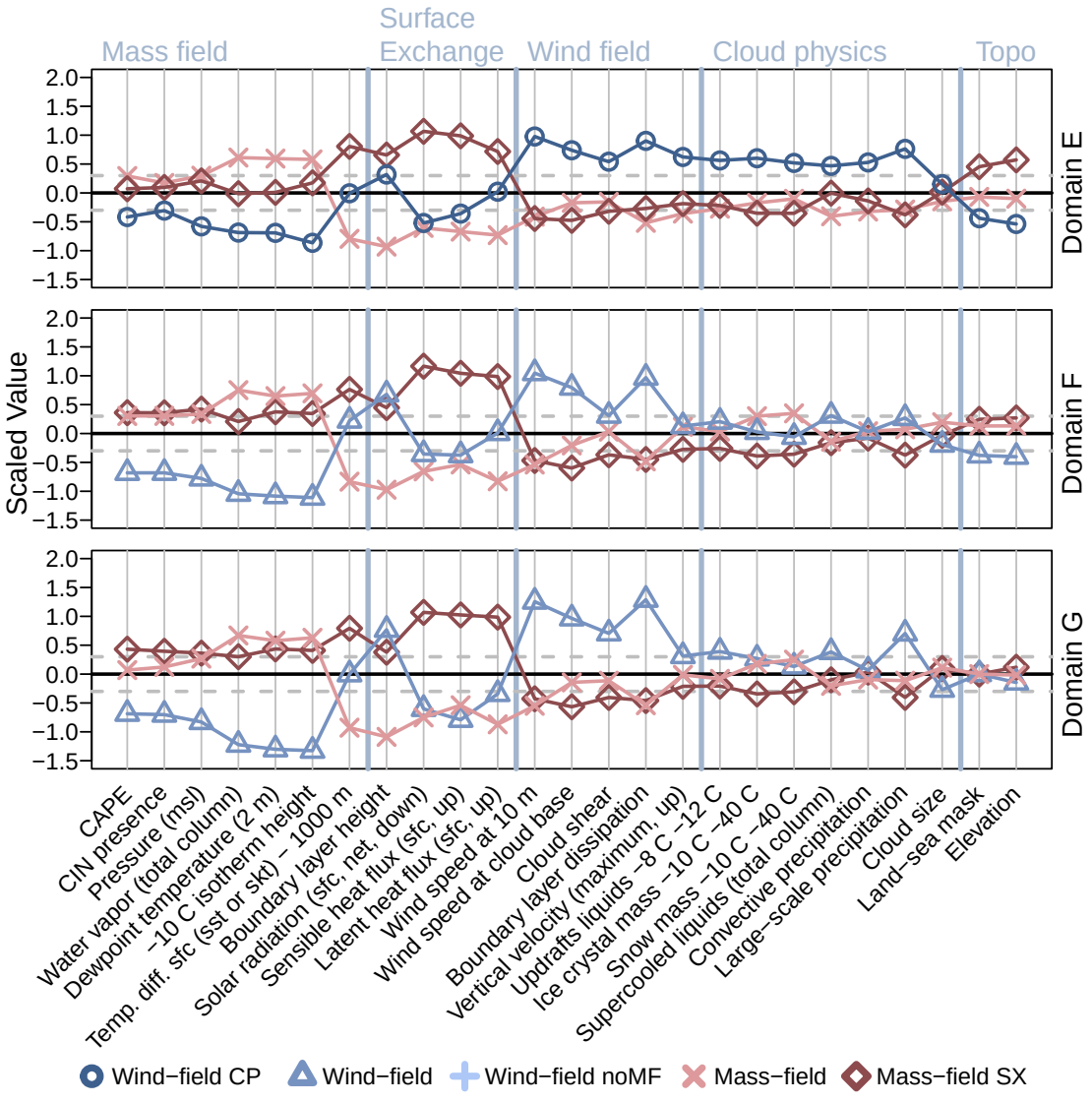

Figure B2. Additional details to Fig. 5. Here: Cluster means for the continental domains. Numbers are given in Table 2.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

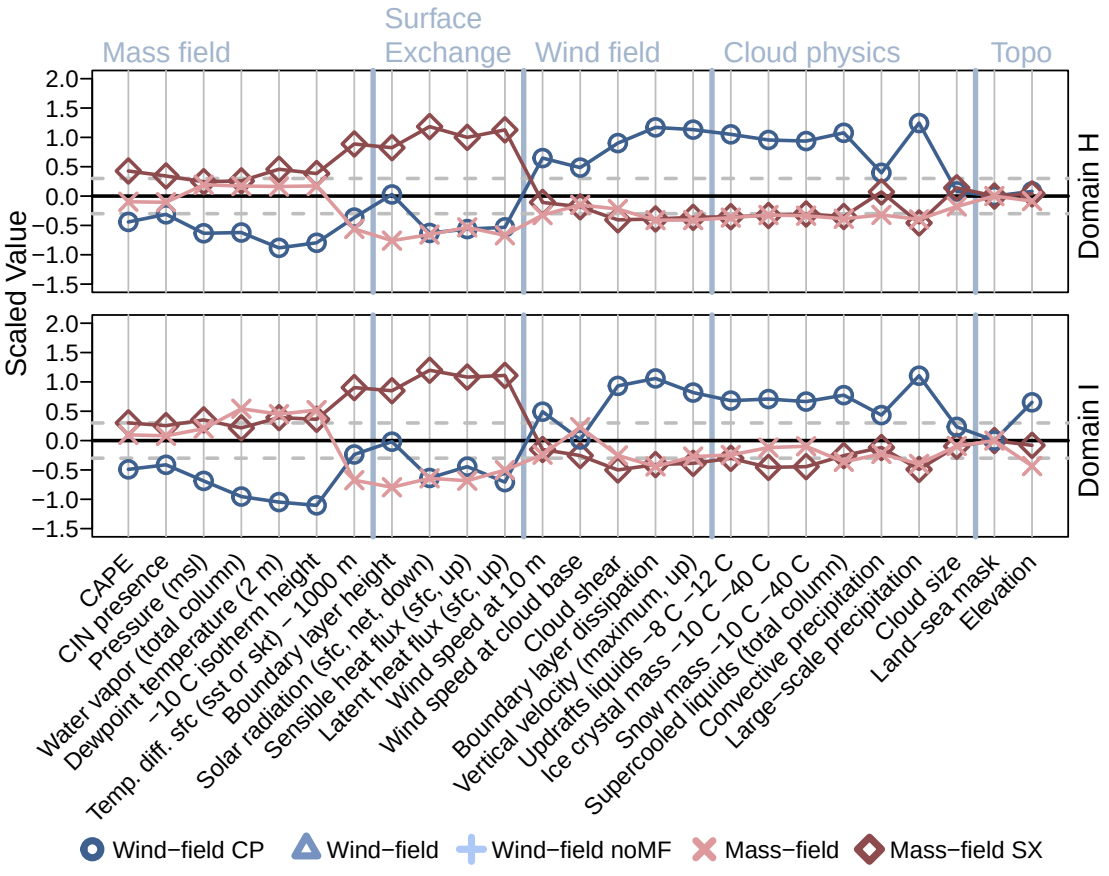

Figure B3. Additional details to Fig. 5. Here: Cluster means for the alpine-central domains. Numbers are given in Table 2.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

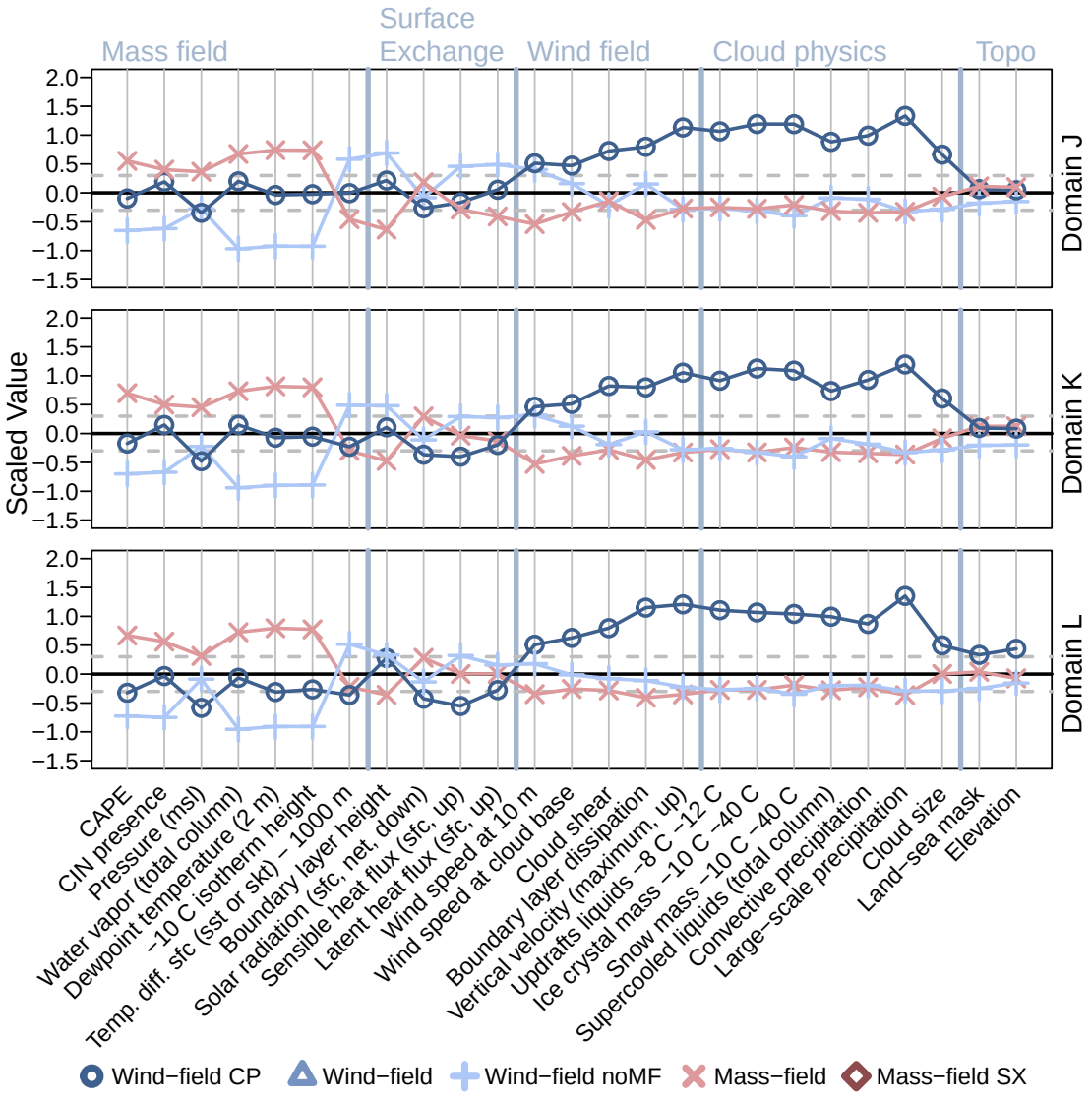

Figure B4. Additional details to Fig. 5. Here: Cluster means for the mediterranean domains. Numbers are given in Table 2.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

Code and data availability. The supplementary material (Morgenstern et al., 2022b) contains a R script to reproduce the core findings and main figures along with the required data for the presented sample, the domain definitions, a precise variable description, and the two tables. ERA5 data are freely available from the Copernicus Climate Change Service (C3S) Climate Data Store (Hersbach et al., 2020; https://cds.climate.copernicus.eu, last access: 30 September 2022). We use ERA5 hourly data 1959 to present on single level and model level (https://doi.org/10.24381/cds.adbb2d47, Hersbach et al., 2018; https://confluence.ecmwf.int/display/CKB/How+to+download+ERA5, last access: 25 November 2022). The results contain modified Copernicus Climate Change Service information for 2010–2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. EUCLID (Poelman et al., 2016; Schulz et al., 2016) data are available on request from ALDIS (Austrian Lightning Detection & Information System, aldis@ove.at) or Siemens BLIDS (Blitzinformationsdienst, fees may apply).

Calculations are performed using R (https://www.R-project.org/, R Core Team, 2021), Python 3 (https://www.python.org/, Van Rossum and Drake, 2009), and CDO (Climate Data Operator; https://doi.org/10.5281/zenodo.3539275, Schulzweida, 2019). Specifically the following packages were used: ncdf4 (https://CRAN.R-project.org/package=ncdf4, Pierce, 2019), sf (simple features; https://r-spatial.github.io/sf/index.html, Pebesma, 2018), stars (https://CRAN.R-project.org/package=stars, Pebesma, 2020), rnaturalearth (https://CRAN.R-project.org/package=rnaturalearth, South, 2017), data.table (https://github.com/Rdatatable/data.table/wiki, Dowle and Srinivasan, 2020), colorspace (https://colorspace.r-forge.r-project.org/, Stauffer et al., 2009), and xarray (Hoyer and Hamman, 2017). The netCDF4 data format is used (https://doi.org/10.5065/D6H70CW6, Unidata, 2020).

Author contributions. DM performed the investigation, wrote the software, visualized the results, and wrote the paper. IS, TS, and DM performed the data curation, built the data set, and derived variables based on ERA5 data. TS contributed coding concepts. GJM provided support for the meteorological analysis, data organization, and funding acquisition. AZ supervised the formal analysis and interpretation of the statistical methods. AZ, GJM, and TS are the project administrators and supervisors. All authors contributed to the conceptualization of this paper, discussed the methodology, evaluated the results, and commented on the paper.

Competing interests. The authors declare no competing interests.

Acknowledgements. We are grateful to Gerhard Diendorfer, Wolfgang Schulz, and Hannes Pichler from ALDIS for data support and discussions about lightning physics and to EUCLID for providing the LLS data.

This research has been supported by the Österreichische Forschungsförderungsgesellschaft (FFG) with grant no. 872656 and the Austrian Science Fund (Fonds zur Förderung der wissenschaftlichen Forschung, FWF) with the grants no. P 31836 and no. P 35780-NBL and with financial support from the State Tirol, Austria.

The computational results presented have been achieved in part using the Vienna Scientific Cluster (VSC).

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

References

Anderson, G. and Klugmann, D.: A European Lightning Density Analysis Using 5 Years of ATDnet Data, Natural Hazards and Earth System Sciences (NHESS), 14, 815–829, https://doi.org/10.5194/nhess-14-815-2014, accessed 2022-11-08, 2014.

Bentley, M. L., Riley, C., and Mazur, E.: A Winter-Season Lightning Climatology for the Contiguous United States, Meteorology and Atmospheric Physics, 131, 1327–1340, https://doi.org/10.1007/s00703-018-0641-2, accessed 2021-05-25, 2019.

Bielec, Z.: Long-Term Variability of Thunderstorms and Thunderstorm Precipitation Occurrence in Cracow, Poland, in the Period 1896–1995, Atmospheric Research, 56, 161–170, https://doi.org/10.1016/S0169-8095(00)00096-X, accessed 2022-11-08, 2001.

Coquillat, S., Pont, V., Lambert, D., Houel, R., Pardé, M., Kreitz, M., Ricard, D., Gonneau, E., de Guibert, P., and Prieur, S.: Six Years of Electrified Convection Over the Island of Corsica Monitored by SAETTA: General Trends and Anomalously Electrified Thunderstorms During African Dust South Flow Events, Atmospheric Research, 275, 106 227, https://doi.org/10.1016/j.atmosres.2022.106227, accessed 2022-11-08, 2022.

Dowle, M. and Srinivasan, A.: data.table: Extension of 'data.frame', https://CRAN.R-project.org/package=data.table, R package version 1.13.2, 2020.

Enno, S.-E., Briede, A., and Valiukas, D.: Climatology of Thunderstorms in the Baltic Countries, 1951–2000, Theoretical and Applied Climatology, 111, 309–325, https://doi.org/10.1007/s00704-012-0666-2, accessed 2022-04-27, 2013.

Enno, S.-E., Post, P., Briede, A., and Stankunaite, I.: Long-Term Changes in the Frequency of Thunder Days in the Baltic Countries, Boreal Environment Research, 19, 452–466, https://helda.helsinki.fi/bitstream/handle/10138/228651/ber19-5-452.pdf?sequence=1, accessed 2022-11-08, 2014.

Enno, S.-E., Sugier, J., Alber, R., and Seltzer, M.: Lightning Flash Density in Europe Based on 10 Years of ATDnet Data, Atmospheric Research, 235, https://doi.org/10.1016/j.atmosres.2019.104769, accessed 2021-06-08, 2020.

Fujii, F., Ishii, M., Saito, M., Matsui, M., and Natsuno, D.: Characteristics of Winter Lightning Threatening Wind Turbines in Coastal Area of the Sea of Japan, Electrical Engineering in Japan, 184, 44–50, https://doi.org/10.1002/eej.22357, accessed 2021-05-24, 2013.

Hayward, L., Whitworth, M., Pepin, N., and Dorling, S.: A Regional Lightning Climatology of the UK and Ireland and Sensitivity to Alternative Detection Networks, International Journal of Climatology, https://doi.org/10.1002/joc.7680, accessed 2022-11-08, 2022.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A.and Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., , and Thépaut, J.-N.: ERA5 Hourly Data on Single Levels from 1979 to Present, data set, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.adbb2d47, accessed 2022-11-25, 2018.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 Global Reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, accessed 2021-04-12, 2020.

Holle, R. R.: A Summary of Recent National-Scale Lightning Fatality Studies, Weather, Climate, and Society, 8, 35–42, https://doi.org/10.1175/WCAS-D-15-0032.1, accessed 2022-12-02, 2016.

Holt, M. A., Hardaker, P. J., and McLelland, G. P.: A Lightning Climatology for Europe and the UK, 1990-99, Weather, 56, 290–296, https://doi.org/10.1002/j.1477-8696.2001.tb06598.x, accessed 2022-11-08, 2001.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

Hoyer, S. and Hamman, J.: xarray: N-D Labeled Arrays and Datasets in Python, Journal of Open Research Software, 5, https://doi.org/10.5334/jors.148, 2017.

Iwasaki, H.: Preliminary Study on Features of Lightning Discharge around Japan using World Wide Lightning Location Network Data, SOLA Scientific Online Letters on the Atmosphere, 10, 98–102, https://doi.org/10.2151/sola.2014-020, accessed 2022-04-27, 2014.

Kolendowicz, L.: The Influence of Synoptic Situations on the Occurrence of Days with Thunderstorms During a Year in the Territory of Poland, International Journal of Climatology, 26, 1803–1820, https://doi.org/10.1002/joc.1348,, accessed 2022-11-08, 2006.

Kotroni, V. and Lagouvardos, K.: Lightning in the Mediterranean and its Relation with Sea-Surface Temperature, Environmental Research Letters, 11, 034 006, https://doi.org/10.1088/1748-9326/11/3/034006, accessed 2022-04-27, 2016.

MacQueen, J. B.: Some Methods for Classification and Analysis of Multivariate Observations, in: Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability, edited by Cam, L. M. L. and Neyman, J., vol. 1, pp. 281–297, University of California Press, https://www.bibsonomy.org/bibtex/25dcdb8cd9fba78e0e791af619d61d66d/enitsirhc, accessed 2020-02-19, 1967.

Mäkelä, A., Saltikoff, E., Julkunen, J., Juga, I., Gregow, E., and Niemelä, S.: Cold-Season Thunderstorms in Finland and Their Effect on Aviation Safety, Bulletin of the American Meteorological Society, 94, 847–858, https://doi.org/10.1175/BAMS-D-12-00039.1, accessed 2022-11-08, 2013.

Mäkelä, A., Enno, S.-E., and Haapalainen, J.: Nordic Lightning Information System: Thunderstorm Climate of Northern Europe for the Period 2002–2011, Atmospheric Research, 139, 46–61, https://doi.org/10.1016/j.atmosres.2014.01.008, accessed 2021-06-08, 2014.

Mallick, C., Hazra, A., Saha, S. K., Chaudhari, H. S., Pokhrel, S., Konwar, M., Dutta, U., Mohan, G. M., and Vani, K. G.: Seasonal Predictability of Lightning Over the Global Hotspot Regions, Geophysical Research Letters, 49, https://doi.org/10.1029/2021GL096489, accessed 2022-11-08, 2022.

Manzato, A., Serafin, S., Miglietta, M. M., Kirshbaum, D., and Schulz, W.: A Pan-Alpine Climatology of Lightning and Convective Initiation, Monthly Weather Review, 150, 2213–2230, https://doi.org/10.1175/MWR-D-21-0149.1, accessed 2022-11-08, 2022.

March, V., Montanyà, J., Fabró, F., van der Velde Oscar, Romero, D., Solà, G., Freijó, M., and Pineda, N.: Winter Lightning Activity in Specific Global Regions and Implications to Wind Turbines and Tall Structures, in: 2016 33rd International Conference on Lightning Protection (ICLP), 25–30 September 2016, Estoril, Portugal, IEEE, https://doi.org/10.1109/ICLP.2016.7791447, accessed 2022-04-09, 2016.

Mardia, K. V., Kent, J. T., and Bibby, J. M.: Multivariate Analysis, Probability and Mathematical Statistics, Academic Press, London, tenth edn., 1995.

Market, P. S., Halcomb, C. E., and Ebert, R. L.: A Climatology of Thundersnow Events over the Contiguous United States, Weather and Forecasting, 17, 1290–1295, https://doi.org/10.1175/1520-0434(2002)017<1290:ACOTEO>2.0.CO;2, accessed 2022-11-08, 2002.

Morgenstern, D., Stucke, I., Simon, T., Mayr, G. J., and Zeileis, A.: Differentiating Lightning in Winter and Summer with Characteristics of the wind field and mass field, Weather and Climate Dynamics, 3, 361–375, https://doi.org/10.5194/wcd-3-361-2022, accessed 2022-04-07, 2022a.

Morgenstern, D., Stucke, I., Simon, T., Mayr, G. J., and Zeileis, A.: Supplementary material: Thunderstorm Types in Europe, https://doi.org/10.5281/zenodo.7436761, 2022b.

Pebesma, E.: Simple Features for R: Standardized Support for Spatial Vector Data, The R Journal, 10, 439–446, https://doi.org/10.32614/RJ-2018-009, 2018.

Pebesma, E.: stars: Spatiotemporal Arrays, Raster and Vector Data Cubes, https://CRAN.R-project.org/package=stars, R package version 0.4-3, 2020.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

Pierce, D.: ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files, https://CRAN.R-project.org/package=ncdf4, R package version 1.17, 2019.

Pineda, N., Montanyà, J., Romero, D., Bech, J., Casellas, E., and González, S.: Meteorological Aspects of Winter Upward Lightning from an Instrumented Tower in the Pyrenees, in: 2018 34th International Conference on Lightning Protection (ICLP), 2–7 September 2018, Rzeszow, Poland, pp. 1–7, https://doi.org/10.1109/ICLP.2018.8503271, accessed 2022-04-27, 2018.

Piper, D. and Kunz, M.: Spatiotemporal Variability of Lightning Activity in Europe and the Relation to the North Atlantic Oscillation Teleconnection Pattern, Natural Hazards and Earth System Sciences, 17, 1319–1336, https://doi.org/10.5194/nhess-17-1319-2017, accessed 2021-06-08, 2017.

Poelman, D. R.: A 10-Year Study on the Characteristics of Thunderstorms in Belgium Based on Cloud-to-Ground Lightning Data, Monthly Weather Review, 142, 4839–4849, https://doi.org/10.1175/MWR-D-14-00202.1, accessed 2022-11-08, 2014.

Poelman, D. R., Schulz, W., Diendorfer, G., and Bernardi, M.: The European Lightning Location System EUCLID - Part 2: Observations, Natural Hazards and Earth System Sciences, 16, 607–616, https://doi.org/10.5194/nhess-16-607-2016, accessed 2020-02-19, 2016.

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

Schulz, W., Cummins, K., Diendorfer, G., and Dorninger, M.: Cloud-to-Ground Lightning in Austria: A 10-year Study using Data from a Lightning Location System, Journal of Geophysical Research: Atmospheres, 110, 20, https://doi.org/10.1029/2004JD005332, accessed 2022-11-08, 2005.

Schulz, W., Diendorfer, G., Pedeboy, S., and Poelman, D. R.: The European Lightning Location System EUCLID - Part 1: Performance Analysis and Validation, Natural Hazards and Earth System Sciences, 16, 595–605, https://doi.org/10.5194/nhess-16-595-2016, accessed 2020-02-19, 2016.

Schulzweida, U.: CDO User Guide, https://doi.org/10.5281/zenodo.3539275, 2019.

Sherburn, K. D. and Parker, M. D.: Climatology and Ingredients of Significant Severe Convection in High-Shear, Low-CAPE Environments, Weather and Forecasting, 29, 854–877, https://doi.org/10.1175/WAF-D-13-00041.1, accessed 2021-05-04, 2014.

Simon, T. and Mayr, G. J.: Lightning Climatology for the Eastern Alpine Region on the Kilometer Scale with Daily Resolution, e & i Elektrotechnik und Informationstechnik, 139, 352–360, https://doi.org/10.1007/s00502-022-01032-1, accessed 2022-11-25, 2022.

Simon, T., Umlauf, N., Zeileis, A., Mayr, G. J., Schulz, W., and Diendorfer, G.: Spatio-Temporal Modelling of Lightning Climatologies for Complex Terrain, Natural Hazards and Earth System Sciences, 17, 305–314, https://doi.org/10.5194/nhess-17-305-2017, accessed 2020-02-19, 2017.

South, A.: rnaturalearth: World Map Data from Natural Earth, https://CRAN.R-project.org/package=rnaturalearth, R package version 0.1.0, 2017.

Stauffer, R., Mayr, G. J., Dabernig, M., and Zeileis, A.: Somewhere over the Rainbow: How to Make Effective Use of Colors in Meteorological Visualizations, Bulletin of the American Meteorological Society, 96, 203–216, https://doi.org/10.1175/BAMS-D-13-00155.1, 2009.

Stucke, I., Morgenstern, D., Diendorfer, G., Mayr, G. J., Pichler, H., Schulz, W., Simon, T., and Zeileis, A.: Thunderstorm Types and Meteorological Characteristics of Upward Lightning, in: 2022 36th International Conference on Lightning Protection (ICLP), 2–7 October 2022, Cape Town, South Africa, pp. 282–288, https://doi.org/10.1109/ICLP56858.2022.9942489, 2022.

Taszarek, M., Czernecki, B., and Kozioł, A.: A Cloud-to-Ground Lightning Climatology for Poland, Monthly Weather Review, 143, 4285–4304, https://doi.org/10.1175/MWR-D-15-0206.1, accessed 2022-11-08, 2015.

https://doi.org/10.5194/egusphere-2022-1453
Preprint. Discussion started: 2 January 2023

Taszarek, M., Allen, J., Púčik, T., Groenemeijer, P., Czernecki, B., Kolendowicz, L., Lagouvardos, K., Kotroni, V., and Schulz, W.: A Climatology of Thunderstorms across Europe from a Synthesis of Multiple Data Sources, Journal of Climate, 32, 1813–1837, https://doi.org/10.1175/JCLI-D-18-0372.1, accessed 2021-05-04, 2019.

Taszarek, M., Allen, J. T., Groenemeijer, P., Edwards, R., Brooks, H. E., Chmielewski, V., and Enno, S.-E.: Severe Convective Storms across Europe and the United States. Part I: Climatology of Lightning, Large Hail, Severe Wind, and Tornadoes, Journal of Climate, 33, 10 239–10 261, https://doi.org/10.1175/JCLI-D-20-0345.1, accessed 2021-06-08, 2020a.

Taszarek, M., Allen, J. T., Púčik, T., Hoogewind, K. A., and 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, Journal of Climate, 33, 10 263–10 286, https://doi.org/10.1175/JCLI-D-20-0346.1, accessed 2021-05-04, 2020b.

Ukkonen, P. and Mäkelä, A.: Evaluation of Machine Learning Classifiers for Predicting Deep Convection, Journal of Advances in Modeling Earth Systems, 11, 1784–1802, https://doi.org/10.1029/2018MS001561, accessed 2021-06-08, 2019.

Unidata: Network Common Data Form (netCDF), Boulder, CO: UCAR/Unidata, https://doi.org/10.5065/D6H70CW6, 2020.

Van Rossum, G. and Drake, F. L.: Python 3 Reference Manual, Python Documentation Manual Part 2, CreateSpace Independent Publishing Platform, Scotts Valley, CA, accessed 2021-03-09, 2009.

Virts, K., Wallace, J., Hutchins, M., and Holzworth, R.: Highlights of a New Ground-Based, Hourly Global Lightning CLimatology, Bull. Am. Meteorol. Soc., 94, 1381–1391, 2013.

Vogel, S., Holbøll, J., López, J., Garolera, A. C., and Madsen, S. F.: European Cold Season Lightning Map for Wind Turbines Based on Radio Soundings, in: 2016 33rd International Conference on Lightning Protection (ICLP), pp. 1–7, https://doi.org/10.1109/ICLP.2016.7791373, accessed 2021-05-24, 2016.

Wapler, K.: High-Resolution Climatology of Lightning Characteristics within Central Europe, Meteorology and Atmospheric Physics, 122, 175–184, https://doi.org/10.1007/s00703-013-0285-1, accessed 2021-05-04, 2013.

Wapler, K. and James, P.: Thunderstorm Occurrence and Characteristics in Central Europe Under Different Synoptic Conditions, Atmospheric Research, 158-159, 231–244, https://doi.org/10.1016/j.atmosres.2014.07.011, accessed 2022-11-08, 2015.

Wilkinson, J. M., Wells, H., Field, P. R., and Agnew, P.: Investigation and Prediction of Helicopter-Triggered Lightning Over the North Sea, Meteorological Applications, 20, 94–106, https://doi.org/10.1002/met.1314, accessed 2022-11-08, 2013.

Zhang, W., Zhang, Y., Zheng, D., Xu, L., and Lyu, W.: Lightning Climatology Over the Northwest Pacific Region: An 11-Year Study Using Data From The World Wide Lightning Location Network, Atmospheric Research, 210, 41–57, https://doi.org/10.1016/j.atmosres.2018.04.013, accessed 2022-11-08, 2018.