Anonymous referee#1: The manuscript is well written even if some clarifications are required. I think some additional parameters should be inserted in the figures. Please find below some specific comments.

-Line 19. One should not forget that the 3D structure of the flashes might be different in winter compared to the one in summer. **This is correctly pointed out by the referee and will be mentioned in the text as an additional potential reason: "An obvious decrease is observed in the percentage of outliers during May-Sept, compared to the other months of the year. This feature could be related to the fact that more sensor upgrades occur during winter or because precipitation of winter thunderstorms is more difficult to detect with the weather radars. In addition, the 3D structure of lightning flashes in winter compared to summer is somewhat different (López et al., 2017), which could increase the difficulty to locate those in winter accurately. [Full reference: López J.A., Pineda N., Montanya J., van der Velde O., Fabro F., Romero D.: Spatio-temporal dimension of lightning flashes based on three-dimensional Lightning Mapping Array., Atmospheric Research, 197, 255-264, 2017].**

-Line 38. Is there a missing word after "more"? **Correct, the word "important" was missing.**

-Lines 64-67. Why only these two regions? And not a larger domain covered by both EUCLID and the European radars? **The following will be added to the summary: "The latter two regions were chosen specifically for their difference in topography and because high spatial and temporal resolution radar data were readily available. However, a similar approach can be performed in the future on a larger spatial scale, based for instance on the radar composite imagery produced by the Eumetnet Operational Programme for the Exchange of Weather Radar Information (OPERA, Huuskonen et al. 2014) and related EUCLID domain" [Full reference: Huuskonen, A., E. Saltikoff, and I. Holleman, 2014: The Operational Weather Radar Network in Europe, Bull. Amer. Meteor. Soc., 95, 897–907, https://doi.org/10.1175/BAMS-D-12-00216.1]**

-Line 98. Please provide some physical and/or technological explanations on your statement that "timing only sensors often increase the number of outliers". **The following will be added to the text: "Timing only sensors often increase the number of outliers if those are used in solutions determined by two or three sensors only."**

-Line 108. What do you mean by "overshooting beams"? Please rephrase. **The text will include following explanation: "Moreover, since the height of the radar beam above ground increases with increasing distance from the radar, precipitation can be underestimated or even undetected at far range by overshooting when precipitation is produced below the height of the lowest radar beam."**

-Lines 108-113. I suspect the precipitation product you have been using has been validated. It might be relevant to provide some references on such validation in your paper. **We agree completely with the referee. The radars used in this study are well calibrated. This can be done in different ways using the internal test signal, intercomparison of multiple radar observations in oberlapping areas, ... Most of those tests are part of the operational work done at RMI and Austro Control and are not published. Anyway, the validation for Belgium of the product used in the manuscript is presented in Goudenhoofdt and Delobbe (J. Hydromet., 2016) where it is referred to as the QPE1 product. The product used for Belgium is not a quantitative precipitation product. It is actually a reflectivity product, with a simple Z-R relation to produce rainfall rates. The product is used in this study to detect the presence of precipitation and not to estimate the intensity. For more information on the validation of the Austrian composite, the interested reader will be referred to Kaltenboeck and Steinheimer (2015).**

-Line 109. Is the Marshall-Palmer relationship valid whatever the precipitation regime? I suspect in your case you are more interested in low precipitation amount where potentially you might find the lightning outliers. So do you think that the radar product used here is sensitive enough to deliver a reliable and accurate product for your investigation? Why did you choose a radar-based precipitation product and not for example the reflectivity composite? Discharges are not only propagating where precipitation occurs (e.g. spider lightning). So I wonder if your choice to use the radar-based precipitation product does not lead to a larger uncertainty. Do you have any comment? **As written above, the radar products used in this work are actually reflectivity products, with a simple Z-R relation to produce rainfall rates. As a lower threshold a reflectivity of 12dBZ is used, which translates to a rainfall rate of 0.2mm/h using the Marshall-Palmer relation. Note that radars easily can measure much smaller values. However, to eliminate non-meteorological echoes the threshold is set at 12dBZ. Precipitation associated with thunderstorms generally produces much larger reflectivity values. We note that 'bolts from the blue' (which are rare events) will be classified as outliers by our method, but will be so as well using for instance satellite products. Moreover, satellite information will introduce uncertainty as well since cloud tops can shift by strong high altitude winds and layers of high clouds might overlay structures below.**

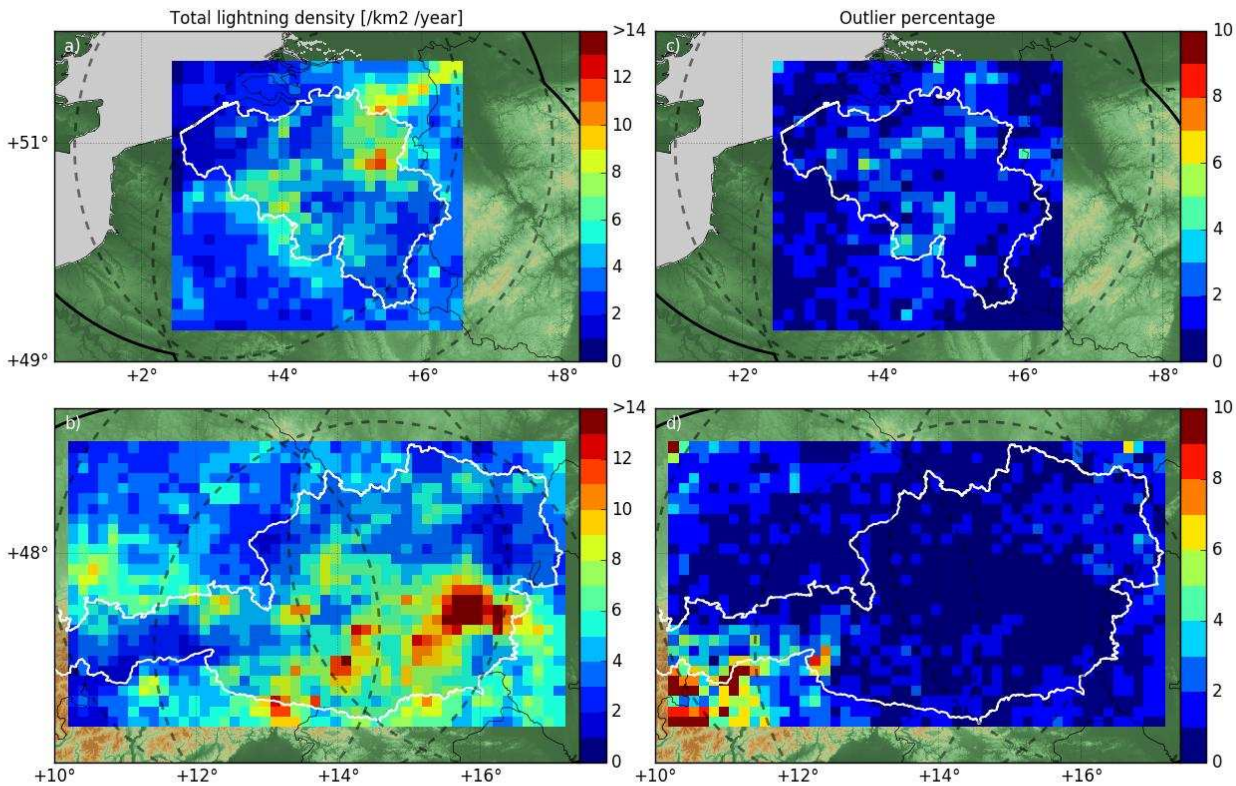
-Lines 116-127. How is the advection taken into account as a 5-min precipitation product is generated? And how do you take into account the advection in your lightning data? At which altitude does the radar-based precipitation product correspond? **No advection correction is performed in Belgium and Austria. However, this is not an issue since different (2, 5, and 10 km) spatial tolerances are used for the distance between precipitation and lightning location. The height of the radar-based precipitation product corresponds to 1500 m above sea level for Belgium. The product used in Austria**

is a vertically maximum surface projection, thus taking any signal into account at any height. This guarantees convective cell detection in complex terrain such as the Alps.

-Lines 130-143. Same questions as for lines 116-127. How is the advection taken into account? At which altitude does the radar-based precipitation product correspond? Is the precipitation product comparable in terms of accuracy for both domains of interest? **See previous comment + the spatial resolution of the radar product is 500m and 1000m in Belgium and Austria, respectively.**

-Line 146. How is the radar-based precipitation distributed in those two domains? Are they geographically uniformly distributed? The precuipitation is not homogeneously distributed over those both areas. **In both regions, precipitation is not homogeneously distributed. More info can be found in Goudenhoofdt and Delobbe (2016) for Belgium and Kaltenboeck and Steinheimer (2015).**

-In Figure 4, you are giving the spatial distribution of the % of outliers. I would have added with iso-contours (in white) the actual lightning distribution from where you computed the %. **Very good idea. However, plotting iso-contours on top of the spatial distribution of percentage of outliers turns out to make the figure hard to read. We therefore opt to change the figure as follows, with on the left panels to total lightning density [/km2/year] and on the right the spatial distribution of the percentage of outliers. For Belgium, densities vary between 0.8 and 11 strokes km^{-2}yr^{-1} with a median value of 3.4 strokes km^{-2}yr^{-1} at 10km x 10km resolution. Overall the densities in Austria are somewhat higher compared to Belgium resulting in a median value of 4.4 strokes km^{-2}yr^{-1}. The highest total lightning densities are found towards the southeast of Austria with a maximum of 22 strokes km^{-2}yr^{-1}.**



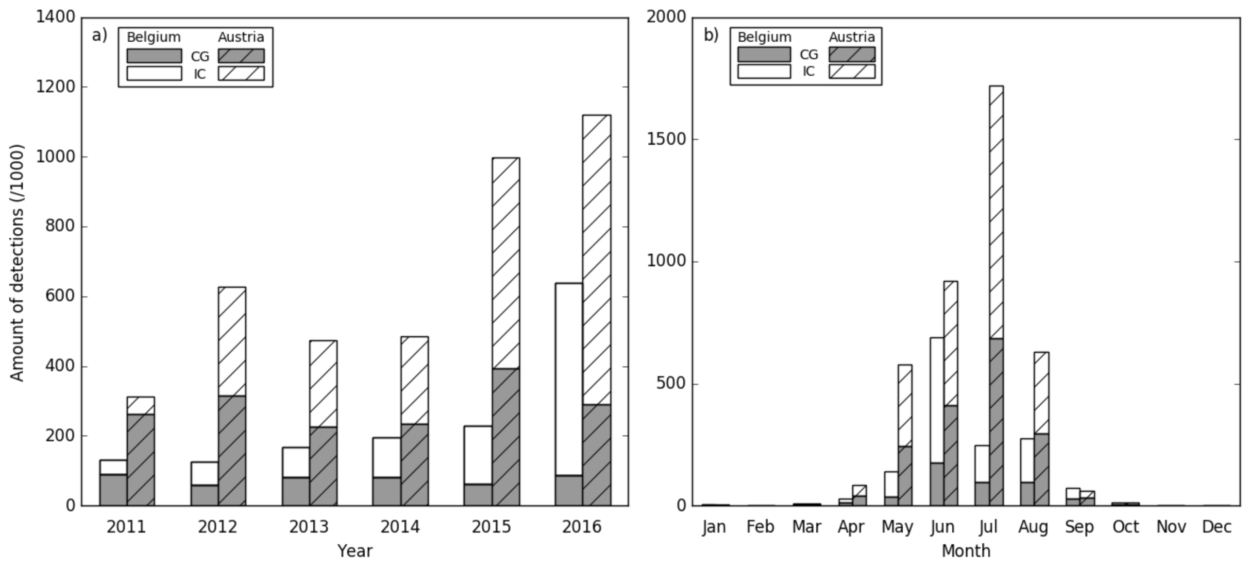
-Lines 148-149. Again how did you take into account the advection of the precipitation and of the lightning activity? The radar product is provided per 5-min period. How did you select the lightning data within that 5 min period and how does it fit with the way the radar product is built? **We have taken the 5min lightning data corresponding to the start and end time of the radar scan. The text will be slightly adapted to make this more clear: "Subsequently, CG strokes and IC pulses with timestamps within the start and end time of the radar scan are superimposed on the corresponding 5-min radar precipitation fields."**

-Lines 151-152. Again this suggests that precipitation at the ground is required to verify the lightning data. How are we sure that precipitation is always required where lightning flashes occur? Please provide some arguments to strengthen your methodology. I think a similar analysis using the reflectivity composite should be performed in order to identify cloud-free outliers to in-cloud non-precipitation outliers. At which stage of the lightning flash do the lightning outliers correspond? **We do make use of "complete" composites, i.e. if one of the radars did not participate in a 5-min composite then this**

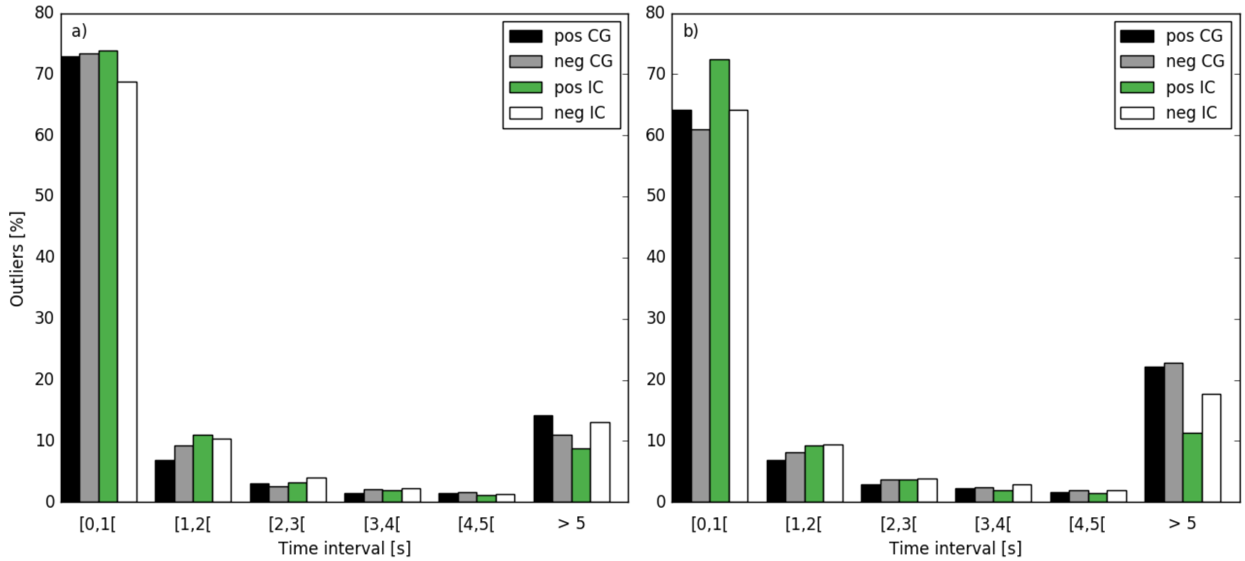
composite is not used at all. Our methodology makes use of radar data, and thus precipitation/reflectivity is required (but not at ground!, see answer related to your remark for Lines 116-127) to discriminate between the outliers and well-located lightning events. However, we admit that lightning produced by “dry” thunderstorms or bolts from the blue will be misclassified as outliers. Nevertheless, we believe that this particular phenomenon is extremely limited. We think that a methodology based on satellite cloudiness products would not allow a proper identification of outliers. Note that the majority of the lightning outliers are single-stroke flashes. The drawback of the used method will now be mentioned in the summary.

-Line 153. I suspect you have projected the lightning data on the same temporal and spatial grid as the radar data, right? **This is Correct.**

-Figure 3. Could you please plot as well the number of CG strokes and IC pulses in order to see how your dataset spreads over the years. **The following figure will be included in the manuscript. It plots the total (CG + IC) amount of detections as a function of a) year and b) month in Belgium and Austria. With regard to the IC detections, one notices a sharp increase in 2016 in Belgium and from 2015 in Austria. This increase is not climatological in nature, but is attributed to the increased amount of LS700x sensors in EUCLID and its capability to detect IC pulses in the low-frequency domain. Looking at the distribution of the total monthly stroke count, it is found that peak in activity is observed in June and July for Belgium and Austria, respectively. For both regions about 95% of all the observed lightning activity occurs between May and September.**



-Lines 161-163. Am I correct when I say that you did not try to associate the outliers with “correct” events based on a time criteria for example? I wonder if a time criteria should not be added to your analysis in order to dissociate isolated events to outliers mis-located from a group of events. **Point well taken. The following figure, which will be added to the paper, plots the distribution of the time difference between the outliers and its closest ‘ok’ event (in time) for a) Belgium and b) Austria. We would argue that all the outliers lying within 1 second of an ‘ok’ event are simply badly located lightning events (~65%-70%), whereas those larger than 1 second (~30-35%) are outliers in time and space or so-called ghost outliers. We find that everything is quite independent of polarity and classification for Belgium and Austria.**



-Line 185. See comment on Line 146 for Figure 4. **Figure 4 has been changed: see above**

-Lines 194-196 Could not we say as well that the October-to-April flashes might have different vertical and horizontal structures that make them detected by EULCID sensors with more difficulties? One should not only criticize the radar sensibility or the sensor upgrade. **True, we will add this as an additional potential reason (see above).**

-Lines 203- 204. I understood that the two geographical domains you have been studying should not really suffer of any long range issue. Am I right? **Correct, the domains are chosen specifically to reduce those kinds of effects to a minimum, if any.**

-Line 208. “which” instead of “wich”. **OK, typo will be adapted in the text.**

-Lines 212-214. I agree with you on the climatology point of view, but then it depends on the application you want to do with your outliers data. **Thanks for your comment. However, we feel this is exactly what is written in L212-213.**

-Lines 217-219. This confirms the interest of considering a temporal criterion in your analysis in order to discriminate isolated (in time and space) outliers to isolated in space only outliers. What do you think? **Thanks for this idea (see fig. above)**

-Figure 6. Without the actual number of events considered for each year, it is difficult to identify how statistically representative is the dataset you are studied in Figure 6. Could you please add that information? **We have added an additional figure showing the number of events per month and per year for both domains (see fig. above)**

-Line 228. What the % of outliers with an absolute current above 20 kA? **“The percentage of CG outliers with absolute peak currents above 20 kA is a factor of about three lower compared to the total percentage of outliers found within |0-20| kA, and is a factor of 15 lower in case of IC. The larger drop in case of IC results from the lower amount of IC events with absolute peak currents larger than 20 kA compared to CG events.”**

-Line 238. I would add “positive current outliers” instead of “positive outliers”. **Thanks for the comment. We will adapt to the terminology of the negative event and write “positive IC and CG outliers”.**

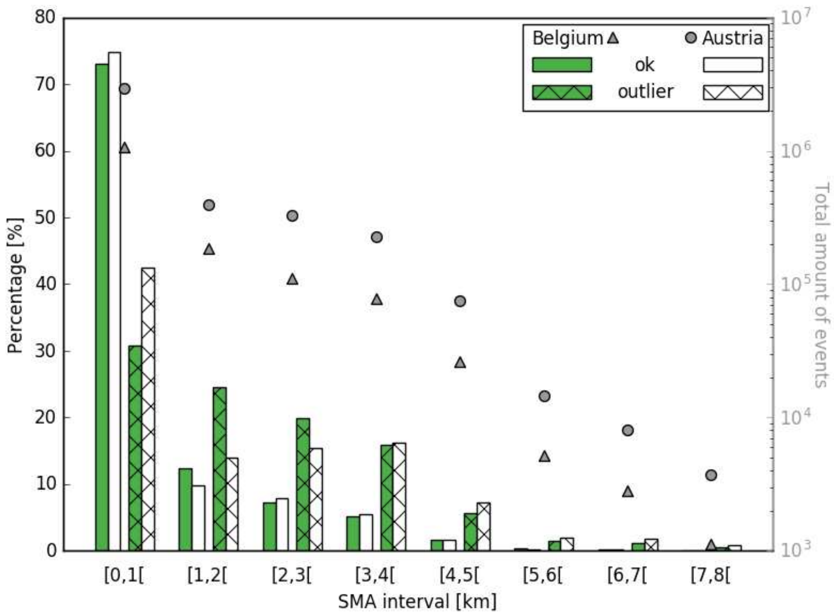
-Lines 228-238. I have a question about all those low current events. How confident are you in their detection and on their classification not only in terms of IC or CG but also in terms of polarity? **We do not present any data on polarity errors because by comparing LLS data with independent E-field measurement data we have never observed such errors since we started those measurements (see Schulz et al. 2016: The European lightning location system EUCLID: Performance analysis and validation). With respect to the IC/CG classification, very recently Kohlmann et al. looked into the classification accuracy (CA) of EUCLID and found that CA varies between 89% and 97% for years 2012 and 2015 in Austria. [reference: H. Kohlmann, W. Schulz, and S. Pedeboy: Evaluation of EUCLID IC/CG classification performance based on ground-truth data; to be presented during the International Symposium on Lightning Protection SIPDA, October 2017]**

-Lines 238-253. How accurate are the event locations when two lightning sensors are only used? Do you usually keep flashes detected with only two EUCLID sensors? Have you plotted the same parameter but by range of current? I would like to see how the number of lightning sensors influences the detection of the outliers according to their estimated current. You could plot it with 2D cumulative distribution. **It is clear that the more sensors participate in a solution, the more accurately detected the lightning event will be. The percentage of events detected by only two sensors is ~10% and ~30% in case of CG and IC, respectively, as can be deduced from Fig. 8. Thus, the majority of the events are detected by more than two sensors. Obviously, the higher the peak current of a lightning event, the higher the number of sensors reporting the event. It is therefore clear that low number of sensors reporting is for low peak currents.**

-Line 257. Do you see any difference between IC and CG outliers separately? **There is a slight difference, which was mentioned in the same paragraph as follows: "Although not shown in this plot, it is found that the average SMA for CG strokes is smaller by a factor of two compared to IC pulses. This is expected since more sensors participate in a solution for CG strokes compared to IC pulses as discussed in Fig. 9."**

-Lines 256-259. Do you have a way to get the number of events that were rejected by the central processor for the period of data you have studied? **Thanks for your comment. However, we don't see exactly what it will bring to our study.**

-Figure 9. The number of samples per SMA range would provide an idea on the statically representativeness of the dataset used here. Please add that parameter in Figure 9. **We have added this information in the following plot:**



-Figure 9 (continued). Similarly to what I suggested for Lines 238-253, have you looked on how the SMA is distributed according to the number of lightning sensors used to locate the different event categories? If yes, what are your main conclusions? If not, please take a look on that plot and provide some information in response to the present comment. **The SMA is indirectly related to the number of sensors reporting. Thus, small SMA values are related to a high number of sensors reporting. From the above plot, it means that ~ 30% and 40% of the outliers exhibit a SMA between [0-1[km for Belgium and Austria, respectively, and are therefore detected by a high number of sensors.**

-Lines 265-267. I do not understand your statement. Please explain it. **To avoid any misunderstanding/confusion and since no results are shown related to $\chi 2$ anyway, we chose to simply remove L266-268 from the text.**

-Lines 280-281. I would insert the actual numbers, i.e. outliers and total number of samples. **The total number of "samples" (as a function of year and month) is now included in an extra plot (see above) and will be mentioned in the text related to the figure.**

%%% End of review

K. Naccarato (Referee)

It is an interesting analysis of lightning solutions provided by the EUCLID network that sometimes do not accurately match the precipitation patterns given by weather radar images. The manuscript is well written, figures are clear and well explained and discussions are comprehensible. Anyway, I have some comments on 3 specific points:

1) In line 84, I really do not understand the sentence: "Note that the latter values are impacted by the strict location quality criteria and correct required stroke classification, i.e. CG versus CG, used in the analysis, as well as temporary sensor outages during the measurements campaign". Please clarify.

We would add following info to the text to clarify what we mean: "To retrieve the latter values, only those strokes are used in the analysis that match certain quality criteria such as χ2, a measure for the correspondence between the different sensor measurements, and semi-major axis of the confidence ellipse, and received a correct stroke classification as CG by the central processor. Those strict criteria, as well as temporary sensor outages during the measurements campaign, can impact the DE estimates given in Schulz et al. (2016). "

2) From line 193 to 213, the authors discuss the results of Figure 5 which mainly shows the seasonal variation of the percentage of outliers. According to the data, clearly during the winter time there is an increase in the number of outliers due to mainly 2 factors: (1) sensor upgrades that provides only TOA solutions during the calibration period; (2) low reflectivity of the precipitating systems due to their smaller size and depth. However, the discussion is confused and I cannot clear understand the apparently 2 opposite effects and their importance (or not): (1) the higher percentage of outliers during winter and (2) the higher absolute number of outliers during summer. This discussion must be rewritten to improve clarity.

- Maybe the confusion was caused by the fact that in L193 (and in the caption of Fig. 5) was written that the "number of outliers" is plotted as well. This is in fact not the true: the absolute number of total detections was plotted. Related to a similar comment of referee 1, we will add following figure to the text, showing the total (CG + IC) amount of detections as a function of a) year and b) month in Belgium and Austria.

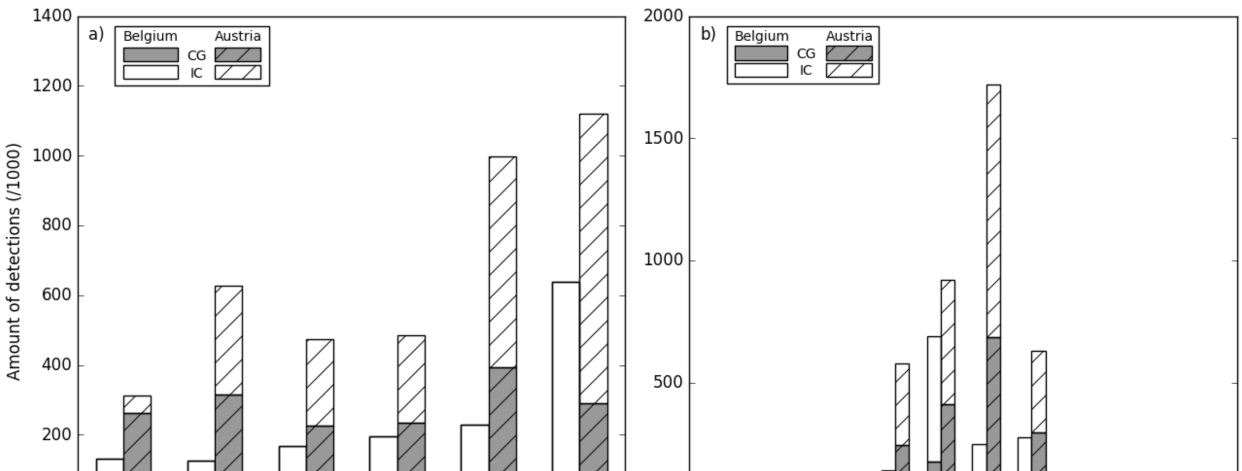


Figure 3: Distribution of the a) annual and b) monthly CG and IC counts as observed within the areas indicated over Belgium and Austria in Fig. 1.

In addition, the original Fig. 5 becomes now:

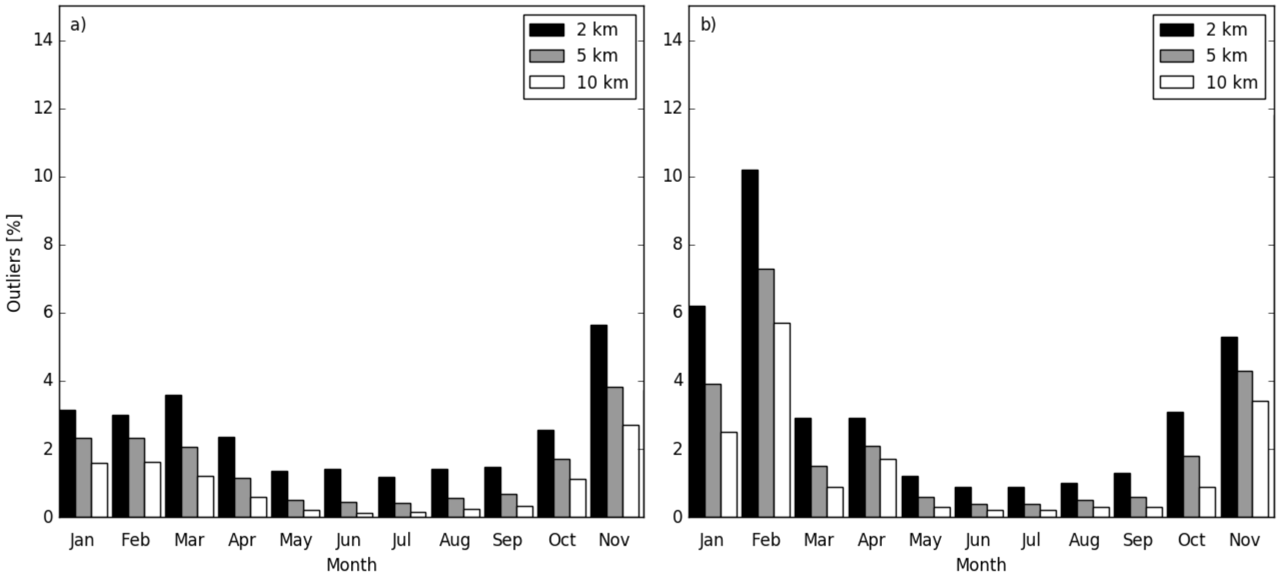


Figure 6: Monthly distribution of the total (CG + IC) percentage of outliers in a) Belgium and b) Austria, for search radii of 2, 5, and 10 km, respectively.

- In order to remove the confusion, we plan to make some changes to the paragraph related to the monthly distribution of outliers as follows: "Fig. 6 illustrates the monthly variation of the percentage of outliers. An obvious decrease is observed in the percentage of outliers during May-Sept, compared to the other months of the year. This feature could be related to the fact that more sensor upgrades occur during winter or because precipitation of winter thunderstorms is more difficult to detect with the weather radars. In addition, the 3D structure of lightning flashes in winter compared to summer is somewhat different (Lopez et al., 2017), which could increase the difficulty to locate those in winter accurately. Regarding the sensor upgrades, those often result in disabled angle information because systematic angle errors, i.e. site errors, are at first unknown and the correction takes a while because lightning data is necessary. Consequently, upgraded sensors start operation with disabled angle information during winter months. With respect to the observation of precipitation, during summer most of the storms are associated with large amounts of precipitation in vertically extended clouds, meaning that those storms are always very well detected by the radars. In contrast, winter storms are generally associated with less intense precipitation cells and with smaller vertical extensions. In some cases winter storms are not detected by the radars at long range. In that case, lightning produced by such undetected winter storms are wrongly classified as outliers. Vice versa, an incorrect classification may also occur when a wrong detection appears by chance in a precipitation area detected by the radar. In this case, a wrong lightning detection is classified as a correct detection. Since radars generally detect less precipitation in winter than in summer (e.g. Hazenberg et al., 2011) such misclassification occurs less in winter than in summer, which means that the classification method will produce more outliers in winter. Thus, the reduced efficiency of precipitation detected by the weather radars in winter is an additional possible source of the observed increase of outlier classifications in winter. Note that Poelman et al. (2016) showed that on average peak current estimates of winter lightning are higher than in summer. One would therefore expect that on average in winter more sensors participate in a lightning event compared to summer, resulting in a good location accuracy. Nevertheless, the absolute number of outliers during winter is much smaller compared to summer, as can be deduced from Fig. 3b. Thus, the increase in percentage of outliers may not be too important for the majority of applications."

3) From Figures 7, 8 and 9, I ask to the authors: all those outliers cannot be considered simply IC discharges misclassified by the network? Note that they mostly present the typical behavior of IC flashes:(1) low peak current values (because they are in majority weaker than the CGs); (2) usually are detected with larger SMA (because are detected by less sensors and has long horizontal extensions inside the clouds leading to major errors in their location (i.e., projection over ground); and (3) present (in a such way) “random” polarity since the ICs can move upward and downward inside the clouds. I’d like to hear more from the authors about this point based on the presented results.

We add a small paragraph at the end of Sect. 3 related to the above question raised: “Looking at Fig. 7 to 9 one could wonder whether those CG outliers could be simply considered as IC discharges misclassified by the network, since IC discharges have on average lower peak currents, hence lower number of contributing sensors and therefore smaller SMA. Although this can be partly true, still a considerable fraction of the CG outliers are found to have large peak currents. It is therefore unlikely that all the CG outliers found with this method are in fact misclassified IC discharges.”

Analysis of lightning outliers in the EUCLID network

Dieter R. Poelman[1], Wolfgang Schulz[2], Rudolf Kaltenboeck[3], Laurent Delobbe[1]

[1]Royal Meteorological Institute of Belgium, Brussels, Belgium
[2]OVE-ALDIS, Vienna, Austria
[3]Austro Control, Innsbruck, Austria

Correspondence to: Dieter R. Poelman (dieter.poelman@meteo.be)

Abstract. Lightning data as observed by the European Cooperation for Lightning Detection network EUCLID are used in combination with radar data to retrieve the temporal and spatial behaviour of lightning outliers, i.e. discharges located on a wrong place, over a 5-year period from 2011 to 2016. Cloud-to-ground stroke and intracloud pulse data are superimposed on corresponding 5-min radar precipitation fields in two topographically different areas, being Belgium and Austria, in order to extract lightning outliers based on the distance between each lightning event and the nearest precipitation. It is shown that the percentage of outliers is sensitive to changes in the network and to the location algorithm itself. The total percentage of outliers for both regions varies over the years between 0.8% and 1.7% for a distance to the nearest precipitation of 2 km, with an average of approximately 1.2% in Belgium and Austria. Outside the European summer thunderstorm season the percentage of outliers tends to increase somewhat. The majority of all the outliers are low peak current events with absolute values falling between 0 to 10 kA. More specifically, positive cloud-to-ground strokes are more likely to be classified as outliers compared to all other type of discharges. Furthermore, it turns out that the number of sensors participating in locating a lightning discharge is different for outliers versus correctly located events, with outliers having the least amount of sensors participating. In addition, it is shown that in most cases the semi-major axis assigned to a lightning discharge as a confidence indicator in the location accuracy is smaller for correctly located events compared to the semi-major axis of outliers.

1 Introduction

Present-day lightning location systems (LLS) are the result of continuous development over the years with improved location accuracy, peak current estimation and type classification for each observed lightning event. However, despite the great progress made to determine those properties amongst others, occasionally some events remain poorly determined by the LLS. For instance, the uncertainty of the measurements related to a low peak current discharge tends to be larger than it is for a high peak current event. In addition, it is still common practice to categorize positive cloud-to-ground (CG) strokes with estimated peak currents smaller than 5 or 10 kA as IC pulses since those are more likely to be of intracloud (IC) nature (Cummins et al., 1998; Wacker and Orville, 1999a, b; Jerauld et al., 2005, Orville et al., 2002; Cummins et al., 2006; Biagi et al., 2007). However, not all the properties are of equal importance for the different users of lightning data. Depending on the customers' application of the LLS data, different performance features are more important, while others are less important, e.g. power utilities normally do not care about the IC detection efficiency (DE) of an LLS, whereas the quality of the CG data is of utmost importance. On the other hand, aviation control and meteorological services which often trigger warning messages based on LLS data favor a good DE of CG as well as IC events coupled to a minimum of events located on a completely wrong position. It is therefore a necessity to gain a thorough knowledge of the LLS at hand.

Comment [DP1]: Referee 1: missing word 'important'

During recent years the performance of LLS got more and more attention (Nag et al., 2015). A direct method to determine the quality of a network, and therefore the values assigned to each lightning event, is by comparing the data against so-called ground-truth observations. Those observations provide valuable information on the DE, location accuracy (LA) and in some cases even the peak current estimates retrieved from an LLS. This is done for instance by examining direct lightning strikes to instrumented towers (Diendorfer et al., 2000a, 2000b; Pavanello et al., 2009; Romero et al., 2011; Schulz et al., 2012; Schulz et al., 2013; Cramer and Cummins, 2014; Azadifar et al., 2016), through the use of rocket triggered lightning (Jerauld et al., 2005; Nag et al., 2011; Chen et al., 2012; Mallick et al., 2014a, 2014b, 2014c), and/or by recording lightning strikes with high-speed video and E-field measurements in open field (Biagi et al., 2007; Chen et al., 2012; Poelman et al., 2013a, Schulz et al., 2016). Although best practice to retrieve robust information on a networks' performance, the aforementioned methods are quite labor intensive in order to acquire a large enough dataset for a statistically reliable output. Other methods exist, such as intercomparing different LLS within regions of overlapping coverage (Said et al., 2010; Pohjola and Mäkelä, 2013; Poelman et al., 2013b). However, the main disadvantage of those studies is the assumption of one network as being "ground-truth". In reality this is hardly the case for any existing LLS, except maybe for the short-baseline lightning mapping arrays (Rison et al., 1999; Thomas et al., 2004; van der Velde et al., 2013; Defer et al., 2015).

In this paper lightning data are combined with radar precipitation observations to analyze the temporal and spatial behavior of lightning outliers in two topographically different regions in Europe. Lightning outliers are sometimes also referred to in the literature as fake or ghost strokes and can be the result of signal interferences from power lines, radio frequencies or other site-specific disturbances or are simply misplaced events by the location algorithm. The results presented here are obtained by combining lightning observations from the European Cooperation for Lightning Detection network EUCLID with radar precipitation data in Belgium and Austria, as described in Section 2. The results of the analysis are presented in Section 3 and summarized in Section 4.

2 Data and Methodology

2.1 Lightning location data

The European Cooperation for Lightning Detection network EUCLID has been operational since 2001 and processes as of January 2017 in real-time data of 164 sensors to provide European wide lightning observations of high and nearly homogeneous quality (Schulz et al., 2016; Poelman et al., 2016). All of the sensors operate over the same low-frequency (LF) range and provide amongst others timing and angle information. The individual raw sensor data are sent in real-time to a single processor, calculating the electrical activity at any given moment. The locations of the EUCLID sensors are displayed in Fig. 1. The network has been tested continuously over the years against ground-truth data from direct lightning current measurements at the Gaisberg Tower in Austria (Schulz et al., 2016), Peisserberg Tower in Germany (Heidler and Schulz, 2016) and Säntis Tower in Switzerland (Romero et al., 2011; Azadifar et al., 2016) and data from E-field and video recordings in Austria, France and Belgium

(Schulz et al., 2016). The latest comprehensive performance analysis of the EUCLID network based on those measurements revealed that the flash and stroke DE for negative CG discharges in different regions of the EUCLID network are greater than 93% and 84%, respectively, while for positive events those are greater than 87% and 84 %, respectively (Schulz et al., 2016). To retrieve the latter values, only those strokes are used in the analysis that match certain quality criteria such as χ^2, a measure for the agreement between the different sensor measurements, and the semi-major axis of the confidence ellipse, and received a correct stroke classification as CG by the central processor. Those strict criteria, as well as temporary sensor outages during the measurements campaign, can impact the DE estimates given in Schulz et al. (2016). In addition, Schulz et al. (2016) showed that the LA dropped steadily over the years down to the present LA in the range of 100 m within the majority of the network. Note that in Schulz et al. (2016) ground truth observations are collected in Austria and Belgium, the same regions of the EUCLID network which are studied in this paper.

Comment [DP2]: Referee3: please clarify

During the time period under consideration, significant changes of the EUCLID network regarding DE and LA were made (Schulz et al., 2016). Those are related to new sensor technology, timing error corrections and a new location algorithm which can influence the outlier behavior. One would think sensor upgrades have always a positive influence on a networks performance. While this is generally true in the long run, the upgrades can cause temporary problems in the beginning since those sensors are awaiting calibration. This is especially true for some sensors in Italy in 2014. From the day of the setup till the sensors were calibrated those sensors were configured to provide timing information only. However, timing only sensors often increase the number of outliers if those are used in solutions determined by two or three sensors only.

Comment [DP3]: Referee 1: provide the physical and/or technical explanation for this

Figure 2a plots the annual distribution of the total stroke count over the years 2011until 2016, as observed within the red boxes in Fig. 1. As expected, the CG distribution experiences a natural annual variability in Belgium as well as in Austria. With regard to the IC detections, one notices a sharp increase in 2015 and 2016. This increase is not climatological in nature, but is attributed to the increased amount of LS700x sensors in EUCLID and its capability to detect IC pulses in the low-frequency domain. The distribution of the total monthly stroke count is shown in Fig. 2b. A peak in activity is observed in June and July for Belgium and Austria, respectively. For both regions about 95% of all the observed lightning activity occurs between May and September.

Comment [DP4]: Referee1: we've added a new Figure 2 in the paper.

2.2 Weather radar data

Weather radar data of the Royal Meteorological Institute of Belgium (RMIB) and of Austro Control in Austria are used in this study. Fig. 1 shows the locations (white stars) and coverage (dashed lines) of the individual radars, as well as the limit of the composite as the outer contour of all the radars (solid lines). The use of radar composites is preferred over the individual radar observations since individual radar observations can be hampered by shielding effects. This is true especially in mountainous regions such as the Alps in Austria, limiting the detection range where the radar data is still considered of sufficient quality. Additionally, since the height of the radar beam above ground increases with increasing distance from the radar, precipitation can be underestimated or even undetected at far range by overshooting when precipitation is produced below the height of the lowest radar beam. Therefore, to eliminate the latter effect, the two geographical areas in this study are limited to the red boxes as indicated in Fig. 1.

Comment [DP5]: Referee 1: include more info on 'overshooting beams'.

Comment [DP6]: Referee 1: added this sentence to emphasis that we limit it to remove any long-range issues with the radar

The composite radar reflectivity threshold is set at 12 dBZ. Following the $Z = 200*R^{1.6}$ relationship from Marshall and Palmer (1948), with Z being the reflectivity and R the rain rate, this threshold corresponds to a rain rate of 0.2 mm/h below which the rain rates are set to zero in this study. This low reflectivity threshold helps to detect convective clouds relevant for lightning generation even in weak cell cores from winter events or upper areas of thunderstorms at far ranges from the radar site.

2.2.1 Belgium

The radar composite used at RMIB consists out of three radars. RMIB owns and operates two of them; the radar at Wideumont in the southeast of Belgium and the radar in Jabbeke located near the west coast which became only operational since 2013. The third weather radar at the center of the composite is located at the airport in Zaventem near Brussels and is operated by Belgocontrol, in charge of the safety of civil aviation. All of the radars are C-band Doppler radars, performing a multiple elevation reflectivity scan every 5 minutes with a resolution of one degree in azimuth and 500 m in range for

Jabbeke and Zaventem and 250 m in range for Wideumont. The maximum range is 300 km for Jabbeke and 250 km for Zaventem and Wideumont. A Doppler filter for clutter elimination is used for the three radars and an additional polarimetric fuzzy logic filter is used for Jabbeke. For each radar, a 2D precipitation product is derived from the volume reflectivity data. The height of those individual products corresponds to 1500 m above sea level. A composite is subsequently produced from these 2D products taking for each pixel the maximum value of the radars covering this pixel. For more information on the validation of the Belgian composite, the interested reader is referred to Goudenhoofdt and Delobbe (2016).

Comment [DP7]: Referee1: wanted to know at which height the product corresponds

Comment [DP8]: Referee1: added reference concerning validation of the composite used in Belgium

2.2.2 Austria

Austro Control, the Austrian civil air service provider is operating five C-band EEC polarized Doppler weather radars in Austria, of which four of them are used in this study. Two of the radar sites are located on the foothills of the Alps close to Vienna and Salzburg (Rauchenwart and Feldkirchen) and the other two radar sites are situated in the west and south of Austria at mountain tops above 2000 m close to Innsbruck and Klagenfurt (Patscherkofel and Zirbitzkogel). The underlying volume scan contains 16 elevations ranging between -1.5° and 67° up to a range of 224 km. Doppler and statistical clutter filters are applied before creating maximum surface projection of reflectivity which combines strongest return from each elevation level. Resulting Austrian composite uses maximum reflectivity in horizontal extent which is provided by one of the 4 radars to avoid shielding effects of the Alps. Temporal and spatial resolution is 5 minutes and 1 km, respectively. For more details, the interesting reader is referred to Kaltenboeck and Steinheimer (2015) and Kaltenboeck (2012a, 2012b). It is important to note that the Austrian weather radar network was upgraded between 2011 and 2013, during which the individual radar gains were modified. This adaptation of the gain could easily influence to some degree the findings in this paper.

2.3 Methodology

To account for border effects of the radar observations as mentioned in Sect. 2.2, only lightning events within the red boxes as indicated in Fig. 1 are used. Those regions correspond approximately to the area where two or more radars participate in the radar image with sufficient distance from the border. Subsequently, CG strokes and IC pulses with timestamps, that fall within the start and end time of the radar scan, are superimposed on the corresponding 5-min radar precipitation fields. In order to have overall homogeneous coverage of the weather radar data, only the time steps were used for which all the radars within the composites were in operation. An event is then categorized as an outlier when no precipitation within a certain distance has been observed. The distance at which an event is classified as outlier is somewhat chosen arbitrary. Different runs are performed applying a distance Δr of 2, 5, and 10 km. An example of this method is visualized in Fig. 3. All the lightning events are superimposed as black dots, whereas the retrieved outliers are in red for clarity. Note that this method is supposed to give a lower limit of the percentage of outliers because some of the outliers will, by chance, be placed in a region with radar reflectivity larger than 12 dBZ. In the remainder of the paper, explanation of the results is based on the findings for a search radius Δr of 2 km, unless otherwise stated explicitly.

Comment [DP9]: added

Comment [DP10]: referee1: text adapted slightly to show how we selected the lightning data corresponding to the radar images.

3 Results

The overall annual percentage of outliers for CG strokes and IC pulses relative to the total number of events, as a function of Δr between the event location and the nearest precipitation, is plotted in Fig. 4 for Belgium and Austria. There are several similarities and differences between the two areas. For example, the total percentage of outliers is of the same order of magnitude for both regions and varies between 0.8% and 1.7% throughout the years for an adopted Δr of 2 km. It is clear that choosing a larger Δr decreases the percentage of outliers, and vice versa, while maintaining the same annual trend. The percentage of the total outliers averaged over six years in Belgium and Austria is approximately 1.2%. In Belgium on average 0.5% of the outliers are of type CG and this value increases up to 0.9% in case of IC, whereas in Austria the level of CG outliers is only slightly higher than that of IC outliers, i.e. 0.8% for CG with respect to 0.5% in case of IC. Shorter baselines in Austria compared to Belgium could be a reason for this discrepancy. The significant higher number of outliers in Belgium in 2011 compared to Austria can be attributed to a timing only sensor located close to Belgium (Den Haag) and another sensor in the Netherlands which was moved and afterwards operated for a longer time period with deactivated angle information (Roermond). From our experience, sensors providing only time information often cause additional outliers. For the vast majority of the sensors which provide angle and time information those measurements have to be consistent since coherence between the latter two reduces the number of outliers. The level of outliers from 2012-2014 is roughly the same for both areas. The lowest level of CG outliers is found in 2016 in both areas. In addition, it is worth mentioning that in Belgium and Austria the majority of the CG outliers are single stroke flashes, while only a minority of the CG outliers belong to a flash with multiplicity larger than one. One could say that assuming a stable radar network, the variation in the percentage of outliers over the years reflects the status of the lightning location network in a certain area. Hence, continuous monitoring of the outliers has the possibility to pick up potential problems in the network, which can be relevant for future automatic forecast applications.

The left panels in Figure 5 display the 6-year mean annual total lightning event (CG strokes + IC pulses) density on a 10 x 10 km^2 grid in Belgium and Austria. For Belgium, densities vary between 0.8 and 11 events km^{-2}yr^{-1} with a median value of 3.4 events km^{-2}yr^{-1} at 10km x 10km resolution. Overall the densities in Austria are somewhat higher compared to Belgium resulting in a median value of 4.4 events km^{-2}yr^{-1}. The highest total lightning densities are found towards the southeast of Austria with a maximum of 22 events km^{-2}yr^{-1}. Fig. 5c and 5d reveal the spatial distribution of the percentage of total outliers as observed in between 2011 and 2016. In Belgium, values range from 0.2% to 3.3%. The distribution of outliers within Belgium is rather uniform with here and there somewhat higher percentage values. The latter are mainly caused by IC outliers, since those contribute the most to the overall outlier percentage in each grid cell. In Austria, grid cell percentages range from 0.1% up to 33%. The majority of the grid cells have low outlier percentage values, except in the southwest corner. This is exactly the place where the Alps disturb the radar observations leading to an increase in outliers with the employed method. From Fig. 5a and 5b we conclude that there is no real correlation between the total lightning event density and the spatial distribution of outliers.

Comment [DP11]: referee1: changed a bit the text to account for the change in the figure.

Comment [DP12]: Referee1: I add this to account for the change in the figure.

Fig. 6 illustrates the monthly variation of the percentage of outliers. An obvious decrease is observed in the percentage of outliers during May-Sept, compared to the other months of the year. This feature could be related to the fact that more sensor upgrades occur during winter or because precipitation of winter thunderstorms is more difficult to detect with the weather radars. In addition, the 3D structure of lightning flashes in winter compared to summer is somewhat different (Lopez et al., 2017), which could increase the difficulty to locate those in winter accurately. Regarding the sensor upgrades, those often result in disabled angle information because systematic angle errors, i.e. site errors, are at first unknown and the correction takes a while because lightning data is necessary. Consequently, upgraded sensors start operation with disabled angle information during winter months. With respect to the observation of precipitation, during summer most of the storms are associated with large amounts of precipitation in vertically extended clouds, meaning that those storms are always very well detected by the radars. In contrast, winter storms are generally associated with less intense precipitation cells and with smaller vertical extensions. In some cases winter storms are not detected by the radars at long range. In that case, lightning produced by such undetected winter storms are wrongly classified as outliers. Vice versa, an incorrect classification may also occur when a wrong detection appears by chance in a precipitation area detected by the radar. In this case, a wrong lightning detection is classified as a correct detection. Since radars generally detect less precipitation in winter than in summer (e.g. Hazenberg et al., 2011) such misclassification occurs less in winter than in summer, which means that the classification method will produce more outliers in winter. Thus, the reduced efficiency of precipitation detected by the weather radars in winter is an additional possible source of the observed increase of outlier classifications in winter. Note that Poelman et al. (2016) showed that on average peak current estimates of winter lightning are higher than in summer. One would therefore expect that on average in winter more sensors participate in a lightning event compared to summer, resulting in a good location accuracy. Nevertheless, the absolute number of outliers during winter is much smaller compared to summer, as can be deduced from Fig. 3b. Thus, the increase in percentage of outliers may not be too important for the majority of applications.

Fig. 7 plots the outlier percentages related to each individual group, e.g. percentage of negative IC is related to the total number of negative IC. Proportionally, the degree of occurrence of positive and

Comment [DP13]: Referee3

Comment [DP14]: Referee 1: Added the extra possible reason of the increase of outliers in winter.

Comment [DP15]: Referee1: “which” instead of “wich”

Comment [DP16]: This paragraph was changed slightly to the comments of referee 3.

Comment [DP17]: Referee3

negative outliers is of the same level, except for 2011, and follows the annual variation as in Fig. 4. Positive CG strokes exhibit the highest percentage of outliers in Belgium and Austria. This could be related to the fact that positive CG strokes are often accompanied with significant IC activity complicating the transmitted electromagnetic fields (Fuquay, 1982; Saba et al., 2009). It is therefore harder to detect and locate correctly such strokes, resulting in a higher percentage of outliers. Furthermore, the percentage of negative CG outliers is roughly half of that of the positive CG outliers for the years 2011-2014. The opposite is found in case of IC pulses, where the percentage of negative IC outliers is higher compared to the positive counterpart. However, the difference between positive and negative CG outliers and/or IC pulses decreases in 2015 and 2016. Thus, the percentage of outliers is more or less unrelated to the polarity of the event. In 2016, it is obvious that the outlier percentages of the individual types are more or less in line with each other. This could also be a result of the improved performance of the latest adopted location algorithm.

In Fig. 8, the percentage of outliers for peak current intervals up to +/- 20 kA is plotted, calculated with respect to the total amount of discharges within each peak current interval. It is seen that the distribution corresponds well between Belgium and Austria. Because positive CG strokes with peak currents below 5 kA are categorized as IC, no data for positive CG below 5 kA exists. First of all, the majority of the outliers for positive CG and IC discharges are found in between [5, 10] kA and [0, 5] kA, respectively, with a decline towards the larger peak current intervals. This is not surprising since the higher the peak current, the more sensors participate on average in locating the event. This is also true for negative IC outliers, whereas negative CG outliers have the highest percentage in the [-10, -5] kA range. Except for the [-5, 0] kA interval, the percentages are similar between negative IC and CG outliers. This is not the case for the positive IC and CG outliers. In addition to what is plotted in Fig. 8, it is found that the percentage of CG outliers with absolute peak currents above 20 kA is a factor of about three lower compared to the total percentage of outliers found within |0-20| kA, and a factor of 15 lower in case of IC. The larger drop in case of IC results from the lower amount of IC events with absolute peak currents larger than 20 kA compared to CG events.

Fig. 9 reveals the cumulative distribution of the number of sensors participating in a solution as a function of event type. First of all, one notices that in case of CG strokes more sensors participate in a

Comment [DP18]: Referee 1

Comment [DP19]: Referee 1: What is the % of outliers with an absolute current above 20 kA?

solution compared to IC events. This is attributed to the fact that in the LF range the amplitude of even the largest IC pulses is significantly lower compared to that of CG return strokes (Weidman et al., 1981). The amplitude difference between CG strokes and IC pulses increases even further with increasing propagation distance between the source and the lightning sensor (Cooray et al., 2000). Hence, more sensors will detect the radiation from a single CG discharge compared to an IC pulse. The resemblance in distribution between Belgium and Austria is not surprising since the lightning sensors in EUCLID are quite homogeneously distributed across the network. In addition, more sensors participate in the location of discharges that are correctly located, than is the case of CG as well as IC outlier events. For instance, 85% of the IC outliers are located by 2 or 3 sensors, whereas this drops to 50% for correctly located IC pulses. For CG strokes on the other hand, only 20% of the outliers are located with more than 6 participating sensors, whereas this is the case for more than 60% for the CG strokes within 2 km of the nearest precipitation. We find that the median amount of sensors participating in a solution for correctly located CG and IC is 8 and 3, respectively, and this drops to 3 and 2 sensors participating in case of CG and IC outliers.

The central processor assigns to each lightning event a value of the semi-major axis (SMA) of the 50% confidence ellipse. This value can be used as a quality indicator of the location accuracy, with smaller values indicating a larger confidence in the assigned location of the event. The distribution of SMA for all the events (CG + IC) is plotted in Fig. 10, separated into and normalized to outlier and correctly located events. In addition, the total number of events per SMA interval for Belgium and Austria is indicated. Note that events with an SMA larger than 7.5 km do not exist in the data since those events are regarded as bad quality events and hence rejected by the location algorithm. First of all, it is striking that the SMA distribution is almost equal for Belgium and Austria. The majority of OK events, i.e. 75%, have SMA values falling in the 0-1 km range, whereas this drops to 40% in case of outliers. The average and median value of the SMA for OK events is 775 m and 200 m, respectively, and this increases to 1.83 km and 1.48 km in case of outliers. Although not shown in this plot, it is found that the average SMA for CG strokes is smaller by a factor of two compared to IC pulses. This is expected since more sensors participate in a solution for CG strokes compared to IC pulses as discussed in Fig. 9.

Comment [DP20]: Referee1: asked to include the amount per SMA interval in the figure

Comment [DP21]: Sentence slightly rephrased

Comment [DP22]: Referee 1: It was decided to delete the sentence about Xi2 because it was confusing and anyway we are not showing any result of it in the paper.

Looking at Fig. 8 to 10 one could wonder whether the CG outliers could be simply considered as IC discharges misclassified by the network, since IC discharges have on average lower peak currents, hence lower number of contributing sensors and therefore smaller SMA. Although this can be partly true, still a considerable fraction of the CG outliers are found to have large peak currents. It is therefore unlikely that all the CG outliers are in fact misclassified IC discharges.

Comment [DP23]: Related to comment of Referee3 (Kleber Naccarato)

Up to now, lightning discharges have been classified into either correctly located strokes or outliers based on their distance to the nearest precipitation. However, using an additional time criterion it is possible to further dissociate the outliers into isolated outliers in space and time from those that are just wrongly located from a group of correctly located events. Fig. 11 plots the distribution of the time difference between the outliers and its closest (in time) correctly located event. Once more, the distribution is found to be similar in Belgium and Austria. The majority of the outliers occur within one second of a correctly located event. One could argue that these are simply bad located lightning events, whereas those that take place later than one second are so-called ghost outliers, i.e. outliers in time and space. Furthermore, from this plot it is found that the outliers behave quite independent of polarity and classification in Belgium and Austria.

Comment [DP24]: Referee1: related to comment to include a time criterion

4 Summary

In this study all lightning events detected by the EUCLID network during 2011 and 2016 that fall within selected areas in and around Belgium and Austria are classified as outliers or correctly located events based on their distance Δr to the nearest precipitation. The latter two regions were chosen specifically for their difference in topography and because high spatial and temporal resolution radar data were readily available. A similar approach can be performed in the future on a larger spatial scale based for instance on the radar composite imagery produced by the Eumetnet Operational Programme for the Exchange of Weather Radar Information (OPERA, Huuskonen et al. 2014) and related EUCLID domain.

Comment [DP25]: Referee1: Why only these two regions? And not a larger domain covered by both EUCLID and the European radars?

The applied methodology makes use of radar data with an adopted lower reflectivity threshold of 12 dBZ. Hence, precipitation is required to discriminate between the outliers and well-located lightning events. Therefore lightning produced by “dry” thunderstorms or bolts from the blue will be misclassified as outliers. However, these particular phenomena are extremely limited and would not influence to a large extent the results presented in this study. We believe that a methodology based on satellite cloudiness products would not allow a proper identification of outliers since cloudiness in Belgium and Austria is mostly not associated with thunderstorms.

Comment [DP26]: Referee 1: related to the comment that precipitation at ground is required to verify lightning data.

Categorizing the lightning events based on radar reflectivity data and comparing the results from different geographical regions is not a straightforward task. The reason is potential calibration issues in the different radar networks with maybe even different technology and local beam blockage problems especially in the mountainous regions in Austria. A workaround, at least for the last problem, is to use composite radar data. Despite those difficulties the overall results in both regions agree quite well. The overall percentage of outliers for both regions varies annually between 0.8% and 1.7% for a distance Δr to the nearest precipitation of 2 km and drops further when a more relaxed Δr is chosen. These values are lower limits since it is possible that an outlier is located in an area with rain. The percentage of outliers is quite small having in mind that a Δr of 2 km is already a quite strict criterion. Outside the European summer thunderstorm season the percentage of outliers tends to increase somewhat. Amongst some of the sources responsible for this increase is the fact that more sensor upgrades occur during winter or that the radar underestimates to some extent precipitation. In addition, the 3D structure of flashes is somewhat different in winter compared to summer, which could increase the difficulty to locate those accurately The majority of all the outliers are low peak current events with absolute values falling between 0 to 10 kA. More specifically, positive CG strokes are more likely to be classified as outliers compared to all other type of discharges. Furthermore, it turns out that the number of sensors participating in locating a lightning discharge is different for outliers versus correctly located events, with outliers having the least amount of sensors participating. In addition, it is found that in general the SMA of non-outliers is much smaller compared to the SMA belonging to outliers.

Comment [DP27]: Referee 1: rewrote this here a bit to account for the extra reason (3D flash structure) in winter/summer.

References

Azadifar, M., Rachidi, F., Rubinstein, M., Paolone, M., Diendorfer, G., Pichler, H., Schulz, W., Pavanello, D., and Romero, C.: Evaluation of the performance characteristics of the European Lightning Detection Network EUCLID in the Alps region for upward negative flashes using direct measurements at the instrumented Säntis Tower, J. Geophys. Res. Atmos., 121, 595–606, doi:10.1002/2015JD024259, 2016.

Biagi, C. J., Cummins, K. L., Kehoe, K. E., and Krider, E. P.: National Lightning Detection Network (NLDN) performance in southern Arizona, Texas, and Oklahoma in 2003-2004, J. Geophys. Res., 112, D05208, doi:10.1029/2006JD007341, 2007.

Cramer, J. A., and Cummins, K. L.: Evaluating location accuracy of lightning location networks using tall towers, in 23rd International Lightning Detection Conference & 5th International Lightning Meteorology Conference, Vaisala Inc., Tucson, Arizona, 2014.

Chen, M., Zheng, D., Du, Y., and Zhang Y.: A statistical method for evaluating detection efficiency of lightning location network and its application, Atmos. Res., 128, 13–23, 2013.

Cooray, V., Fernando, M., Sörensen, T., Götschl, T., and Pedersen, A.: Propagation of lightning generated transient electromagnetic fields over fintely conducting ground, J. Atmos. Terres. Phys., vol. 62, pp. 583-600, 2000.

Cummins, K., Murphy, M., Bardo, E., Hiscox, R., Pyle, W. L., and Pifer, A.: A combined TOA/MDF technology upgrade of the U.S. National Lightning Detection Network, Journal of Geophysical Research, 103(D8), 9035-9044, 1998.

Cummins, K., Cramer, J., Biagi, C., Krider, E., Jerauld, J., Uman, M., and Rakov, V.: The U.S. National Lightning Detection Network: Post-upgrade status, Second Conference on Meteorological Applications of Lightning Data, Atlanta, Georgia, American Meteorological Society, 2006.

Defer, E., Pinty, J.-P., Coquillat, S., Martin, J.-M., Prieur, S., Soula, S., Richard, E., Rison, W., Krehbiel, P., Thomas, R., Rodeheffer, D., Vergeiner, C., Malaterre, F., Pedeboy, S., Schulz, W., Farges, T., Gallin, L.-J., Ortéga, P., Ribaud, J.-F., Anderson, G., Betz, H.-D., Meneux, B., Kotroni, V., Lagouvardos, K., Roos, S., Ducrocq, V., Roussot, O., Labatut, L., and Molinié, G.: An overview of the lightning and atmospheric electricity observations collected in southern France during the HYdrological cycle in Mediterranean EXperiment (HyMeX), Special Observation Period 1, Atmos. Meas. Tech., 8, 649–669, doi:10.5194/amt-8-649-2015.

Diendorfer, G., Mair, M., Schulz, W., and Hadrian, W.: Lightning current measurements in Austria—Experimental setup and first results, in 25th Intl. Conf. on Lightning Protection, pp. 44–47, ICLP Centre, Rhodes, Greece, 2000a.

Diendorfer, G., Schulz, W., and Mair M.: Evaluation of a LLS based on lightning strikes to an instrumented tower, in 16th International Lightning Detection Conference, Tucson, Ariz., 2000b.

Fuquay, D.M.: Positive cloud-to-ground lightning in summer thunderstorms, J. Geophys. Res. Ocean., vol. 87, 7131-7140, 1982.

Goudenhoofdt, E., and Delobbe, L.: Generation and verification of rainfall estimates from 10-yr volumetric weather radar measurements, J. Hydrometeorol., 17, 1223-1242, 2016.

Hazenberg, P., Leijnse, H., and Uijlenhoet, R.: Radar rainfall estimation of stratiform winter precipitation in the Belgian Ardennes, Water Resour. Res., 47, W02507, doi:10.1029/2010WR009068, 2011.

Heidler, F., and Schulz, W.: Lightning current measurements compared to data from the lightning location system BLIDS, in International Colloquium on Lightning and Power Systems (CIGRE), Bologna, Italy, 2016.

Huuskonen, A., Saltikoff, E., and Holleman, I.: The Operational Weather Radar Network in Europe, Bull. Amer. Meteor. Soc., 95, 897–907, 2014.

Comment [DP28]: Referee1: reference related to "Why only these two regions? And not a larger domain covered by both EUCLID and the European radars?"

Jerauld, J., Rakov, V. A., Uman, M. A., Rambo, K. J., Jordan, D. M., Cummins, K. L., and Cramer, J. A.: An evaluation of the performance characteristics of the U.S. National Lightning Detection Network in Florida using rocket-triggered lightning, J. of Geophys. Res., 110, D19106, doi:10.1029/2005JD005924, 2005.

Kaltenboeck, R., and Steinheimer, M.: Radar-based severe storm climatology for Austrian complex orography related to vertical wind shear and atmospheric instability, Atmos. Res., vol. 158-159, 216-230, 2015.

Kaltenboeck, R.: New generation of dual polarized weather radars in Austria, in 7th Euoprean Conference on Radar in Meteorology and Hydrology (ERAD), 2012.

Kaltenboeck, R.: Das österreichische Wetterradarnetzwerk, OEGM Bull., vol. 2, pp. 14-22, 2012.

López, J. A., Pineda, N., Montanya, J., van der Velde, O., Fabro, F., Romero, D.: Spatio-temporal dimension of lightning flashes based on three-dimensional Lightning Mapping Array., Atmospheric Research, 197, 255-264, 2017.

Mallick, S., Rakov, V., Ngin, T., Gamerota, W., Pilkey, J., Hill, J., Uman, M., Jordan, D., Nag, A., and Said, R.: Evaluation of the GLD360 performance characteristics using rocket-and-wire triggered lightning data, Geophys. Res. Lett., 41, 3636–3642, doi:10.1002/2014GL059920, 2014a.

Mallick, S., et al.: Performance characteristics of the NLDN for return strokes and pulses superimposed on steady currents, based on rocket-triggered lightning data acquired in Florida in 2004–2012, J. Geophys. Res. Atmos., 119, 3825–3856, doi:10.1002/2013JD021401, 2014b.

Mallick, S., et al.: Performance characteristics of the ENTLN evaluated using rocket-triggered lightning data, Electr. Power Syst. Res., 118, 15–28, doi:10.1016/j.epsr.2014.06.007, 2014c.

Marschall, J. S., and Palmer, W. Mc K.: The distribution of raindrop size, J. Meteorol., no. 5.4, pp. 165-166, 1948.

Murphy, M. J., and Cummins, K. L.: Two-dimensional and three-dimensional cloud discharge detection, presented at the Int. Lightning Detection Conf., Tucson, AZ, 1998.

Nag, A., et al.: Evaluation of US National Lightning Detection Network performance characteristics using rocket-triggered lightning data acquired in 2004–2009, J. Geophys. Res., 116, D02123, doi:10.1029/2010JD014929, 2011.

Nag, A., Murphy, M. J., Schulz, W., Cummins, K. L.: Lightning Location Systems: Insights on characteristics and validation techniques, Earth Sp. Sci., no. 2, pp. 65-93, 2015.

Orville, R., Huffines, G., Burrows, W., Holle, R., and Cummins, K.: The North American Lightning Detection Network (NLDN) – First results: 1998-2000, Monthly Weather Review, 130, 2098-2108, 2002.

Pavanello, D., Rachidi, F., Janischewskyj, W., Rubinstein, M., Shostak, V., Nucci, C. A., Cummins, K. L., Hussein, A. M., and Chang, J. S.: On the current peak estimates provided by lightning

Comment [DP29]: Referee 1: Added extra reference, related to the different 3D structure of winter/summer flashes

detection networks for lightning return strokes to tall towers, IEEE Trans. Electromagn. Compat., 51(3), 453–458, 2009.

Poelman, D. R., Schulz, W., and Vergeiner, C.: Performance characteristics of distinct lightning detection networks covering Belgium, J. Atmos. Oceanic Technol., 30, 942–951, doi: 10.1175/JTECH-D-12-00162.1, 2013a.

Poelman, D. R., Honoré, F., Anderson, G., and Pedeboy, S.: Comparing a regional, subcontintental, and long-range lightning location system over the Benelux and France, J. Atmos. Oceanic Technol., 30, 2394–2405, doi:10.1175/JTECH-D-12-00263.1, 2013b.

Poelman, D. R., Schulz, W., Diendorfer, G., and Bernardi, M.: The European lightning location system EUCLID – Part 2: Observations, Nat. Hazards Earth Syst. Sci., vol. 16, pp. 607-616, 2016.

Pohjola, H., and Mäkelä, A.: The comparison of GLD360 and EUCLID lightning location systems in Europe, Atmos. Res., 123, 117–128, 2013.

Rison, W., Thomas, R. J., Krehbiel, P. R., Hamlin, T., and Harlin, J.: A GPS-based three-dimensional lightning mapping system: initial observations in central New Mexico, Geophys. Res. Lett., 26, 3573–3576, 1999.

Romero, C., Paolone, M., Rachidi, F., Rubinstein, M., Rubinstein, A., Diendorfer, G., Schulz, W., Bernardi, M., and Nucci, C. A.: Preliminary comparison of data from the säntis tower and the euclid lightning location system, XIth International Symposium on Lightning Protection (SIPDA), Fortaleza, Brazil, 3–7 Oct, 2011.

Saba, M. M. F., Campos, L. Z. S., Krider, E. P., and Pinto., O. J.: High-speed video observations of positive ground flashes produced by intracloud lightning, Geophys. Res. Lett., vol. 36, no. 12, p. L12811, 2009.

Said, R., Inan, U., and Cummins, K.: Long-range lightning geolocation using a VLF radio atmospheric waveform bank, J. Geophys. Res., 115, D23108, doi:10.1029/2010JD013863, 2010.

Schulz, W., Vergeiner, C., Pichler, H., Diendorfer, G., and Packet, S.: Validation of the Austrian Lightning Location System ALDIS for negative flashes, in Proc. CIGRE Symposium, 2012.

Schulz, W., Pichler, H., Diendorfer, G., Vergeiner, C., and Pack, S.: Validation of detection of positive flashes by the Austrian Lightning Location System (ALDIS), in 12th Intl. Symposium on Lightning Protection (XII SIPDA), paper 2.2, Belo Horizonte, Brazil, 2013.

Schulz, W., Diendorfer, G., Pedeboy, S., and Poelman, D. R.: The European lightning location system EUCLID – Part 1: Performance validation, Nat. Hazards Earth Syst., vol. 16, no. 2, pp.595-605, 2016.

Thomas, R. J., Krehbiel, P. R., Rison, W., Hunyady, S. J., Winn, W. P., Hamlin, T., and Harlin, J.: Accuracy of the lightning mapping array, J. Geophys. Res., 109, D14207, doi:10.1029/2004JD004549, 2004

van der Velde, O., and Montanya, J.: Asymmetries in bidirectional leader development of lightning flashes, J. Geophys. Res. Atmos., 118, 13, 504–13, doi:10.1002/2013JD020257, 2013.

Wacker, R., and Orville, R.: Changes in measured lightning flash count and return stroke peak current after the 1994 U.S. National Lightning Detection Network Upgrade: 1. Observations, Journal of Geophysical Research, 104(D2), 2151-2157, 1999a.

Wacker, R., and Orville, R.: Changes in measured lightning flash count and return stroke peak current after the 1994 U.S. National Lightning Detection Network Upgrade: 2. Theory, Journal of Geophysical Research, 104(D2), 2159-2162., 1999b.

Weidman, C. D., Krider, E. P., and Uman, M. A.: Lightning amplitude spectra in the intveral from 100 kHz to 20 MHz, Geophys. Res. Lett., vol. 8, no. 8, pp. 931-934, 1981.

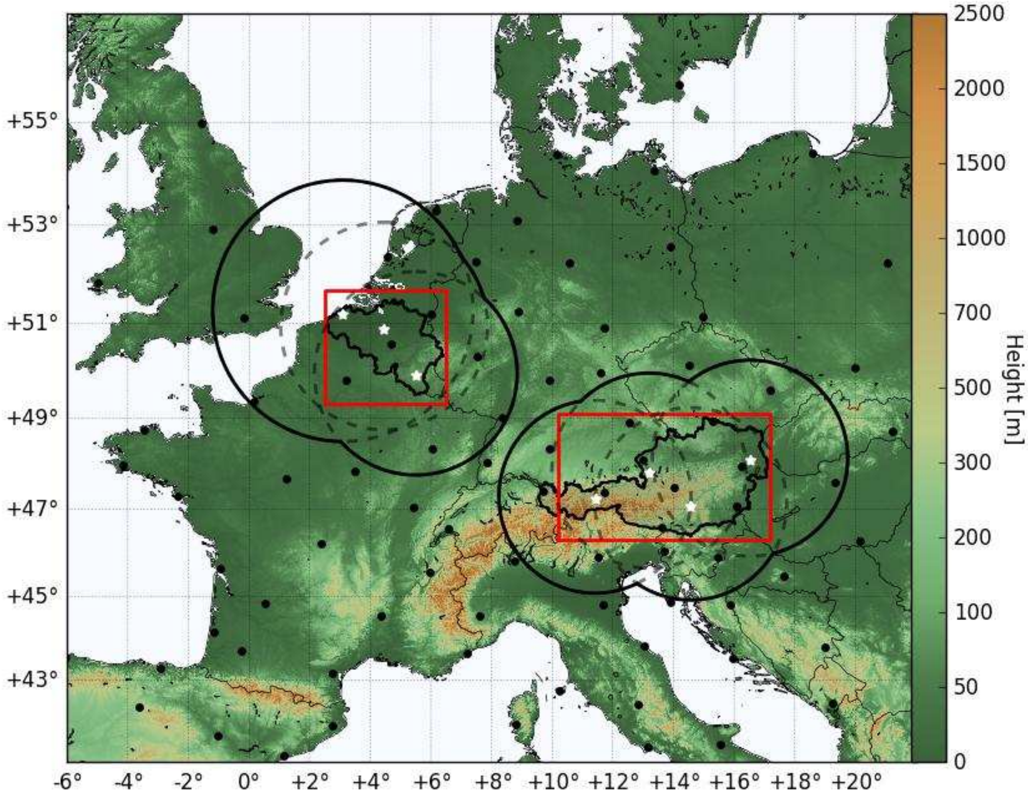


Figure 1: The locations of the EUCLID sensors in the domain are indicated (black dots), as well as the positions of the radars (white stars) together with the respective collective detection range in Belgium and Austria. The red boxes indicate the two areas that are used in this study.

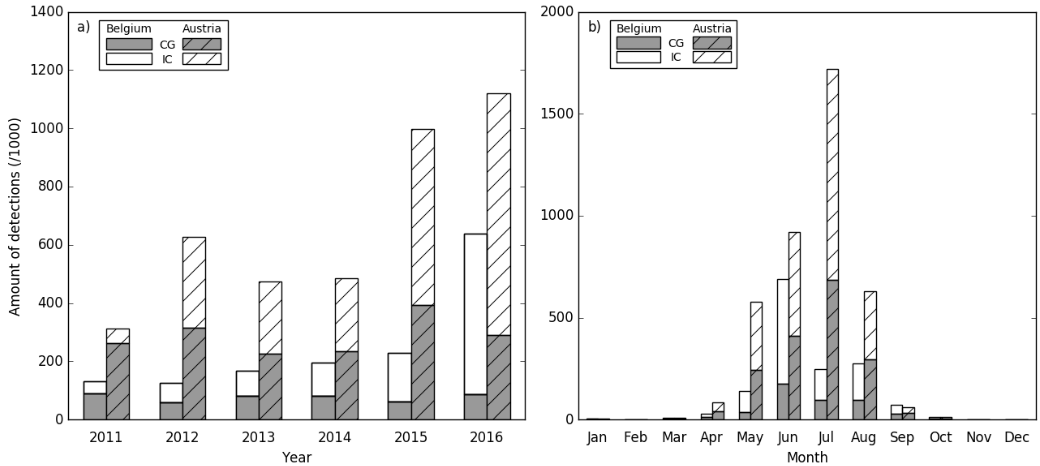


Figure 2: Distribution of the a) annual and b) monthly CG and IC counts as observed within the areas indicated over Belgium and Austria in Fig. 1.

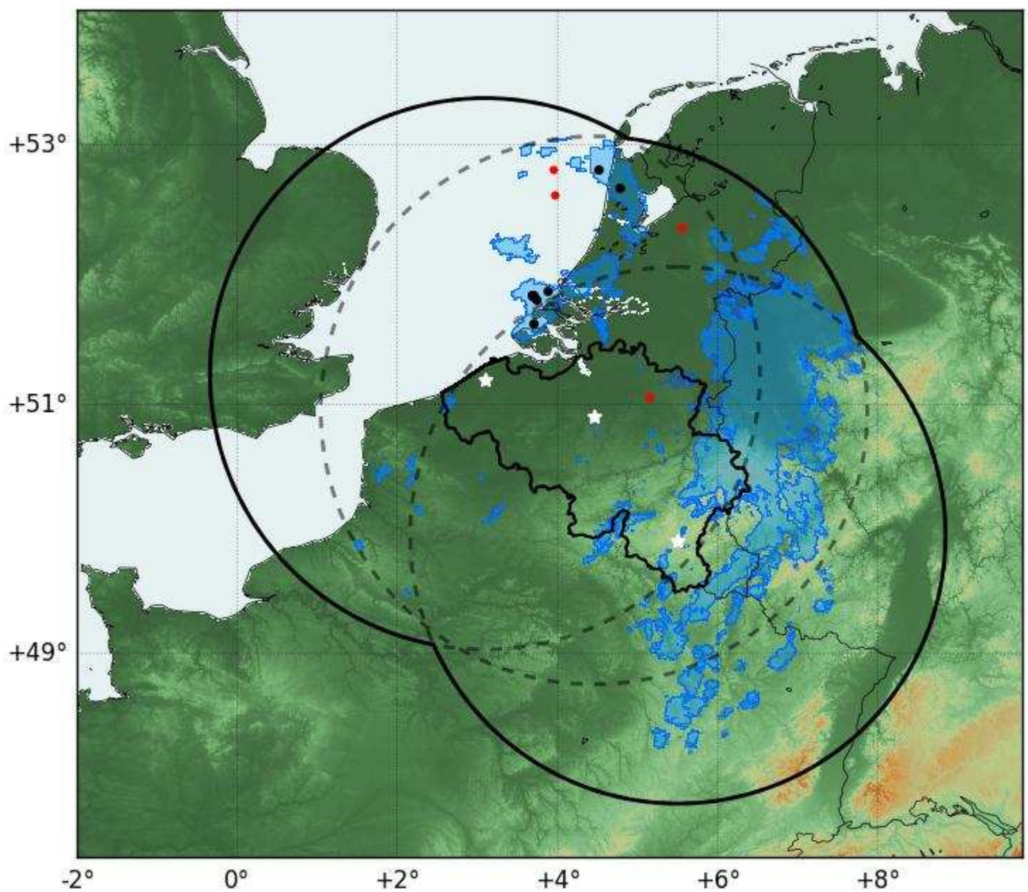


Figure 3: Example of a 5-min precipitation field, superimposed with the lightning events within the time interval. The correctly located events are indicated as black dots, whereas the derived outliers are plotted in red. For clarity the underlying precipitation field has been given the same value above the applied threshold of 0.2 mm/h everywhere.

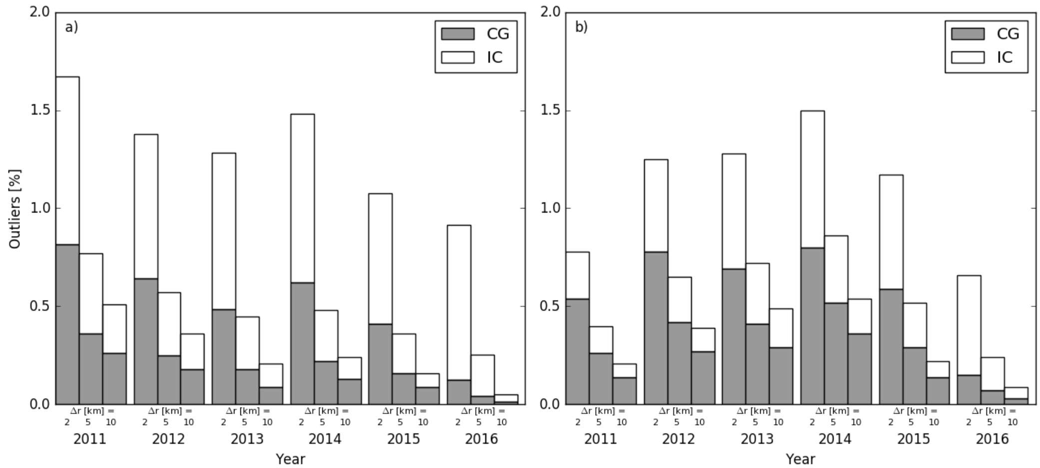


Figure 4: Annual variation of outliers in a) Belgium and b) Austria, based on cloud-to-ground (CG) and intracloud (IC) events, for search radii of 2, 5, and 10 km, respectively.

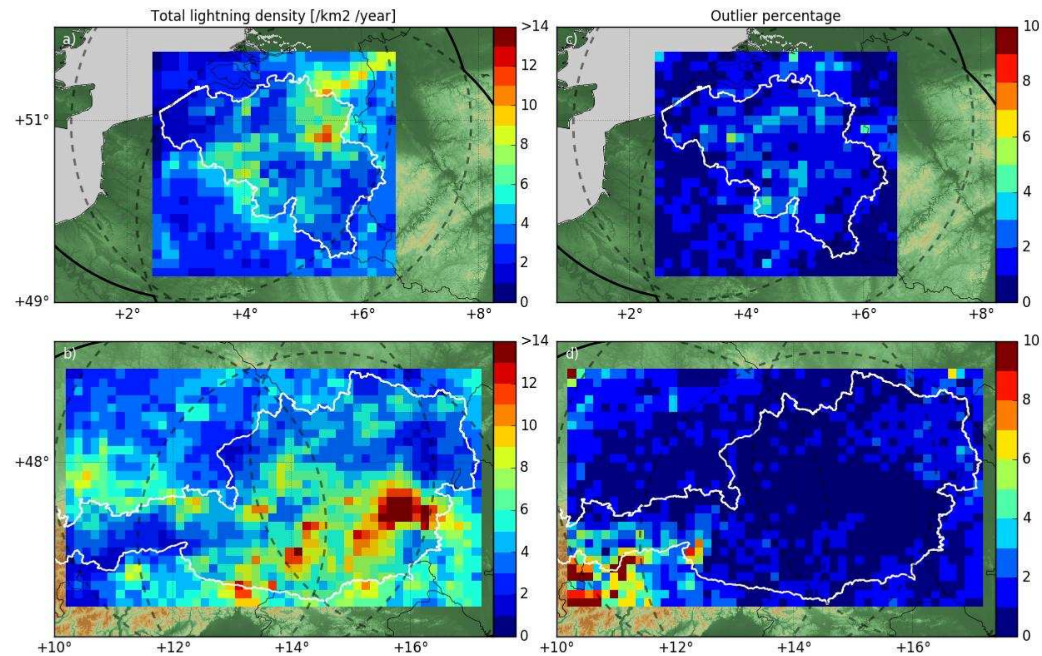


Figure 5: The left display the 6-year mean annual total lightning event (CG strokes + IC pulses) density on a 10 x 10 km^2 grid in Belgium (top) and Austria (bottom), whereas the spatial distribution of the total percentage of outliers is plotted in the right panels for an adopted search radius of 2 km.

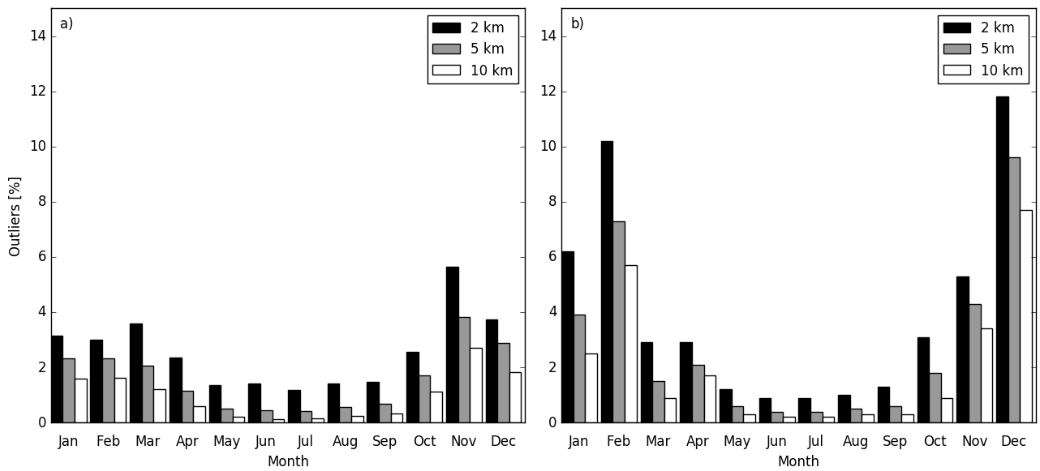


Figure 6: Monthly distribution of the total (CG + IC) percentage of outliers in a) Belgium and b) Austria, for search radii of 2, 5, and 10 km, respectively.

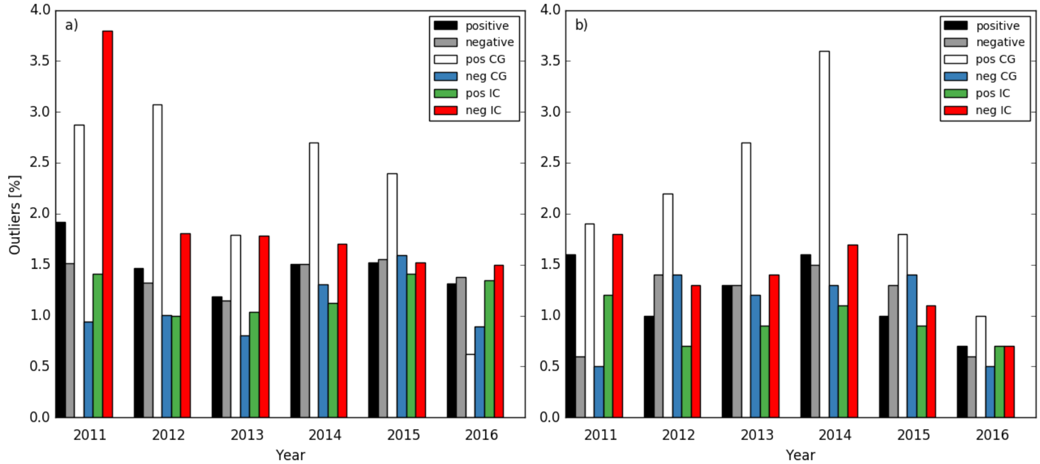


Figure 7: Percentage of outliers versus event type in a) Belgium and b) Austria, for a search radius of 2 km.

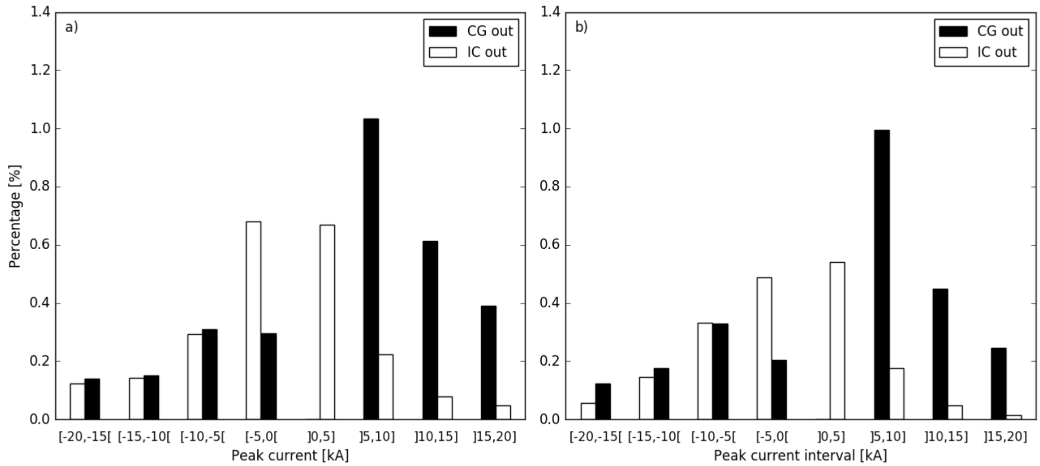


Figure 8: Percentage of outliers as a function of peak current in a) Belgium and b) Austria, for a search radius of 2 km.

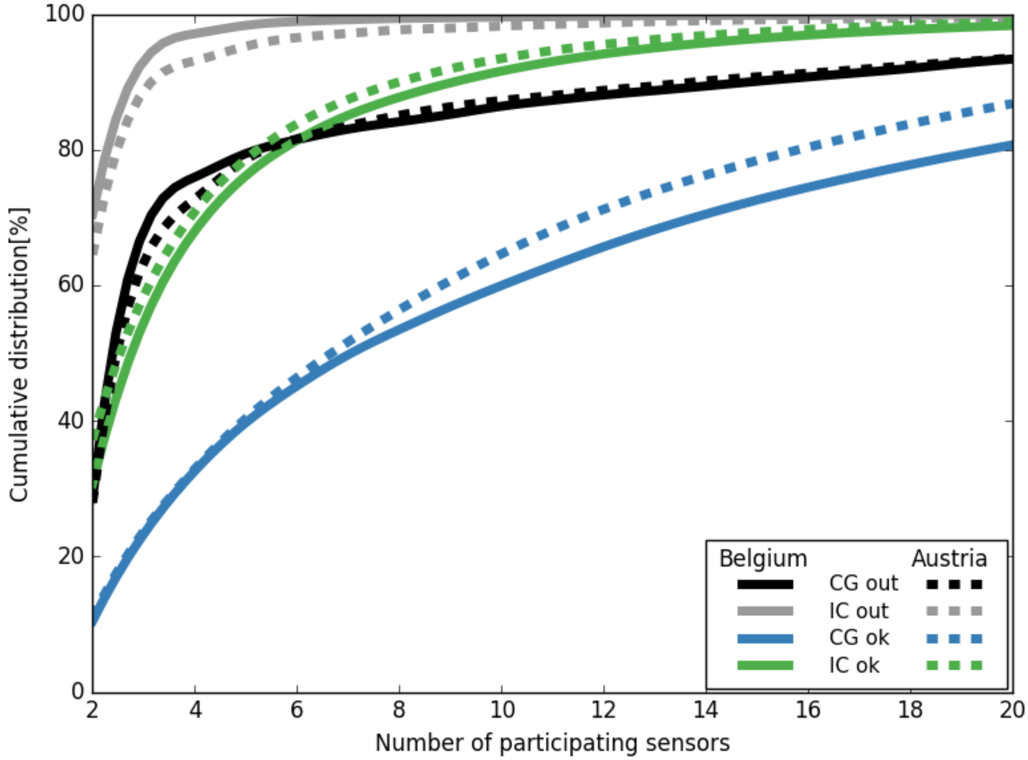


Figure 9: Cumulative distribution of the number of sensors participating in a solution for CG and IC outliers ('out')) and correctly ('ok') located events.

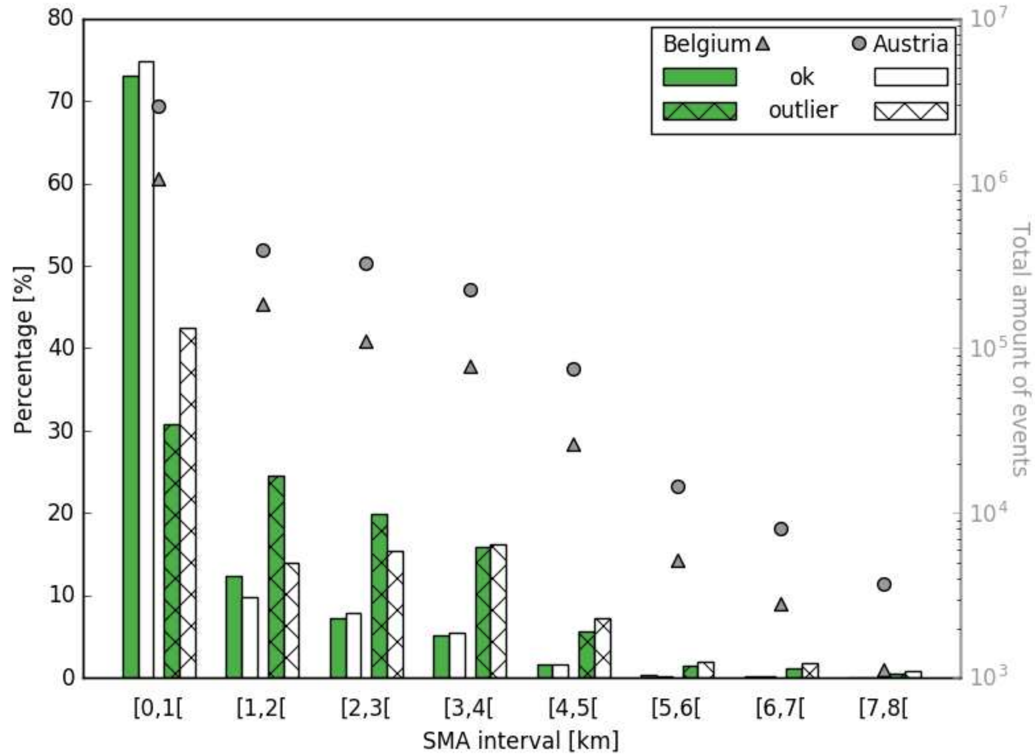


Figure 10: Distribution of the semi-major axis (SMA) of the outliers and ok events in Belgium and Austria, for a search radius of 2 km. In addition, the total amount of events per SMA interval is indicated as grey circles and triangles.

Comment [DP30]: added

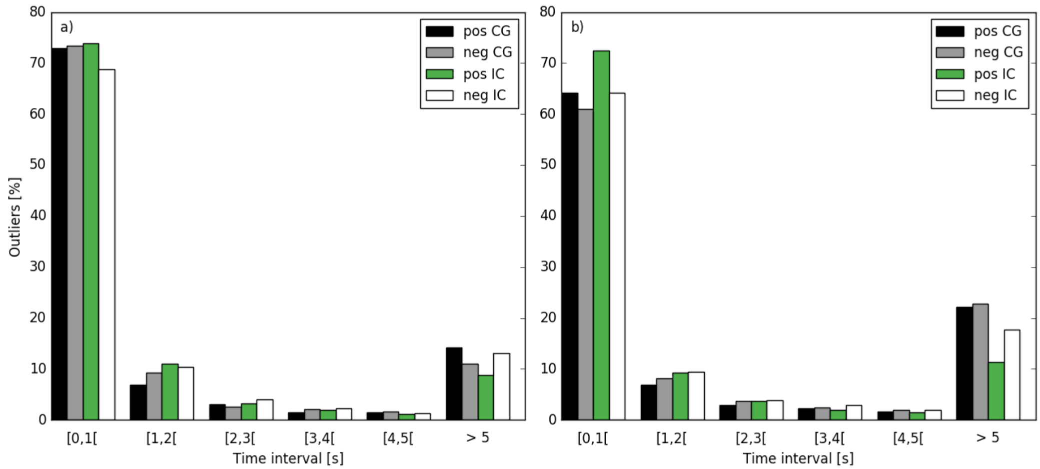


Figure 11: Distribution of the time difference between the outliers and its closest (in time) correctly located event.