Abstract  Lightning flashes are potentially hazardous, albeit locally rare events. Despite this rareness, generalized additive models (GAMs) have succeeded in producing a climatology of lightning occurrence for the eastern Alps and surrounding lowlands with an unprecedented resolution of 1 km$^2$ for each day from April to September based on data from the ALDIS lightning location system (2010–2020). This resolution is achieved due to the GAM incorporating information from cells with similar characteristics, such as region, altitude, season, and elevations within the cell. The probability of a cloud-to-ground discharge within 1 km$^2$ on a given day during the warmer seasons is typically less than 1%, with a rapid increase in spring, followed by a plateau in summer and a gentle decrease in fall. Early in the season, probabilities are lower in the highest regions but increase once the snow cover is gone, becoming higher than in the valleys. Regional patterns of lightning probability also vary with season. The details are complex, but generally, higher values shift towards the south in the course of the season. Grid cells with a jagged topography have a higher probability of lightning.

Keywords  Lightning location system · Climatology of lightning occurrence · Generalized additive model · Complex terrain · Machine learning · Alps · ALDIS

Ostalpine Blitzklimatologie auf der Kilometerskala in Tagesauflösung

Zusammenfassung  Blitze sind potenziell gefährliche, aber auch lokal seltene Ereignisse. Trotz dieser Seltenheit ist es gelungen, mithilfe von verallgemeinerten additiven Modellen (GAM) eine Blitzklimatologie für die Ostalpen und das umliegende Vorland in einer bisher unerreichten Auflösung von 1 km$^2$ für jeden einzelnen Tag von April bis September basierend auf Daten des ALDIS-Blitzmessnetzes (2010–2020) zu erstellen. Diese Auflösung wird dadurch ermöglicht, dass ein GAM auch Informationen aus Zellen mit ähnlichen Eigenschaften wie Region, Höhe, Jahreszeit und Erhebungen innerhalb der Zelle mitverwendet. Die Wahrscheinlichkeit einer Blitzentladung in 1 km$^2$ an einem bestimmten Tag während der warmen Jahreszeiten liegt typischerweise unter 1%. Sie steigt im Frühjahr rasch an, erreicht ein breites Plateau im Sommer und fällt im Herbst sanft ab. Blitzwahrscheinlichkeiten in den höchsten Regionen sind zu Saisonbeginn niedriger als in den Tälern und im Vorland, nehmen dann aber zu, sobald die Schneedecke geschmolzen ist, und übersteigen schließlich die Werte in den Tälern. Die regionalen Muster der Blitzwahrscheinlichkeiten verändern sich ebenfalls saisonal. Die Details sind complex, höhere Werte wandern aber generell im Lauf der Saison südwärts. Das Vorhandensein von prägenden Erhebungen innerhalb einer Gitterzelle erhöht die Blitzwahrscheinlichkeit.

Schlüsselwörter  Blitzmessnetz · Blitzklimatologie · Verallgemeinertes additives Modell · Komplexes Gelände · Maschinelles Lernen · Alpen · ALDIS
1 Introduction

Lightning affects many fields of research and aspects of everyday life. Cloud-to-ground lightning strikes may damage equipment and structures such as wind turbines [1, 17] and power lines [4], start fires [6, 22] and injure or kill people [9, 23]. It produces NO\textsubscript{x}, which in turn affects the concentration of greenhouse gases [19]. While lightning during cold season occurs with an unusually strong wind field [18], it is closely connected to strong convection in warm season, which adds flash floods, large hail and damaging winds as further hazards. Reliable climatologies of lightning thus aid in assessing all these risks and the understanding of processes associated with strong convection.

Lightning location systems (LLS) measure lightning discharges continuously in both space and time unlike any other atmospheric measurement system. Detection efficiencies often exceed 90% with location accuracies well below 1 km [e.g. 20]. Lightning climatologies are usually based on a “cell-count” method, where an area and a period span each cell. However, at any given point, lightning is a rare event with typically only a few discharges per square kilometer over the whole year. [5] treated lightning as a Poisson-process to provide a theoretical limit for a 90% confidence interval of the flash density estimate with the cell-count method and found that approximately 80 flashes in a cell over the whole observation period are needed for this interval width to be ± 20%. Previous climatologies therefore required an aggregation to large cells, typically about 10–100 km\textsuperscript{2} times 1–12 months. We will use the same spatio-temporal definition of “cell” throughout the paper but carry it to a much finer resolution.

Two promising approaches have been used to achieve spatial cell sizes of 1 km\textsuperscript{2} and temporal sizes of 1 month or shorter. The first one [2, 11] uses the location uncertainty instead of the precise locations of flashes, which corresponds to performing a kernel density estimation around the location of lightning discharge under an assumed independence between the discharges. The second one [27] exploits the similarity in seasonal, regional and altitudinal characteristics between cells using the statistical method of generalized additive models [GAMs, 34] to achieve even smaller cell sizes of 1 km\textsuperscript{2} times 1 day and additionally extract the functional dependence of lightning on these underlying characteristics. Their study was intended as a proof of concept for the suitability of GAMs to achieve high-resolution lightning climatologies and thus limited to a small region (the state

| Name          | Latitude | Longitude | Elevation [m] | Roughness [m] |
|---------------|----------|-----------|---------------|---------------|
| A Saentis     | 47.2494  | 9.3433    | 2214          | 271           |
| B Munich      | 48.1513  | 11.4921   | 575           | 13            |
| C Gaisberg    | 47.8056  | 13.1125   | 1063          | 268           |
| D Bohemian Forest | 49.1182 | 13.3071   | 1235          | 100           |
| E North of Vienna | 48.4452 | 15.8566   | 268           | 11            |
| F Central Alps N | 46.8854 | 10.8672   | 3593          | 155           |
| G Central Alps S | 46.3654 | 10.5872   | 3213          | 270           |
| H Range NW of Udine | 46.1327 | 12.5675   | 1545          | 127           |
| I Udine       | 46.0301  | 13.0013   | 127           | 5             |
| J Graz basin  | 46.9968  | 15.6558   | 479           | 49            |
of Carinthia in Austria) and basic effects of season, elevation and region.

We will use the GAM method to compute a lightning climatology for the entire eastern Alpine region and the surrounding lowlands and include a more comprehensive set of shared characteristics between cells. To our knowledge it is the first lightning climatology at a cell resolution of 1 km² times 1 day for such a large region. Other climatologies for (parts of) this region [e.g. 7, 31, 33] have a considerably coarser—either in space and/or in time—cell resolution.

2 Data

The climatology of lightning days is based on lightning detection data from the Austrian Lightning Detection & Information System [ALDIS, 26], complemented by data from the digital elevation model TanDEM-X [24] to account for topographic effects.

2.1 Lightning location data

For this study we use ALDIS cloud-to-ground flashes for the summer months April–September. Eleven years of data, 2010–2020, have been on hand. The ALDIS data are cropped slightly to a domain that covers around a quarter of a million square kilometers (Fig. 1). The domain includes the Eastern Alps and extends northwards to the lower mountain ranges of the Swabian and Franconian Jura and the Bohemian Forest.

The spatio-temporal dimensions of a cell used for our climatology are an area of 1 square kilometer and a period of 24 hours (one day). The shape of the 1 square kilometer area is taken to be hexagonal and a day is taken to start at 06 UTC, the approximate time of the diurnal lightning minimum. The study region has an area of 248 308 km² and the study period from one day is taken to start at 06 UTC, the approximate time of the diurnal lightning minimum. The shape of the underlying process, we utilize a GAM [34] that incorporates data not only for single data points but leverages information of data points with similar settings, e.g. same region, similar time of the year, similar topographic conditions, which are represented by a choice of explanatory variables. For these explanatory variables non-linear transformations are learned from the data. GAMs have proven their abilities for lightning climatologies in complex terrains [27]. The GAM used for this study is set up as follows,

\[
\logit(\pi) = \beta_0 + f_1(y_{\text{day}}) + f_2(\log(\text{topo.y}_{\text{day}})) + f_3(\text{lon.lat.y}_{\text{day}}) + f_4(\log(\text{roughness})).
\]

The natural logarithm of the mean topography (Fig. 1a) will enter the statistical climatology model as explanatory variable. Note that both the mean topography reveals a right skewed distribution. After transformation with the natural logarithm the variable is closer to a normal distribution which eases its usage within the statistical model. Further, the natural logarithm of the difference between the maximum and the average topography serves as proxy for the roughness of the terrain within a hexagonal square kilometer (Fig. 1b). Though various proxies for the roughness of the terrain are conceivable, we prefer a proxy that especially takes terrain peaks into account as terrain peaks favour thunderstorm formation.

3 Methods

Each of the 45.2 million spatio-temporal cells—1 square kilometer hexagons spatially times 1 day temporally for each day from April through September—is one data point for our statistical climatology. Each cell contains the number of observed thunderstorm days over the 11 years for this specific day of the year and hexagonal area. These numbers of observed thunderstorm days {0, 1, 2, ..., 11} can be seen as realizations of Bernoulli’s urn problem with replacement. For each cell, 11 trials are conducted, where each trial has two possible outcomes: lightning and no lightning, respectively.

The outcome of these trials follows a binomial distribution, which is determined by a probability \( \pi \) that lightning occurs. In order to estimate the probability \( \pi \) of the underlying process, we utilize a GAM [34] that incorporates data not only for single data points but leverages information of data points with similar settings, e.g. same region, similar time of the year, similar topographic conditions, which are represented by a choice of explanatory variables. For these explanatory variables non-linear transformations are learned from the data. GAMs have proven their abilities for lightning climatologies in complex terrains [27]. The GAM used for this study is set up as follows,

\[
\logit(\pi) = \beta_0 + f_1(y_{\text{day}}) + f_2(\log(\text{topo.y}_{\text{day}})) + f_3(\text{lon.lat.y}_{\text{day}}) + f_4(\log(\text{roughness})).
\]

On the left hand side is the logit-transformed probability \( \pi \). The logit maps \( \pi \) from the probability scale \([0,1]\) to a real number \([-\infty, \infty]\). On the right hand side is the additive predictor that combines the intercept \( \beta_0 \) of the regression and multiple, potentially non-linear terms \( f_\star \).

- \( f_1(y_{\text{day}}) \) is a function of the day of the year and thus represents a baseline annual cycle valid for all locations of the domain.

354 Lightning climatology for the eastern Alpine region on the kilometer scale with daily resolution

\[\sigma_{\beta_0} = \sigma_{\beta_0} \]

\[\sigma_{\beta_0} = \sigma_{\beta_0} \]

\[\sigma_{\beta_0} = \sigma_{\beta_0} \]
• $f_2(\log\text{topo}, \text{yday})$ is a function of the log-topography and the day of the year. Thus this term allows for deviations of the annual cycle from its baseline conditioned on the topography.
• $f_3(\text{lon}, \text{lat}, \text{yday})$ additionally allows deviations from the baseline annual cycle conditioned on the geographical location.
• $f_4(\log\text{roughness})$ adds a correction for the roughness of the topography.

With the GAM all these functions $f$, are modelled using regression splines such as cubic regression splines and thin plate regression splines. The spline bases for functions with two ($f_2$) or three ($f_3$) covariates are set up using the tensor product of the univariate spline bases. For the technical details on splines within GAMs the reader is referred to [34].

This setup leads to approximately 100 regression coefficients of the GAM. These coefficients are estimated via penalized maximum likelihood. Here the amount of smoothing of the potentially non-linear functions $f_*$ is determined by generalized cross-validation as implemented in the mgcv extension for R. Estimating such a flexible regression model for the large data set on hand with 45,192,046 data points is feasible with a fitting algorithm for giga data [35, implemented in the mgcv extension for R].

The resulting probabilities $\pi$ can be converted into expected number of thunderstorm days within $k$ years at a specific location. Since one can assume that—for a specific day of the year and location—the occurrence of lightning from one to the next year is independent from each other, the expected number of thunderstorm days within $k$ can be computed by the product,

$$E(\text{ts\_days}) = k \cdot \pi.$$  

4 Results

Lightning in the eastern Alps is rare locally. Only 5.47% of the 45.2 million spatio-temporal cells of 1 km$^2$ times 1 day in the 182 days from months April through September had lightning occurrence in at least one of the eleven years 2010–2020. Fig. 2 shows that the maximum probability for lightning to strike within a grid box of 1 square kilometer on a particular day barely exceeds 2%. The maps for day 15 of each month demonstrate that the probability of lightning waxes from spring through the end of July and then wanes into autumn. In spring (May 15, top) the likelihood for lightning strikes north of the Alps is regionally uniform below about 0.5%. The probability along the northern Alpine rim slightly exceeds these values. In the Alps, valleys and mountain ranges stand out (cf. Fig. 1, top) with probabilities in valleys considerably higher than over mountain ranges, which are still mostly snow-covered at this time of the year. By mid-June these differences blur and valleys are almost no longer distinguishable from mountain ranges. By now, lightning in the higher terrain north of the Alps strikes more readily than in the surrounding flat lands. Differences south of the Alps are even starker. The flat land around location I (Udine) in Fig. 1 has the highest probability in the whole study region. This is exceptional as lower terrain everywhere else has lower probabilities than higher terrain in its vicinity. By mid-July light-
ning activity favors higher terrain. Lightning probability again clearly differs between valleys and mountain ranges, however now reversed to mid-May. The probability in valleys is considerably lower than over the mountains. Overall, the strongest activity has shifted to the south of the Alpine crest and probabilities even exceed the ones in the former hotspot in the flat land in the vicinity of location I (Udine). One month later in mid-August the overall pattern is the same but probabilities are lower—with the exception of location I (Udine) where the probabilities remain relatively high even into mid-September, by which time lightning activity is low everywhere, most pronouncedly so north of the Alps.

Contrary to the results from the GAM method, results from the cell-count method in Fig. 2b are noisy and lack details despite averaging the results over an 11-day window centered on that particular day. Grid boxes with probabilities of more than 10% occur next to boxes with 0%, which necessitated a different color scale in the figure. Only the major features of lower lightning probabilities/frequencies over the high mountains early in the lightning season, the shift into the high mountains and south of the Alpine crest, and the hotspot near location I (Udine, Fig. 1) are visible, admittedly more easily so when one has previously seen the results from the GAM method. Sample sizes in the grid boxes are simply far too small to obtain reliable frequencies of lightning occurrence, which is a rare event of only about 1% per day in a 1 square-kilometer grid box. The uncertainty of the cell-count method can be determined using the modified Wilson method \[3\] to estimate the width of the 95% confidence interval for a binomial distribution (lightning yes/no) for a sample size of 11—the number of years in our data set. The width is a staggering 27% for an event with a probability of 1%. Even increasing the sample size by a factor of 11 to \[11 \times 11 = 121\] using a time-window of 11 days instead of 1 day, as is done in Fig. 2b, still yields a width of 4.9 percentage points. Therefore finding a lot of noise in the results from cell-count method comes as no surprise.

The power of the GAM method and its ability to produce detailed and smooth results even at such high spatial and temporal resolution lies in its harnessing of information from grid boxes that share seasonal, topographical and regional characteristics. Extracting the functional form of these characteristics in turn leads to increased understanding of the contributors to the overall spatio-temporal distribution of the rare event “lightning occurrence”. Fig. 3 shows the effects of each of the additive terms in Eq. 1 as they contribute to \[\text{logit}(\pi)\]. Seasonally (Fig. 3a), the probability of lightning for the whole region increases rapidly in spring and slowly during mid-summer, then tapers off in autumn.

The elevation effect (Fig. 3b) varies seasonally. Note that the effect is logarithmic (cf. Eq. 1) but the ordinate axis has been labeled in meters to ease interpretation. This effect shows a reduced likelihood for lightning at the highest elevations early in the season, which is reversed in the second half of June. The sign reverses last at the very highest elevations, where snow melts last. Snow-covered ground greatly reduces sensible heat flux and thus the chances for convection to occur. The seasonal elevation effect also changes sign in elevations typical of intra-alpine valleys. It is increased in spring, when the valleys are already free of snow, and reduced in summer and early autumn. The elevation effect is positive from mid-May onwards for the lowest elevations, found mostly at the southern edge of the study region.
The regional effect in Fig. 3c accounts for adjustments needed to add to the effects of season and elevation. At first glance, the patterns do not seem to vary seasonally but a closer look reveals a shift of enhanced and reduced regions from spring into summer and fall. In spring the lightning probability in some parts of the lowlands and rims on either side of the Alps and the high mountains in southeastern Switzerland is enhanced. By mid-July lightning activity along the whole southern side of the Alps is enhanced as well as in the westernmost and easternmost parts of the lowlands north of the Alps. By mid-September, at the end of convective season in most of the region, the northwestern rim and lowlands and especially the southeastern rim and lowlands have enhanced lightning probabilities.

Protruding topography such as peaks increases the likelihood of lightning. Again, the effect is logarithmic but elevation difference in Fig. 3d is given in units of meters. Here the difference between the maximum elevation at 90 m horizontal resolution to the average elevation in the one-square-kilometer grid box serves as proxy for how jagged the terrain is. Overall, where that difference exceeds about 100 m, lightning likelihood increases somewhat, which is in alignment with findings from lightning research [e.g. 7, 11, 21].

Time series of lightning probability at specific locations in the eastern Alps in Fig. 4 exemplify the sum of the seasonally varying regional and elevation effects and the roughness effect from Fig. 3 and fill in the times between the regional maps of Fig. 2. These locations are shown in Fig. 1. They represent the north side and the northern rim of the Alps (A–E) both in flatlands and high terrain. The two exposed peaks Saenits (A) and Gaisberg (C) are included since measurements from the instrumented towers there have featured prominently in numerous papers. Locations F and G straddle the Alpine crest, and H–J exemplify the southern rim of the Alps. Locations I and H in the Alpine lightning hotspot represent flatland and adjacent mountain range, respectively.

The lightning season starts about three weeks later at the highest mountains, exemplified by locations F (Central Alps N) and G (Central Alps S), than at other locations (cf. Fig. 1 and Table 1 for their coordinates). The difference in probability between the two locations F (Central Alps N) and G (Central Alps S) stems from their difference in elevation. Similar peak probabilities are reached by locations north of the Alps at A (Saentis) and D in the Bohemian Forest, although they are considerably lower. Their seasonal cycles are shifted by about one week, with D (Bohemian Forest) starting out one week earlier and A (Saentis) lasting one week longer. Lightning season at location C (Gaisberg), just south of D (Bohemian Forest) at the northern rim of the Alps starts about at the same time but peak probabilities are about one third lower. Both A (Saentis) and C (Gaisberg) are locations with a long history of direct lightning measurements from instrumented towers. Although both are at the northern rim of the Alps, seasonally varying regional differences (Fig. 3c) and an elevation difference of 1.1 km result in different shapes and peaks of their seasonal lightning probability distributions. Flat terrain north of the Alps, exemplified by locations B (Munich) and E (North of Vienna), experiences lightning less frequently than the mountainous counterparts A (Saentis), C (Gaisberg) and D (Bohemian Forest). They also reach a first peak in probability earlier and then increase slightly to a second peak at the end of July. Locations in the flat lands south of the Alpine crest (E – North of Vienna – and I – Udine) show a similar bimodal distribution. Location I (Udine), which is situated in the flatlands south of the Alps has the most abrupt increase of lightning probability of all 10 locations shown. Despite its low elevation, its peak probabilities exceed the ones of the highest mountains near the Alpine crest. It is located in the hotspot so clearly visible in the regional probability maps in Fig. 2. However, probabilities at location H (Range NW of Udine) at the first mountain range north of I (Udine) are even higher.

5 Discussion and Conclusions

Lightning location systems (LLS) are unique among atmospheric measurement systems in that they provide continuous measurements in both time and space—two-dimensional or three-dimensional depending on LLS type. Since they measure events that are rare and infrequent at any given location,
translating these measurements into high-resolution climatologies is difficult. Most climatologies so far have used the cell-count method of counting how many flashes occurred within a particular time-space cell, which necessitates making the spatial and/or temporal dimensions of such a cell large in order to achieve a sufficient signal to noise ratio. Frequently cell sizes of 10–100 square kilometers by 1–12 months have been used [e.g. 7, 20, 25]. [2, 11, 27] demonstrated how the cell-count method is unsuitable for smaller cell sizes corroborating the theoretical considerations in [5], who used the Poisson distribution to derive the minimum number of flashes in a cell required for reaching a specified accuracy. The Poisson distribution is the limiting case of the binomial distribution, which describes the binary event of the occurrence of lightning (yes/no). We therefore use the binomial distribution to compute the 95% confidence interval for a lightning probability of 1%. It has a width of staggering 27% for a sample size of 11 for a particular day in our 11-year data set. Increasing the time dimension of a cell from 1 day to averaging over 11 days to obtain a sample size of 121 still gives an uncertainty of 4.9%, five times larger than what we need to detect. Fig. 2b illustrates the high noise level the cell-count method produces for high-resolution cells of 1 km² times 1 day. Fig. 2a, on the other hand, produced with the generalized additive model (GAM) method, is devoid of noise and yet provides intricate spatial details for any given day in the lightning season (of which the map shows five). Similarly spatially detailed climatological maps can be achieved using a bivariate Gaussian error distribution for each lightning discharge instead of a precise location [2, 11].

The advantage of the GAM method is its ability to harness regional, elevational and seasonal characteristics that cells share that need not be immediately “next” to another and thus make a high combined spatio-temporal resolution possible. To our knowledge, this paper presents the first daily-resolved kilometer-scale lightning climatology over a large region. GAMs have another advantage: the resulting functional forms can be explicitly included since it suppresses convection. In our analysis it is implicitly contained in the seasonally varying elevation effect where lightning season at the highest elevations with the longest snow cover has a higher probability when the snow has melted. The most pronounced increase on the southern side of the Alps (vicinity of location I – Udine) with maxima in June and late August/early September is found at the very lowest elevations.

This regional hotspot was already known from earlier cell-count climatologies [e.g. 7, 25, 31] and is possibly linked to the propagation and moisture transport of the daytime seabreeze front from the Adriatic Sea [14, 32]. With GAMs we could additionally identify how regional differences vary with season—in extension to [27]—when different elevations are already accounted for.

The approach of using similar characteristics to compute high-resolution yet non-noisy lightning climatologies with GAMs pioneered by [27] can be expanded by including further characteristics beyond region, elevation and time of year. Snow cover could be explicitly included since it suppresses convection. In our analysis it is implicitly contained in the seasonally varying elevation effect where lightning season at the highest elevations with the longest snow cover duration has a delayed start. Further characteristics of earth’s surface such as slope angle and exposure, soil moisture, amount and type of vegetation cover, and land/water could be included as well as the presence of tall structures, which can trigger lightning [21]. The GAM approach can also use further additive terms with output from numerical weather prediction models in order to provide forecasts of the probability of thunderstorm occurrence [28].

Our GAM models the probability whether lightning occurs on a given specific day of the season and location. These probabilities can be intuitively converted to expected number of thunderstorm days as shown by the alternative scales in Figs. 2 and 4. Flash densities could also be estimated with a GAM, though this would require a different set up: The ALDIS data have to be aggregated to counts of cloud-to-ground discharges in each cell. These counts would serve as target variable in the GAM and thus a parametric count data distribution has to be applied for the estimation. As several parametric count data distributions have been proposed, e.g. Poisson [27] or negative binomial [29], further research is needed to find amounts of moisture and heat that provide lifting and destabilization of the atmosphere conducive to the formation of convection [e.g. 8]. Valley and slope winds have similar effects on smaller scales. Mountains can also modify larger-scale flow to create favorable conditions for convection [e.g. 10]. Since both thermally and mechanically driven flows vary seasonally it is highly advantageous that the GAM method makes it possible to implement a seasonally varying elevation effect. It finds an increased probability in spring at elevations in lowlands and intra-Alpine valleys compared to higher elevations (above approx. 1200 m msl), which are still (partly) snow covered. This pattern reverses later and high elevations have a higher probability when the snow has melted. The most pronounced increase on the southern side of the Alps (vicinity of location I – Udine) with maxima in June and late August/early September is found at the very lowest elevations.

The purely seasonal effect (Fig. 3a) shows a rapid increase of lightning probability in late spring but a more gradual tapering off in late summer and early fall, which was previously also found for a smaller region in the eastern Alps [27]. The peak, however, is less sinusoidal but more plateau-like, probably due to the larger region where different subdomains peak at different times (cf. Fig. 4).

Lightning probability also varies with elevation (Fig. 3b). [30] and [27] found an overall increase with elevation. This is due to thermally and mechanically forced lifting along higher topography [e.g. 12]. Daytime winds from the plains to the mountains (“Alpine pumping” 15, 16] lead to updrafts and additional

Originalarbeit
the distribution that fits best the data on hand. Another challenge is handling the data size as the count data distributions require a different optimization approach for the class of distributional regression models [13].

6 Computational Details

The results in this study were achieved using R, a software environment for statistical computing and graphics. The add-on packages mgcv were used for building the statistical model.

Acknowledgements This work was funded by the Austrian Science Fund (FWF, grant no. P31836) and the Austrian Research Promotion Agency (FFG, grant no. 872656). We thank the two reviewers for their careful and constructive feedback. Finally, we thank Gerhard Diendorfer and Wolfgang Schulz from ALDIS for a fruitful and pleasant collaboration on lightning-related research and for providing the data for this study.

Funding Open access funding provided by University of Innsbruck and Medical University of Innsbruck.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence and indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material, are not included in the article’s Creative Commons licence and must be obtained from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Becerra M, Long M, Schulz W et al (2018) On the estimation of lightning incidence to offshore wind farms. Electr Power Syst Res 157:211–226. https://doi.org/10.1016/j. EPSR.2017.12.008
2. Bourscheidt V, Pinto O, Naccarato KP (2014) Improvements on lightning density estimation based on analysis of lightning location system performance parameters: Brazilian case. IEEE Trans Geosci Remote Sens 52(3):1648–1657. https://doi.org/10.1109/TGRS.2013.2253109
3. Brown LD, Cai TT, DasGupta A (2001) Interval estimation for a binomial proportion. Stat Sci 16(2):101–133. https://doi.org/10.1214/ss/1009213286
4. Cummins K, Krider E, Malone M (1998) The US National Lightning Detection Network and applications of cloud-to-ground lightning data by electric power utilities. IEEE Trans Electromagn Compat 40(4):465–480. https://doi.org/10.1109/15.736207
5. Diendorfer G (2008) Some comments on the achievable accuracy of local ground flash density values. 29th International Conference on Lightning Protection, Uppsala, Sweden, 23–26 June 2008.
6. Dowdy AJ, Mills GA (2012) Atmospheric and fuel moisture characteristics associated with lightning-attributed fires. J Appl Meteorol Climatol 51(11):2025–2037. https://doi.org/10.1175/jamc-d-11-0219.1
7. Feudale L, Manzato A, Micheletti S (2013) A cloud-to-ground lightning climatology for north-eastern Italy. Adv Sci Res 10(1):77–84. https://doi.org/10.5194/asr-10-77-2013
8. Heim C, Panosetti D, Schlemmer L et al (2020) The influence of the resolution of orography on the simulation of orographic moist convection. Mon Weather Rev 148(6):2391–2410. https://doi.org/10.1175/mwr-d-19-0247.1
9. Holle RL (2016) A summary of recent national-scale lightning fatality studies. Weather Clim Soc 8(1):35–42. https://doi.org/10.1175/WCAS-D-15-0032.1
10. Houze RA (2012) Orographic effects on precipitating clouds. Rev Geophys 50(1):1–47. https://doi.org/10.1029/ 2011RG000365
11. Kingfield DM, Calhoun KM, de Beurs KM (2017) Antenna structures and cloud-to-ground lightning location: 1995–2015. Geophys Res Lett 44(10):5203–5212. https://doi.org/10.1002/2017gl073449
12. Kirshbaum D, Adler B, Kalthoff N et al (2018) Moist orographic convection: physical mechanisms and links to surface-exchange processes. Atmosphere 9(3):80. https://doi.org/10.3390/atmos9030080
13. Kneib T, Silbersorff A, Säkken B (2021) Rage against the mean—a review of distributional regression approaches. Econometr Stat. https://doi.org/10.1016/j.ecosta.2021.07.006
14. Kotroni V, Lagouvardos K (2016) Lightning in the Mediterranean and its relation with sea-surface temperature. Environ Res Lett 11(3):34006. https://doi.org/10.1088/1748-9326/11/3/034006
15. Langhans W, Schmidli J, Fuhrer O et al (2013) Long-term simulations of thermally driven flows and orographic convection at convection-parameterizing and cloud-resolving resolutions. J Appl Meteorol Climatol 52(6):1490–1510. https://doi.org/10.1175/jamc-d-12-0167.1
16. Lugauer M, Winker P (2005) Thermal circulation in south bavaria climatology and synoptic aspects. Meteorol Z 14:15–30. https://doi.org/10.1127/0941-2948/2005/0014-0015
17. Montanya J, van der Velde O, Williams ER (2014) Lightning discharges produced by wind turbines. J Geophys Res Atmos 119(3):1455–1462. https://doi.org/10.1002/2013jd020225
18. Morgenstern D, Stucke I, Simon T et al (2022) Differentiating lightning in winter and summer with characteristics of wind-field and mass-field. Weather Clim Dyn. https://doi.org/10.5194/wcd-2021-68
19. Murray LF (2016) Lightning NOx and impacts on air quality. Curr Pollut Rep 2(2):115–133. https://doi.org/10.1007/ s40726-016-0031-7
20. Poelman DR, Schulz W, Diendorfer G et al (2016) The European lightning location system EUCLID—part 2: observations. Nat Hazards Earth Syst Sci 16(2):607–616. https://doi.org/10.5194/nhess-16-607-2016
21. Rakov VA, Uman MA (2003) Lightning: physics and effects. Cambridge University Press, Cambridge https://doi.org/10.1017/CBO9781107340886
22. Reineking B, Weibel P, Conedera M et al (2010) Environ mental determinants of lightning- v. human-induced forest fire ignitions differ in a temperate mountain region of Switzerland. Int J Wildland Fire 19(5):541–557. https://doi.org/10.1071/WF08206
23. Ritenour AE, Morton MJ, McManus JG et al (2008) Lightning injury: a review. Burns 34(5):585–594. https://doi.org/10.1016/j.burns.2007.11.006

Springer Lighting climatology for the eastern Alpine region on the kilometer scale with daily resolution 359
24. Rizzoli P, Martone M, Gonzalez C et al (2017) Generation and performance assessment of the global TanDEM-X digital elevation model. ISPRS J Photogramm Remote Sens 132:119–139. https://doi.org/10.1016/j.isprsjprs.2017.08.008
25. Schulz W, Cummins K, Diendorfer G et al (2005) Cloud-to-ground lightning in Austria: a 10-year study using data from a lightning location system. J Geophys Res 110:D9101. https://doi.org/10.1029/2004JD005332
26. Schulz W, Diendorfer G, Pedeboy S et al (2016) The European lightning location system EUCLID part 1: performance analysis and validation. Nat Hazards Earth Syst Sci 16(2):595–605. https://doi.org/10.5194/nhess-16-595-2016
27. Simon T, Umlauf N, Zeileis A et al (2017) Spatio-temporal modelling of lightning climatologies for complex terrain. Nat Hazards Earth Syst Sci 17(3):305–314. https://doi.org/10.5194/nhess-17-305-2017
28. Simon T, Fabsic P, Mayr GJ et al (2018) Probabilistic forecasting of thunderstorms in the Eastern Alps. Mon Weather Rev 146(9):2999–3009. https://doi.org/10.1175/MWR-D-17-0366.1
29. Simon T, Mayr GJ, Umlauf N et al (2019) NWP-based lightning prediction using flexible count data regression. Adv Stat Clim Meteorol Oceanogr 5(1):1–16. https://doi.org/10.5194/ascmo-5-1-2019
30. Smorgonskiy A, Rachidi F, Rubinstein M et al (2013) On the relation between lightning flash density and terrain elevation. 2013 International Symposium on Lightning Protection (XII SIPDA). https://doi.org/10.1109/sipda.2013.6729216
31. Taszarek M, Allen J, Púčik T et al (2019) A climatology of thunderstorms across Europe from a synthesis of multiple data sources. J Clim 32(6):1813–1837. https://doi.org/10.1175/JCLI-D-18-0372.1
32. Wang CC, Kirshbaum DJ (2015) Thermally forced convection over a mountainous tropical island. J Atmos Sci 72(6):2484–2506. https://doi.org/10.1175/jas-d-14-0325.1
33. Wapler K (2013) High-resolution climatology of lightning characteristics within Central Europe. Meteorol Atmos Phys 122(3–4):175–184. https://doi.org/10.1007/s00703-013-0285-1
34. Wood SN (2017) Generalized additive models: an introduction with R, 2nd edn. Texts in statistical science. Chapman & Hall/CRC, Boca Raton
35. Wood SN, Li Z, Shaddick G et al (2017) Generalized additive models for gigadata: modeling the U.K. black smoke network daily data. J Am Stat Assoc 112(519):1199–1210. https://doi.org/10.1080/01621459.2016.1195744

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Thorsten Simon, graduated 2009 with a Diploma in Physical Oceanography from the Institute of Marine Sciences in Kiel, Germany. In 2015 he received his PhD in Atmospheric Sciences from the University of Bonn, Germany. Scince 2015 he is holding a PostDoc position at the University of Innsbruck, Austria. His research intersects atmospheric sciences and applied statistics with a focus on weather predictions and climate assessment of lightning and convection.

Georg Mayr, received his Master’s degree from the University of Innsbruck, Austria, in 1987 and his PhD from Colorado State University, USA, in 1993. He has held positions at Colorado State University and the University of Innsbruck, where he has been working as an associate professor since 2001. He researches mountain-related weather and climate, convection and lightning, and methods to provide objective end-user forecasts, and blends atmospheric science with statistics and machine learning.