

Earth and Space Science



TECHNICAL REPORTS: METHODS

10.1029/2021EA001713

Key Points:

- A LMA radiation source connection method considering the source density and repeated discharges in channels is proposed
- The stability of lightning channel scale under different radiation source densities caused by different detection capabilities is improved
- The obtained discharge scale can track the lightning discharge process and better characterize the impact range of lightning discharge

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y. Zhang and Y. Zhang,
zhangyang@cma.gov.cn;
zhangyijun@fudan.edu.cn

Citation:

Li, Y., Zhang, Y., Zhang, Y., & Krehbiel, P. R. (2021). A new method for connecting the radiation sources of lightning discharge extension channels. *Earth and Space Science*, 8, e2021EA001713. <https://doi.org/10.1029/2021EA001713>

Received 19 FEB 2021

Accepted 23 MAY 2021

A New Method for Connecting the Radiation Sources of Lightning Discharge Extension Channels

YuRui Li^{1,2}, Yang Zhang² , YiJun Zhang¹, and Paul R. Krehbiel³ 

¹Department of Atmospheric and Oceanic Sciences and Institute of Atmospheric Sciences, Fudan University, Shanghai, China, ²State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, China, ³Langmuir Laboratory for Atmospheric Research, Geophysical Research Center, New Mexico Institute of Mining and Technology, Socorro, NM, USA

Abstract The connections of lightning radiation sources along channels are greatly affected by the radiation source density, and the channel length is geometrically scaled when neglecting repeated discharges in the same channel. Based on lightning mapping array (LMA) three-dimensional radiation source location data of two lightning flashes, this study presents a radiation source connection method considering the source density and repeated discharges in channels that includes two steps: the connection of radiation sources and the connection of segments. After increasing the spatial connection threshold determined by the source density, the stability of the channel scale under different detection capabilities (source densities) is improved compared with that of the traditional fixed spatial threshold method. The channel growth rate of the low-density case reaches 120.99%, which is close to the real situation, and the connection shape is consistent with the real situation. For repeated discharge paths, by limiting the time interval of the radiation source connection, the optimized threshold obtained in this paper can distinguish between discharges occurring in the same channel at different times. Compared with the geometric scale, the discharge scale is significantly larger (~2.8 times the geometric scale in one case) and can better characterize both the lightning discharge process and the affected area. These comprehensive results show that the proposed method can reduce the number of incorrect connections, increase the channel length, and obtain a more realistic discharge scale.

1. Introduction

Each lightning flash takes place within a certain spatial range and lasts for a period of time, with repeated breakdowns and charge transfers in some channels. The various effects of lightning discharges are closely related to the discharge range, which is often regarded as the lightning scale. To date, many studies have been performed on the correlations between the lightning scale and chemical effects, charge transfer amount and thunderstorm configuration.

The amount of nitrogen oxides greatly depends on the lightning scale, and thus, the lightning scale must be accurately characterized before calculating the amount of nitrogen oxides (Finney et al., 2016; Koshak et al., 2014; Wild, 2007). The charge transfer amount and extent of damage during the lightning process are also related to the lightning scale (Hager et al., 2007; Jacobson & Krider, 1976; Krehbiel et al., 1979). The larger the lightning scale is, the greater the threat to electronic equipment, tall buildings, windmills, and people (Yasuda et al., 2012). The lightning scale is usually determined by the thunderstorm configuration: in a supercell thunderstorm, the scale is often smaller near the strong ascending zone and normally larger near the front cloud anvil (Bruning & MacGorman, 2013; Z. Zhang et al., 2017). Moreover, there is an inverse correlation between the diurnal variation of the lightning scale and that of the lightning frequency (Chronis et al., 2015).

At present, research on the scale of lightning is based mostly on optical data and very high-frequency (VHF) positioning data. With the development of lightning mapping array (LMA) technology, the channel characteristics of lightning can be readily characterized, and both the two-dimensional convex hull area of the lightning projection and the three-dimensional convex hull volume can be obtained (Bruning & Thomas, 2015). Bruning and Thomas reported that using the lightning coverage area (using either the convex hull volume or the convex hull area) to estimate the lightning channel length is more stable than other existing approaches; that is, the channel length will not vary considerably as the number of stations and

© 2021. The Authors. Earth and Space Science published by Wiley Periodicals LLC on behalf of American Geophysical Union.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](https://creativecommons.org/licenses/by/4.0/), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

chi-square value change. Hager et al. (2007) developed a pulse graph algorithm to calculate the amount of charge transfer that directly connects each spatially adjacent radiation source and then calculates the sum of the connected line segments as the lightning channel length. Lapiere et al. (2014) used the above algorithm to calculate the extension length of the lightning channel during the continuous current phase and concluded that the occurrence and duration of continuous current may not be determined by the growth rate of the channel length of the positive leader. Koshak et al. (2014) and R. Zhang et al. (2014) extracted the mean coordinates of the radiation source in the three-dimensional space cell grid and connected them to the channel to calculate the geometric length. In their study, the space occupied by the radiation sources is divided into a series of three-dimensional grid cells of equal size. When the number of radiation sources in a grid cell is equal to or more than a threshold, the arithmetic average coordinate values of the sources in the cell are calculated, and then, the least distance radiation source is connected among the averaged radiation sources. Because the connection process uses the average position in the grid, it may lead to the loss of some small branches of the lightning channel. Although the length of small branches is very short, the number is large, so the total length cannot be ignored. Besides, when the time threshold is not considered, only the geometric scale of the channel can be obtained.

Based on optical data from the Lightning Imaging Sensor (LIS), which can identify the luminous lightning channel and determine the area of lightning both on land and at sea, Peterson, Deierling, et al. (2017) proposed three parameters, namely, elongation, characteristic propagation, and observed propagation, to describe the scale of lightning and concluded that the extension of a lightning channel is longer over the ocean than on land. More recently, using data from the Geostationary Lightning Mapper (GLM), Peterson, Lang, et al. (2020) regarded the horizontal distance and duration of lightning as the lightning scale and updated the existing records with the largest known lightning scale to date. Nevertheless, although LIS and GLM data can image the luminous channel, the optical method cannot completely characterize the lightning scale due to the occlusion of clouds and the limitation on the luminous intensity of weak discharge channels.

The above methods focus on characterizing the scale of geometric channels. However, the extension of a lightning channel is related to the physical processes of lightning, i.e., not only to the forward extension but also to repeated local breakdown and recoil processes in the existing channel (e.g., the K process and the dart leader before the subsequent return stroke). The traditional convex hull area and volume expand the range of discharge, whereas the traditional spatial geometric channel neglects repeated discharges in the same channel. Therefore, we propose a new method to characterize all the lightning extension pathways that can better link the lightning scale with the physical processes occurring during lightning discharge, thereby obtaining a more accurate extension channel length.

2. Data

In this study, a new method for connecting lightning radiation sources is proposed based on the three-dimensional location data of two natural lightning processes from a lightning imaging array (LMA) managed by the Langmuir Laboratory of the New Mexico Institute of Mining and Technology. More than 10 LMA stations were operated over an area with a diameter of 60 km surrounding the Langmuir Laboratory on the day in question. The LMA data have a temporal resolution of 80 μ s and a positioning accuracy of tens of meters within the station network (Rison et al., 1999; Thomas et al., 2004) and a strict standard of noise reduction is used (number of stations ≥ 7 ; reduced chi-square ≤ 2) to ensure the reliability of the location data. The continuous broadband interferometer (INTF) system (Stock et al., 2014) is used to precisely locate the entire discharge process with high temporal resolution (submicrosecond). Combined with the fast electric field waveform captured by this system, a detailed analysis of the lightning discharge process can be performed, the results of which can be used to test the method for connecting radiation sources as an objective fact. The two lightning processes used in this paper occurred at 19:54:43 UTC and 20:02:56 UTC on August 5, 2015 (hereinafter referred to as Flash1 and Flash2, respectively), and plots illustrating the height variation over time and the spatial distributions of the radiation sources are provided in Figure S1. Both Flash1 and Flash2 occurred in the middle of the LMA network, so they were positioned with high accuracy. Flash1 is a typical cloud-to-ground (CG) flash, lasting ~ 600 ms, with 1,652 radiation sources being located extending 6 km in the east-west direction and 12 km in the north-south direction. Flash2 is a typical cloud-to-cloud (CC) flash,

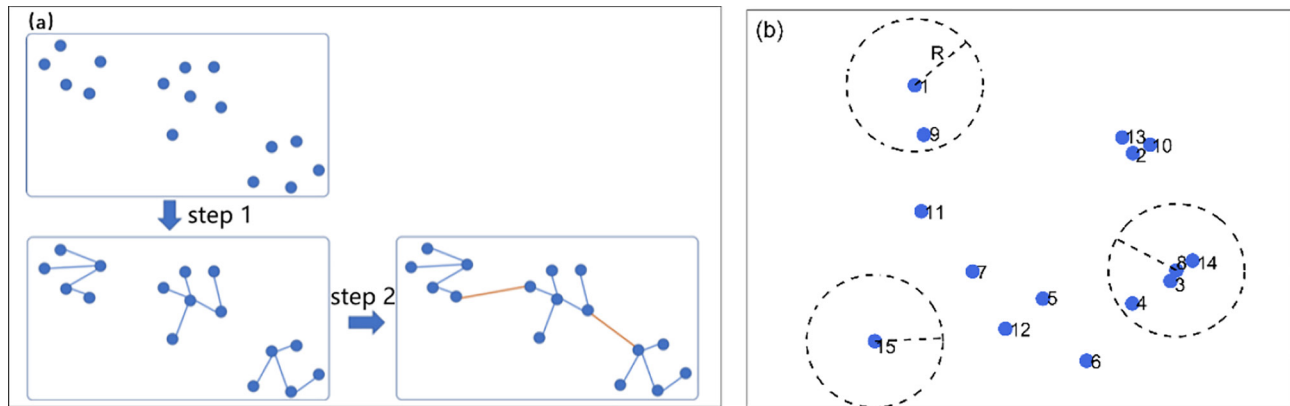


Figure 1. Schematic diagrams of connecting radiation sources and the definition of lightning density. The radiation sources are represented as blue dots. (a) Schematic diagram of connecting radiation sources. The blue and orange lines represent the first and second steps of the connection algorithm, respectively. (b) Schematic diagram of the lightning density definition. R is the radius, and the point densities of vertices 1 through 15 are 1, 2, 3, 2, 1, 0, 1, 3, 1, 2, 0, 2, 2, 2, and 0. The total density of these 15 scattered points is 1.47.

lasting ~ 500 ms, with 1,980 radiation sources being located extending 7 km in the east-west direction and 9 km in the north-south direction.

3. Method for Connecting Radiation Sources for the Discharge Channel Extension

Referring to the pulse graph algorithm for connecting LMA radiation sources (Hager et al., 2007), a new method for connecting radiation sources is developed in this study. This method comprehensively considers both the connection space threshold determined by the radiation source density and the time threshold restricted by the lightning discharge process and connects as many lightning discharge channels as possible, thus obtaining a discharge channel instead of a geometric channel. In this study, the analysis method of graph theory is introduced into the description of lightning channels (Harary, 1969), and simple graphs are used to reflect complex structural relationships. For a flash, the radiation sources and the connecting lines are regarded as mathematical objects known as “vertices” and “edges,” respectively, and the edges are connected to these vertices in two steps.

The first step is shown in Figure 1a, in which adjacent vertices (radiation sources) in a specific space are connected into line segments. When connecting adjacent vertices, the connected vertices are limited by both the time threshold and the connection space threshold. Of these two thresholds, the time threshold should be able to partition the subdischarge process, separate the radiation sources that discharge in the same channel, and avoid connecting radiation sources that discharge in close time at both ends of the channel.

During the test of time threshold sensitivity, the algorithm will connect the nearest radiation source within the minimum time threshold. Because the time interval of discharge at both ends of the lightning channel may be very short, when the time threshold is set to a minimum, there are few close-range radiation sources around the discharge at one end of the channel within this extremely small-time threshold, which makes it possible that the radiation source at one end of the channel is connected with the radiation source at the other end of the channel which discharges almost at the same time (Figure S4 also confirms it). When the time threshold increases, there will be more close radiation sources near the discharge end, and the algorithm will select these close radiation sources near the discharge end instead of the radiation sources discharged at the other end of the channel to connect. In this way, the appropriate time threshold can avoid the connection of radiation sources discharged at the same time at both ends of the channel. For the spatial threshold, since the number and density of radiation sources will be affected by many factors (such as the radiation power, station network layout, and number of stations), if only uniform spatial parameters are used to connect the radiation sources, the connections for low-density lightning cases will be very different from the real situation. Therefore, to reduce the influence of the radiation source density on the channel

connection results, we adopt a spatial connection threshold based on the lightning radiation source density. The definition of the lightning radiation source density is as follows: first, we define the point density, which refers to the number of sources in a sphere with a given point as the center of the sphere and a radius of $r = 100$ m; thus, each source is given a point density. Then, the total density of lightning radiation sources is defined as the sum of the point densities of all sources divided by the total number of radiation sources. For example, in Figure 1b, the point densities of vertex 1, vertex 8, and vertex 15 are 1, 3, and 0, respectively, and the total density of these 15 scattered points is 1.47. This density definition assigns each vertex a corresponding point density, so the spatial connection threshold parameter can be subsequently set according to the density, the size of the radius will affect only the density value of the point, and the relative relationship among the point densities will not change substantially. However, if the total number of radiation sources is divided by the area of the convex hull as the lightning density, the result is greatly affected by the lightning shape; in particular, when the shape of lightning tends to be a straight line, there is no area, and thus, the density cannot be calculated. In addition, although the spatial threshold can partially avoid the connection of the radiation sources at both ends of the channel, when the distance between the two ends of the channel is less than the spatial threshold and the time threshold is also extremely small, the density-based spatial threshold may not be able to avoid the connection of radiation sources that discharge at a similar time at both ends of the channel. Therefore, different discharge processes can be effectively distinguished by considering the temporal and spatial thresholds of different density lightning.

The second step is to connect the small segments connected in the first step, i.e., to connect the nearest vertices within 2 km of the two segments. Similar to the separation of segments described by Hager et al. (2007), we also believe that modifying the spatial connection threshold (node_connect) in the first step will not only change the connections of most sources but also affect the connections of small segments in the second step of the algorithm. When the vertices are connected differently, the distances between small segments will be affected; hence, the connections in the second step of the algorithm will be different. Therefore, in this paper, we consider changes in the spatial connection threshold with the density only in the first step.

4. Results

4.1. Connections Using a Density-Based Spatial Threshold

Based on Flash1 and Flash2, we study the specific scheme of adjusting parameters according to the density and compare the results with the real connections and scale (main channel length) of the pulse graph algorithm.

First, the radiation sources of the two flashes are downsampled to construct low-density radiation source cases. For the high-density case, the pulse graph algorithm is considered capable of obtaining more real connections (Hager et al., 2007). Therefore, the connections in the high-density case (not downsampled) are regarded as the benchmark for comparing the connection results of the other samples and is called the real connection in the following discussion. We sample Flash1 and Flash2 every 1, 3, 7, 15, and 31 radiation sources to obtain samples with different densities. The sample obtained after sampling every radiation source is called sample1, the sample obtained after sampling every three radiation sources is called sample3, and so on, such that the sample after sampling every 15 radiation sources is called sample15. According to the density, the sample is divided into high, medium, and low densities: greater than or equal to 1 but less than 3 is low density; greater than or equal to 3 but less than 5 is medium density, and greater than or equal to 5 is high density. Through a sensitivity test of these samples, the specific parameters of the spatial threshold are obtained. For the high-density case, the spatial connection threshold (node_connect) is 1,000 m, which is consistent with the original parameters of the pulse graph algorithm; for the medium-density case, the spatial connection threshold (node_connect) is set to 1,240 m; for the low-density case, the spatial connection threshold (node_connect) is set to 2,070 m. The process of adjusting these parameters to obtain a more reasonable threshold and the comparison results can be found in Figures S2 and S3.

Based on the new density-based spatial threshold, different sample radiation sources are connected to calculate the geometric scale. As shown in Table 1, the density of A_sample1 is 3.49, and that of B_sample1 is 3.25; both are classified as medium density. The geometric scales of A_sample1 and B_sample1 are 64.56 and 75.04 km, respectively, by using the original fixed parameters, while the geometric scales of A_sample1

Table 1
The Number of LMA Radiation Sources, the Density, and the Scale (Main Channel Length) Before and After Adjusting the Spatial Threshold Parameter of Flash1 and Flash2 and the Samples With Different Densities

ID	Number	Density	Main channel length (previous parameters)	Main channel length (density-based parameters)
Flash1	1,652	6.05	93.23	93.23
A_sample1	826	3.49	64.56	65.63
A_sample3	413	2.12	50.61	54.6
A_sample7	207	1.47	43.19	49.71
A_sample15	104	1.15	32.13	40.77
A_sample31	52	1	13.33	29.47
Flash2	1,980	5.81	97.96	97.96
B_sample1	990	3.25	75.04	76.05
B_sample3	495	1.96	57.87	59.93
B_sample7	248	1.42	48.58	51.60
B_sample15	124	1.23	29.82	37.85
B_sample31	62	1.10	16.47	25.19

and B_sample1 obtained by using the density-based parameters are 65.63 and 75.05 km, respectively, and the growth rates of the scale are 6.95% and 0.86%, respectively. The densities of A_sample3, A_sample7, A_sample15, and A_sample31 are 2.21, 1.47, 1.15, and 1.00, respectively, which are all classified as low density. The corresponding geometric scales are 50.61, 43.19, 32.13, and 13.33 km by using the original fixed parameters and 54.60, 49.71, 40.77, and 29.47 km, respectively, by using the density-based parameters, and the growth rates of the scale are 7.30%, 15.08%, 26.90%, and 120.99%, respectively. The densities of B_sample3, B_sample7, B_sample15, and B_sample31 are 1.96, 1.42, 1.23, and 1.10, respectively. The corresponding geometric scales are 57.87, 48.58, 29.82, and 16.47 km by using the original fixed parameters and 59.93, 51.60, 37.85, and 25.19 km, respectively, by using density-based parameters, and the growth rates of the scale are 3.57%, 6.21%, 26.92%, and 52.94%, respectively. The lower the density of the sample is, the larger the scale difference before and after adjusting the parameters; moreover, the larger the growth rate of the scale after adjusting the parameters, the closer the scale after sampling is to that before sampling.

In addition to judging the similarity with the real situation from the scale, the number of isolated sources, the correlation coefficient, and other numerical values, the consistency between the connection results with both the original fixed parameters and the density-based parameters in the connection graph and the real connection is further analyzed. Figure 2 shows a comparison between the partial connections of A_sample31 (Figure 2a) and B_sample31 (Figure 2b) by using the pulse graph (fixed parameters) algorithm and the algorithm with density-based parameters and the real connections (both classified as low density). In Figure 2a, the black line is the real connection of the selected part, and the blue line is the connection based on the fixed threshold algorithm; the latter completely deviates from the real connection. In contrast, the orange line is the connection obtained by using the density-based parameters, and the results are more consistent with the real connection.

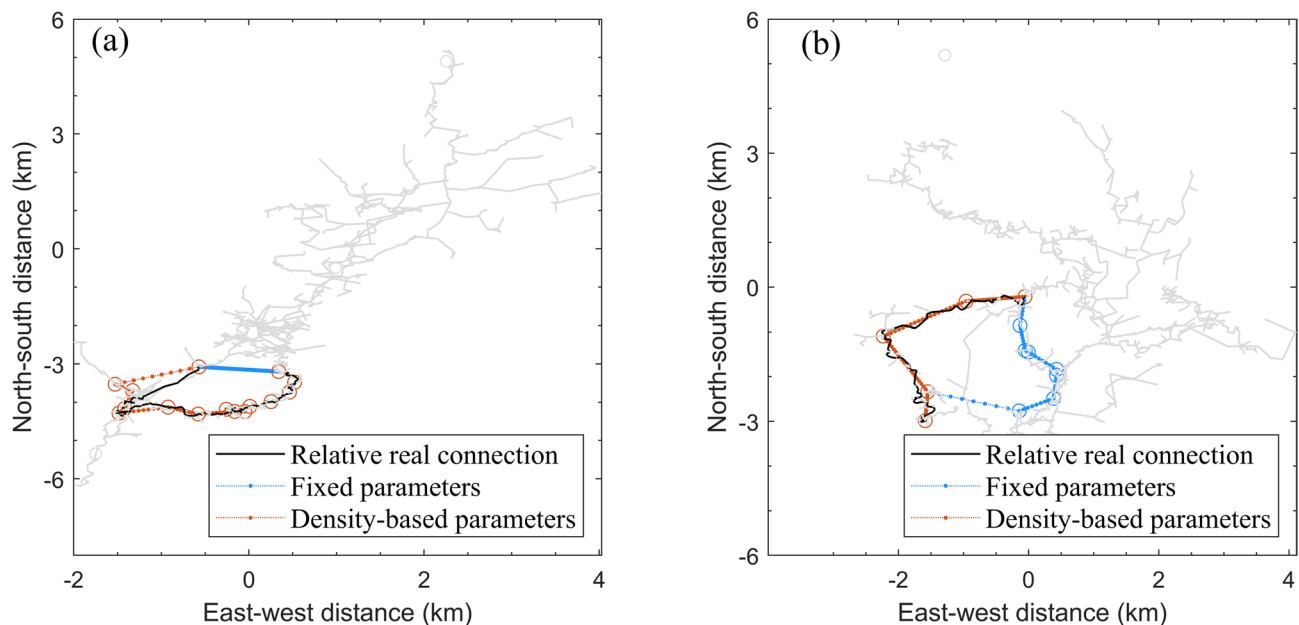


Figure 2. Comparison between the connections of low-density samples before and after adjusting the parameters. The black line is the actual connection, the blue line is the connection with fixed parameters, and the red line is the connection with the density-based parameters. (a) Comparison of the connections of A_sample31 sampled from Flash1 using the original fixed parameters and the density-based parameters superimposed onto the real connection (grayscale image) of Flash1. (b) Same as (a) but for B_sample31 sampled from Flash2 using the original fixed parameters and the density-based parameters.

Moreover, Figure 2b indicates that the connection based on the density-based parameters are more consistent with the real connection, while the connection based on fixed parameters deviates from the real connection.

4.2. Connections Considering Repeated Discharges

Lightning discharges often occur repeatedly in an existing channel. Therefore, the geometric scale of the lightning channel cannot characterize the whole lightning discharge process. When connecting radiation sources, the durations of subdischarge events should be considered to distinguish different discharges in existing channels. First, we analyze the 100 ms discharge process from 20:02:57.2 UTC to 20:02:57.3 UTC in Flash2 to check whether the connection of the radiation source using the reference algorithm (Hager et al., 2007) is reasonable by combining the LMA and INTF data (during the K process and repeated discharges within the same channel, LMA data can locate only a few radiation sources, while INTF data can locate more radiation sources, which is conducive to analyzing the real discharge process). The criteria for determining the reasonableness of a source connection are as follows: (a) simultaneous discharges at both ends of a discharge channel cannot be connected and (b) the connection channel of a unidirectional extension corresponds to a single discharge event.

Figures 3a–3c show the INTF observation result within 100 ms during Flash2 and the corresponding fast antenna (FA) waveform. The radiation sources of the discharge event correspond to an obviously fast change in the electric field. Combining the FA waveform with the spatial distribution of the radiation source reveals that the channel developed in the cloud and that five K discharge processes occurred successively, almost all within the same channel. The detailed development of the K process is depicted in Figures 3d–3h. For each K process, an increase in the elevation angle of the radiation source indicates that the discharge developed from the far end to the observation position. These K processes clearly developed within the same channel or along part of the same channel. The LMA radiation sources with serial numbers of 1,882 and 1,885, 1,933, 1,941 and 1,950 occurred in the first, third, fourth, and fifth K processes, respectively, corresponding to the 148th and 484th, 880th, 1067th, and 1265th INTF radiation sources, respectively, within >100 ms.

By using the pulse graph algorithm (without considering the time threshold), some of the LMA radiation sources (serial numbers 1882, 1885, 1933, 1941, and 1950 occurring in multiple different K processes) are directly connected. Figure 3i shows the connections of the LMA sources within 100 ms; the serial numbers of the irrationally connected sources are marked near the sources. Figure 3j shows the distribution of LMA sources and the corresponding FA waveform. According to Figure 3i, the 1882nd source is connected to the 1941st and 1950th sources, and the 1885th and 1933rd sources are also connected. Through the above analysis, the 1882nd, 1941st, and 1950th sources occurred in different K processes; the 1885th and 1933rd sources also occurred in different K processes. The above connections ignore multiple discharge paths of repeated K discharge processes.

Based on the above analysis, we postulate that the radiation sources within a certain interval correspond to a single channel and occur within a certain spatial range. In this way, the discharges occurring repeatedly within an existing channel can be separated from the previous discharge on the same path; as a result, the connection is somewhat closer to the real connection, and the scale of the tracked lightning discharge can be obtained.

Taking Flash2 as an example, the connections under different time thresholds (1, 5, 10, and 15 ms) are discussed. The different connections vary mainly north of 3 km and along the channel of repeated discharges. When the time threshold is 1 ms, the connection is limited by time, causing the sources to develop synchronously at both ends of the channel to be connected. When the time threshold is 5 ms, the above problem is partially improved, but there are still cases where the sources at both ends of the channel that develop simultaneously are connected. When the time threshold is further increased, the above problems disappear, but a new problem appears: the sources with long intervals in the same repeated discharge channel will be connected due to the short spatial distance. Intuitive figure and table of the different parts of the connections under different time thresholds can be seen in Figure S4 and Table S1. The differences are concentrated primarily in the 250 ms interval from 20:02:56.96 UTC to 20:02:57.21 UTC. The following is a detailed analysis based on the continuous INTF data.

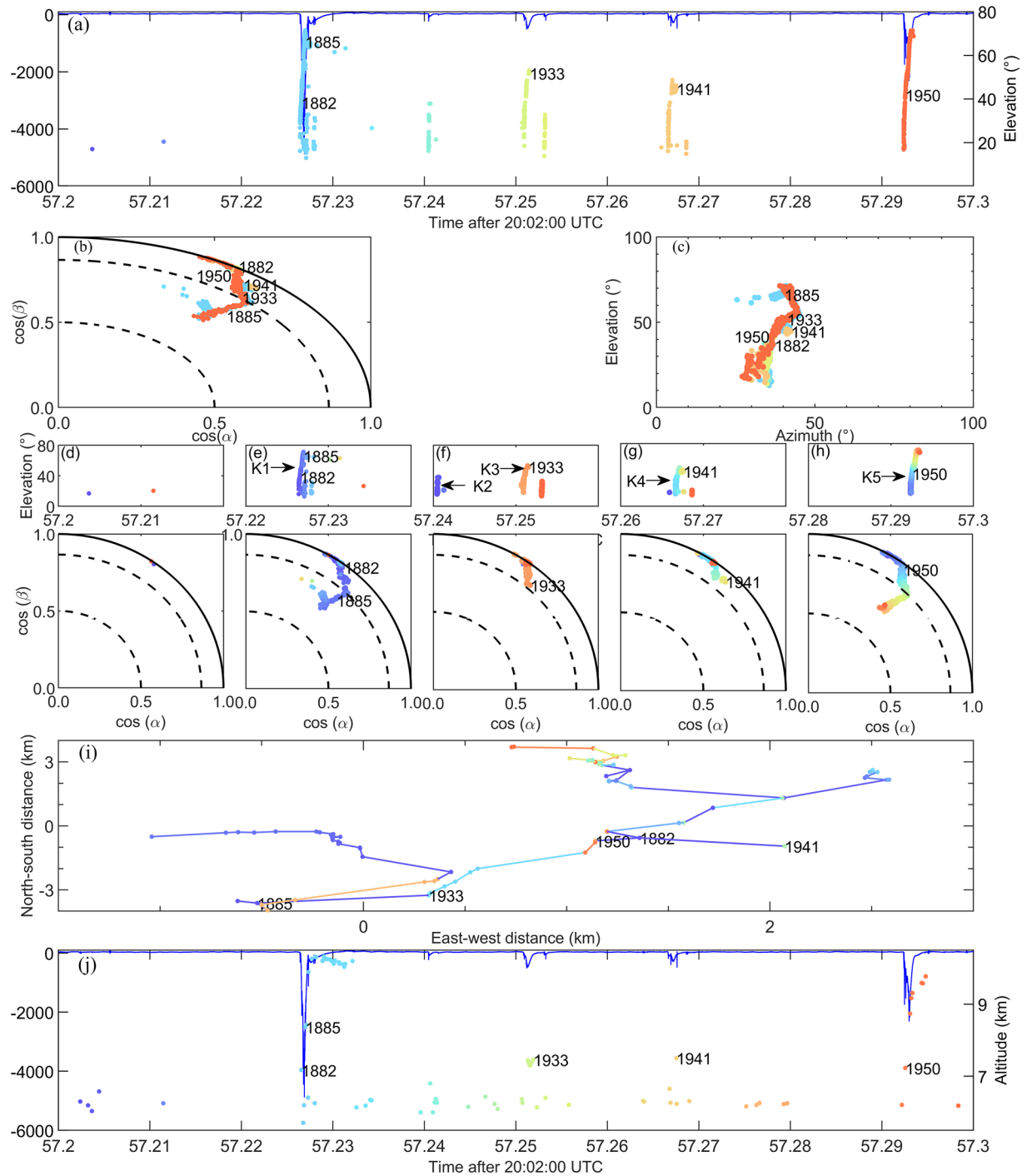


Figure 3. (a–c) INTF observations of Flash2 from 20:02:57.2 UTC to 20:02:57.3 UTC and the corresponding fast antenna (FA) waveform. (a) Elevation angle versus time superimposed onto the FA waveform (blue line). (b) Hemispherical projection. (c) Azimuth-elevation distribution. (d–h) Elevation angle versus time plots and hemispherical projections of the INTF radiation sources within 100 ms from 20:02:57.3 UTC to 20:02:57.3 UTC. (d) 57.20–57.22 UTC; (e) 57.22–57.24 UTC; (f) 57.24–57.26 UTC; (g) 57.26–57.28 UTC; (h) 57.24–57.26 UTC. (i) Connection diagram of the lightning mapping array (LMA) radiation sources from 20:02:57.2 UTC to 20:02:57.3 UTC in Flash2 using the pulse graph algorithm. (j) Altitude versus time and corresponding FA waveform from 20:02:57.2 UTC to 20:02:57.3 UTC. Radiation sources are colored by time, and the number is the serial number of the radiation source.

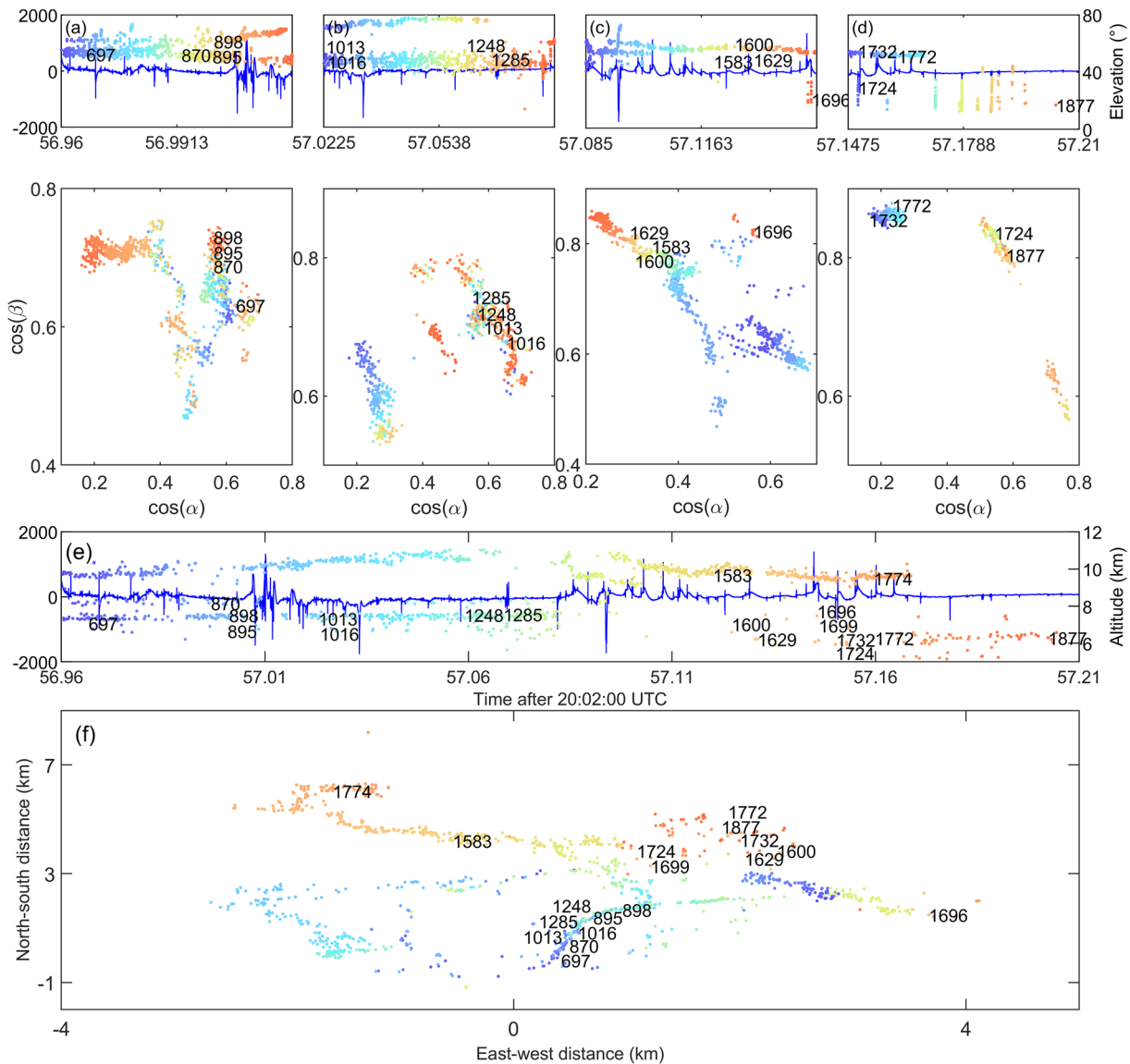


Figure 4. (a–d) INTF observations of Flash2 from 20:02:56.96 UTC to 20:02:57.21 UTC and the corresponding FA waveform (blue line). (a) 56.96–57.0225 UTC; (b) 57.0225–57.085 UTC; (c) 57.085–57.1475 UTC; (d) 57.1475–57.21 UTC. (e and f) LMA observations of Flash2 from 20:02:56.96 UTC to 20:02:57.21 UTC and the corresponding FA waveform (blue line). (e) Altitude versus time and corresponding FA waveform. (f) E-W and N-S spatial distributions. Radiation sources are colored by time, and serial numbers are marked near the radiation sources of interest.

When $T_{th} = 10$ ms, the 1772nd source is connected to the 1732nd source, the 1600th source is connected to the 1629th source, the 1699th source is connected to the 1724th source, the 1013rd source is connected to the 1016th source, and the 1248th source is connected to the 1285th source. Moreover, according to Figure 4, these radiation sources are in the development process of channel extension, so the above connection is reasonable. When $T_{th} = 1$ ms, an unreasonable connection is caused mainly by the direct connection of synchronous radiation sources at both ends of the same channel (the 1772nd and 1774th sources). Figure 4f shows that the 1772nd and the 1774th sources discharged at both ends of the same channel (similar to the concept of the bidirectional development of lightning leaders proposed by Kasemir [1960]); when $T_{th} = 5$ ms, the connection error is similar to that when $T_{th} = 1$ ms. In addition, Figure 4f shows that the 1696th and 1699th sources obviously discharged at both ends of the same channel, so such a connection is unreasonable. When $T_{th} = 15$ ms, the error is due mainly to the connection of different discharge events at different times (the 697th and 1016th sources). According to Figures 4a and 4b, the 1016th source discharged after a period of time near the position where the 697th sourced discharged; when $T_{th} = 20$ ms,

the connection error is similar to that when $T_{th} = 15$ ms. Additionally, Figures 4a and 4b show that the 1285th source also discharged after a period of time near the position where the 895th source discharged, so the above connection is unreasonable. Ultimately, setting the time threshold to 10 ms may optimally reflect the real lightning discharge situation and obtain a more accurate lightning discharge scale. (The concept of discharge scale is proposed for the first time in this study.)

For some paths of repeated discharges ($L1$, $L2$, and $L3$, details can be seen in Figure S5) in Flash2, we use the reference pulse graph algorithm and the algorithm considering repeated discharge channels to connect the radiation sources, and we compare their channel lengths (L). When the connection does not consider a repeated discharge, i.e., the time threshold is not considered, $L1$ is 3.65 km, $L2$ is 18.33 km, and $L3$ is 9.58 km; the corresponding L values after adding the time threshold are 10.08, 28.12 and 20.12 km, respectively, which are 2.76 times, 1.53 times, and 2.10 times those without the time threshold. The channel lengths obtained by using the time threshold are much larger than those obtained without considering the time threshold and thus can more accurately describe the various discharge effects of lightning.

4.3. Comprehensive Result of the Method of Connecting the Radiation Sources of Discharge Channel Extensions

After discussing the connections considering density-based parameters and repeated discharge channels, we use the radiation sources in Flash1 from 19:54:43.9 UTC to 19:54:44.0 UTC within 100 ms to test the comprehensive result of the new method of connecting the radiation sources of discharge channel extensions. Figures 5a and 5b show the observation results of the LMA and INTF data, respectively, superimposed onto the corresponding FA waveforms, and the serial numbers of the LMA radiation sources to be analyzed are marked near the sources. During this time period, 113 real LMA sources were located with a source density of 2.8939, denoting a low-density lightning case. A total of 1,623 INTF sources were located in the same time period. The hemispherical projection and azimuth-elevation distribution of the entire 100 ms process are shown in Figures 5c and 5d, respectively. Almost the entire process occurred in the previous discharge channel.

Figure 6 shows the connection results of the LMA sources in this period using both the reference pulse graph algorithm and the new algorithm, and the different parts of the connection are listed in Table S2. Figure 6 show that when the pulse graph algorithm is used, the 4th source is isolated, the 22nd source is connected to the 89th source, and the 69th source is connected to the 111th source. When using the new algorithm to connect the radiation sources, the 4th source is no longer isolated but is instead connected to the 5th source, the 22nd source is connected to the 21st source, and the 111th source is connected to the 110th source. Since the INTF sources between the 10th and 11th LMA sources are deleted after reducing noise, the 10th and 11th LMA sources correspond to the same INTF source. Similarly, the 51st and 54th LMA sources correspond to the same INTF source, as do the 58th and 59th LMA sources. Here, these sources are considered to occur during the same physical process, so we will not discuss them in detail. In the following section, we analyze the discharge process to judge whether other connections are reasonable.

Combining the LMA and INTF observation results with Figure 5, the detailed discharge process from 19:54:43.9 UTC to 19:54:44.0 UTC is analyzed. During 43.90–43.92 s, the radiation source first propagated in the cloud (the 4th and 5th LMA sources correspond to this process), and a K process lasting ~ 250 μ s occurred at 43.906 s. During this period, the elevation of the INTF source changed rapidly, which corresponds to the recoil of the radiation source in the existing channel. After a period of quiescence, a low-frequency pulse lasting ~ 400 μ s occurred at 43.911 s; at this time, the discharge was still in the existing channel. Afterward, a change in the signal of the electric field lasting ~ 900 μ s appeared at 43.915 s; the elevation angle first increased rapidly and then decreased rapidly, as shown in the time-elevation plot, which corresponded to the recoil process transmitting to the observation station first and recoiling on the existing channel, during which the 21st and 22nd LMA sources appeared. Only the obvious pulse of a low-frequency signal was present for ~ 200 μ s during 43.92–43.94 s corresponding to the radiation sources. Within ~ 18 ms after the above low-frequency signal, there was no radiation source and no obvious signal in the FA waveform until a change in the signal of the electric field appeared at 43.948 s, lasting ~ 0.45 ms; this constitutes a typical K process. A radiation source was generated again in the existing channel, and the 69th radiation source appeared during this process. Subsequently, several small low-frequency signals appeared within 20 ms, and a

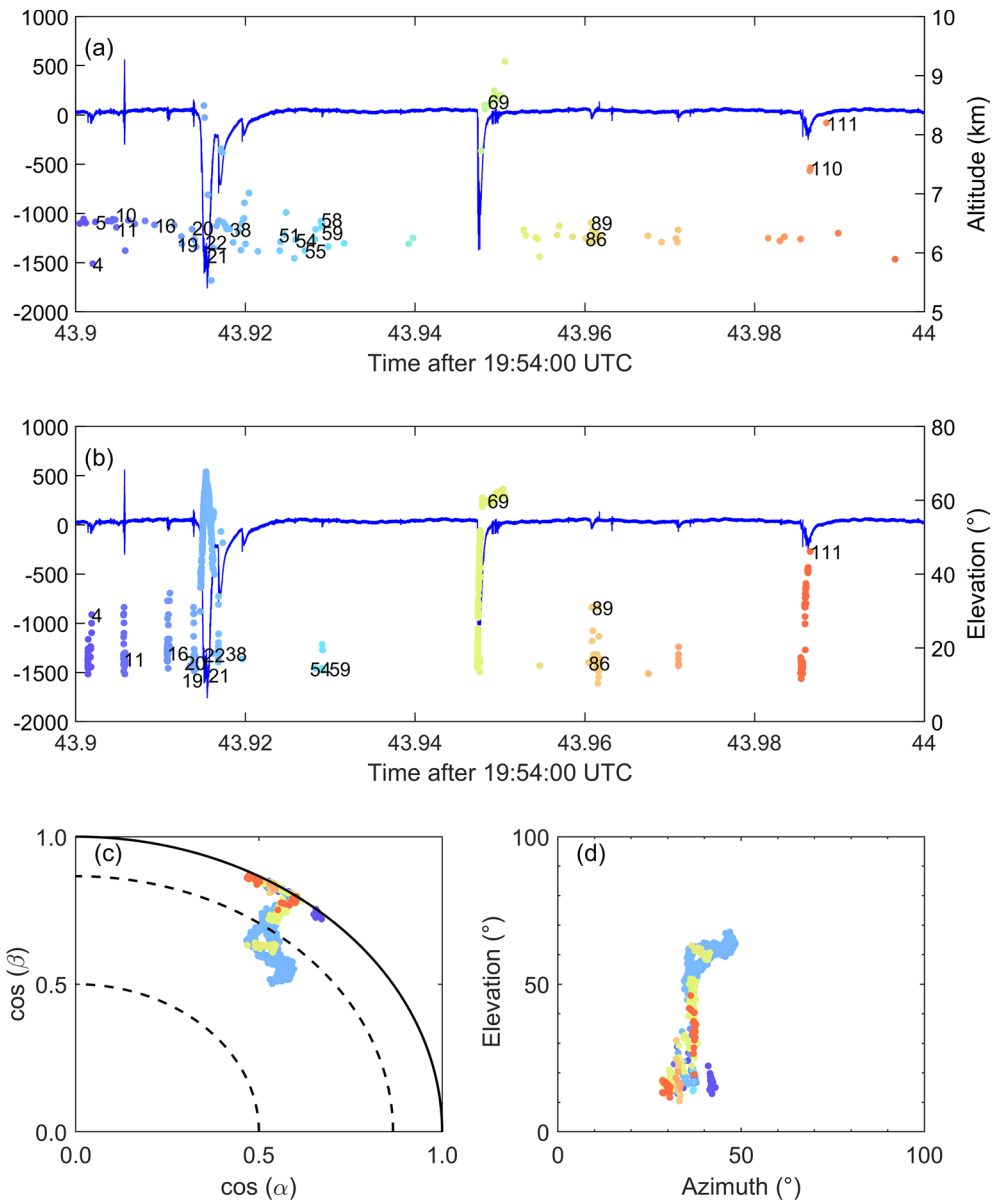


Figure 5. INTF and LMA observations of Flash1 from 19:54:43.9 UTC to 19:54:44.0 UTC and the corresponding FA waveform (blue line). Radiation sources are colored by time, and serial numbers are marked near the radiation sources of interest. (a) Altitude-time distribution of the LMA observations and the corresponding FA waveform. (b) Elevation-time distribution of the INTF observations and the corresponding FA waveform. (c) Hemispherical projection of the INTF observations; (d) azimuth-elevation angle distribution of the INTF observations.

radiation source appeared again at the end of the formed channel, discharging along the existing channel; the 89th LMA source appeared during this process. In the final 20 ms, the 111th LMA source appeared. An FA pulse with a duration of $\sim 180 \mu\text{s}$ appeared, and a radiation source was again generated at the previous discharge position.

Combined with the discharge process analyzed above, the 22nd and 89th sources and the 69th and 111th LMA sources did not discharge during the same physical process, so it is unreasonable to connect them by using the pulse graph algorithm. In contrast, the 4th and 5th sources, the 21st and 22nd sources, and the 110th and 111th sources all occurred in the same discharge process, and thus, they are reasonably connected. Finally, the geometric scale obtained by the pulse graph algorithm is 20.51 km, whereas the discharge

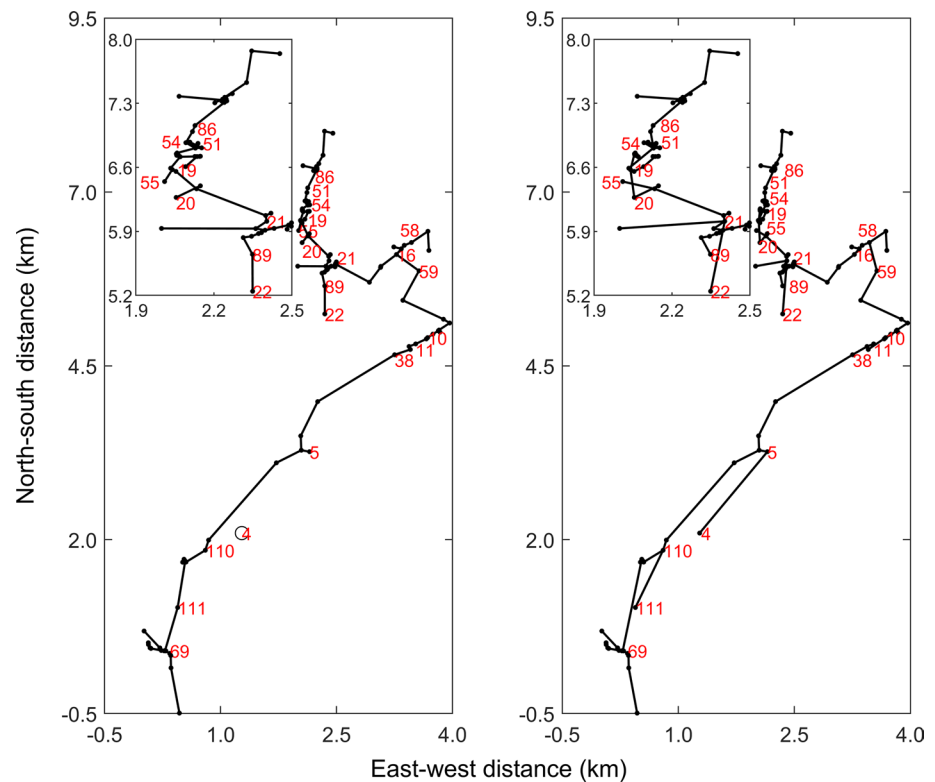


Figure 6. Comparison between the connection diagrams of the LMA sources from 19:54:43.9 UTC to 19:54:44.0 UTC using the pulse graph algorithm and proposed discharge channel extension algorithm. The upper left corner shows an enlarged view north of 10 km, the black line segments are connecting lines, serial numbers marked near the radiation sources of interest, and the hollow circles represent isolated sources. (a) Radiation source connection diagram using the pulse graph algorithm; (b) radiation connection diagram using the proposed discharge channel extension algorithm.

scale calculated by the new method of connecting radiation sources for discharge channel extension is 26.07 km.

More connection results based on more cases with different radiation source densities and comparisons with the current method are shown in Table S3. We use the current method and the new method to connect eight lightning flashes with different radiation source densities that occurred at different times. The connection comparison results of these eight lightning flashes are consistent with the above case analysis. The results also show that for the cases of low and medium densities (density less than 5), the conventional methods cannot connect the radiation sources as much as possible by using the fixed spatial threshold. But the proposed new method can connect more isolated radiation sources and make the length of the lightning channel closer to the real situation. The number of isolated points is reduced to 0 at most, and the maximum increase of channel length is 4.94 times of the original. For the cases of high density (density greater than 5), adjusting the spatial threshold has little effect on the isolated source, but the new method can trace the lightning discharge path as much as possible. Compared with the conventional method, we can connect the path of multiple discharges in the same channel as much as possible to obtain a more realistic lightning discharge scale and the maximum increase in channel length is 13.44 km.

5. Conclusions and Discussion

In this paper, a new method of connecting radiation sources is proposed based on refined 3D total flash LMA data. The influences of the spatial density of radiation sources and repeated discharge processes are considered by varying the spatial threshold and time threshold determined by the discharge process. The rationality of this method is verified by comparison with more refined INTF positioning data. The proposed

method reduces some incorrect connections under a low source density and obtains more real discharge pathways.

Different from the traditional method of a fixed spatial threshold, the proposed method selects the connection spatial threshold according to the density of radiation sources, thereby improving the stability of the channel scale under different detection capabilities (radiation source densities). The radiation source density is divided into high, medium, and low levels to determine the spatial threshold. The connection results show that the channel shape determined via the proposed method is more consistent with the real connection than the results of existing methods, and the channel scale is closer to the real scale by using the density-based threshold than by using the traditional fixed threshold.

Furthermore, the repeated discharge process is considered comprehensively. For the subdischarge processes occurring at different times in the same channel, multiple discharge channels are connected by limiting the duration of each discharge, thereby distinguishing the paths of multiple discharge events. The time threshold for this study is 10 ms. Compared with the traditional geometric channel scale, the discharge scale obtained in this study can track the lightning discharge process and better characterize the impact range of lightning discharge. The results also show that the discharge scale is much larger than the geometric scale: the geometric scales are 3.65, 18.33, and 9.58 km, while the corresponding discharge scales are 10.08, 28.12, and 20.12 km, respectively.

Of course, when the radiation source density is very small, it remains difficult to yield results close to the actual situation by using the proposed method of connecting radiation sources with a density-based spatial threshold. Therefore, the application of this method imposes a minimum requirement on the density of radiation sources: only flashes whose density exceeds a certain value can be connected and calculated reliably. In addition, the distinction of subdischarge events is very dependent on the specific lightning process. The goal of future research is to develop an adaptive method to obtain the time connection threshold based on low-frequency waveform and radiation source location data, which would make the calculation of the discharge scale more universal.

Data Availability Statement

This work complies with the AGU data policy; the data used in this paper can be obtained from the website: <https://doi.org/10.5281/zenodo.4524718>.

Acknowledgments

This research was funded by the National Key Research and Development Program of China (2017YFC1501501, 2019YFC1510103), the National Natural Science Foundation of China (41775009, 41775007, and 41905004), and the Basic Research Fund of Chinese Academy of Meteorological Sciences (Grant 2021Z011). We are very grateful to the lightning team of NMT for providing LMA and INTF data.

References

- Bruning, E. C., & MacGorman, D. R. (2013). Theory and observations of controls on lightning flash size spectra. *Journal of the Atmospheric Sciences*, 70(12), 4012–4029. <https://doi.org/10.1175/JAS-D-12-0289.1>
- Bruning, E. C., & Thomas, R. J. (2015). Lightning channel length and flash energy determined from moments of the flash area distribution. *Journal of Geophysical Research: Atmospheres*, 120, 8925–8940. <https://doi.org/10.1002/2015JD023766>
- Chronis, T., Lang, T., Koshak, W., Blakeslee, R., Christian, H., McCaul, E., & Bailey, J. (2015). Diurnal characteristics of lightning flashes detected over the São Paulo lightning mapping array. *Journal of Geophysical Research: Atmospheres*, 120, 11799–11808. <https://doi.org/10.1002/2015JD023960>
- Finney, D. L., Doherty, R. M., Wild, O., Young, P. J., & Butler, A. (2016). Response of lightning NO_x emissions and ozone production to climate change: Insights from the atmospheric chemistry and climate model intercomparison project. *Geophysical Research Letters*, 43, 5492–5500. <https://doi.org/10.1002/2016GL068825>
- Hager, W. W., Sonnenfeld, R. G., Aslan, B. C., Lu, G., Winn, W. P., & Boeck, W. L. (2007). Analysis of charge transport during lightning using balloon-borne electric field sensors and lightning mapping array. *Journal of Geophysical Research*, 112, D18204. <https://doi.org/10.1029/2006JD008187>
- Harary, F. (1969). *Graph theory*. Addison-Wesley Publishing Company.
- Jacobson, E. A., & Krider, E. P. (1976). Electrostatic field changes produced by Florida lightning. *Journal of the Atmospheric Sciences*, 33(1), 103–117. [https://doi.org/10.1175/1520-0469\(1976\)033<0103:EFCPBF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1976)033<0103:EFCPBF>2.0.CO;2)
- Kasemir, H. W. (1960). A contribution to the electrostatic theory of a lightning discharge. *Journal of Geophysical Research*, 65(12), 1873–1878. <https://doi.org/10.1029/JZ065i007p01873>
- Koshak, W., Peterson, H., Biazar, A., Khan, M., & Wang, L. (2014). The NASA lightning nitrogen oxides model (LNOM): Application to air quality modeling. *Atmospheric Research*, 135–136, 363–369. <https://doi.org/10.1016/j.atmosres.2012.12.015>
- Krehbiel, P. R., Brook, M., & McCrory, R. A. (1979). An analysis of the charge structure of lightning discharges to ground. *Journal of Geophysical Research*, 84(C5), 2432–2456. <https://doi.org/10.1029/JC084iC05p02432>
- Lapierre, J. L., Sonnenfeld, R. G., Edens, H. E., & Stock, M. (2014). On the relationship between continuing current and positive leader growth. *Journal of Geophysical Research: Atmospheres*, 119, 12479–12488. <https://doi.org/10.1002/2014JD022080>
- Peterson, M., Deierling, W., Liu, C., Mach, D., & Kalb, C. (2017). The properties of optical lightning flashes and the clouds they illuminate. *Journal of Geophysical Research: Atmospheres*, 122, 423–442. <https://doi.org/10.1002/2016JD025312>

- Peterson, M., Lang, T. J., Bruning, E. C., Albrecht, R., Blakeslee, R. J., Lyons, W. A., et al. (2020). New world meteorological organization certified megaflash lightning extremes for flash distance (709 km) and duration (16.73 s) recorded from space. *Geophysical Research Letters*, 47, e2020GL088888. <https://doi.org/10.1029/2020GL088888>
- Rison, W., Thomas, R. J., Krehbiel, P. R., Hamlin, T., & Harlin, J. (1999). A GPS-based three-dimensional lightning mapping system: Initial observations in central New Mexico. *Geophysical Research Letters*, 26(23), 3573–3576. <https://doi.org/10.1029/1999GL010856>
- Stock, M. G., Akita, M., Krehbiel, P. R., Rison, W., Edens, H. E., Kawasaki, Z., & Stanley, M. A. (2014). Continuous broadband digital interferometry of lightning using a generalized cross-correlation algorithm. *Journal of Geophysical Research: Atmospheres*, 119, 3134–3165. <https://doi.org/10.1002/2013JD020217>
- Thomas, R. J., Krehbiel, P. R., Rison, W., Hunyady, S. J., Winn, W. P., Hamlin, T., & Harlin, J. (2004). Accuracy of the lightning mapping array. *Journal of Geophysical Research*, 109, D14207. <https://doi.org/10.1029/2004JD004549>
- Wild, O. (2007). Modelling the global tropospheric ozone budget: Exploring the variability in current models. *Atmospheric Chemistry and Physics*, 7, 2643–2660. <https://doi.org/10.5194/acp-7-2643-2007>
- Yasuda, Y., Yokoyama, S., Minowa, M., & Satoh, T. (2012). Classification of lightning damage to wind turbine blades. *IEEE Transactions on Electrical and Electronic Engineering*, 7(6), 559–566. <https://doi.org/10.1002/tee.21773>
- Zhang, R., Zhang, G., Li, Y., Wang, Y., Wu, B., Yu, H., & Liu, Y. (2014). Estimate of NO_x production in the lightning channel based on three-dimensional lightning locating system. *Science China Earth Sciences*, 57(7), 1613–1625. <https://doi.org/10.1007/s11430-013-4812-1>
- Zhang, Z., Zheng, D., Zhang, Y., & Lu, G. (2017). Spatial-temporal characteristics of lightning flash size in a supercell storm. *Atmospheric Research*, 197, 201–210. <https://doi.org/10.1016/j.atmosres.2017.06.029>