ADVANCING EARTH AND SPACE SCIENCE

Earth and Space Science

METHOD

10.1029/2021EA002081

A Comparison of Processing Methods for the Oklahoma Lightning Mapping Array

V. C. Chmielewski[1,2], **J. Blair**[1,2,3], **D. Kennedy**[2], **D. MacGorman**[1,3], **and K. M. Calhoun**[2]

[1]Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma, Norman, OK, USA, [2]NOAA/OAR National Severe Storms Laboratory, Norman, OK, USA, [3]School of Meteorology, University of Oklahoma, Norman, OK, USA

Key Points:

- The Oklahoma Lightning Mapping Array can be processed as one unified network or the combination of two independent networks
- A comparison is made of these methods throughout the passage of a mesoscale convective system and isolated storm
- Processing as two networks saved computational time and increased the signal-to-noise ratio resulting more realistic mapping

Supporting Information:

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

Correspondence to:

V. C. Chmielewski,
vanna.chmielewski@noaa.gov

Citation:

Chmielewski, V. C., Blair, J., Kennedy, D., MacGorman, D., & Calhoun, K. M. (2022). A comparison of processing methods for the Oklahoma lightning mapping array. *Earth and Space Science*, *9*, e2021EA002081. https://doi.org/10.1029/2021EA002081

Received 13 OCT 2021
Accepted 14 FEB 2022

Author Contributions:

Conceptualization: V. C. Chmielewski, K. M. Calhoun
Data curation: D. Kennedy
Formal analysis: V. C. Chmielewski, J. Blair
Funding acquisition: D. MacGorman
Investigation: V. C. Chmielewski, D. Kennedy
Methodology: V. C. Chmielewski, J. Blair
Resources: D. MacGorman
Supervision: D. MacGorman

Abstract The Oklahoma Lightning Mapping Array (OKLMA) has consisted of 17 Very High Frequency (VHF) sensors in two separate clusters separated by roughly 150 km since 2012. This large footprint of sensors is expected to benefit the spatial coverage, sensitivity and accuracy of the network (e.g., Chmielewski & Bruning, 2016, https://doi.org/10.1002/2016jd025159; Koshak et al., 2018, https://doi.org/10.1175/jtech-d-17-0041.1). The unique configuration of sensors facilitates the evaluation of treating the two clusters as a single, unified network. Network operators may wish to do this when processing networks in close proximity or for temporary network extensions during field projects. The 0000–0600 UTC period on 16 June 2019 included millions of VHF sources and thousands of flashes from isolated storms and a mesoscale convective system. The 0300–0700 UTC period on 12 June 2019 contained a single storm of interest proceeding scattered activity. The VHF solutions and their grouped flashes from the unified OKLMA network are compared to those generated from the combination of the two clusters each operating as separate, independent networks. The unified network processing not only required the substantially longer processing time, but also resulted in lower sensitivity and more noise, suggesting that the theoretical model of high accuracy and sensitivity with a large LMA footprint may not apply to this configuration. Additionally, relatively few (<10%) of the VHF sources observed by a single cluster of sensors matched a VHF source observed by the opposite cluster, suggesting that the sources mapped by each cluster of sensors were most often unique.

Plain Language Summary The Oklahoma Lightning Mapping Array (OKLMA) can be treated as two separate networks of instruments or one large one. Theoretically, treating it as one large network, as has been done historically, should be better since the distances between stations and the large number of stations can better solve for where the lightning occurred in the cloud. We test both treating it as one large network and two separate networks during two periods of busy lightning activity. Treating the OKLMA as one large network took longer to compute and produced much worse results than treating it as the combination of two separate networks. The longer computation time was expected, but the worse results were not. In fact, it was relatively infrequent when the lightning signals recorded by each separate network were identical, which may be part of why the performance is worse because it is expected that all stations observe the same lightning signals using the historical treatment of a single large network.

1. Introduction

Lightning Mapping Arrays (LMA; Rison et al., 1999) have become the standard over the past 20 years for detection and location of very high frequency (VHF) emissions from lightning. The arrays typically consist of at least eight VHF receivers in a 50–100 km diameter area. The systems then use the time-of-arrival of the VHF radiation to geolocate lightning channels. Each array is capable of mapping VHF sources with an accuracy of tens of nanoseconds and meters over the network with decreasing accuracy and sensitivity with distance from the network (Koshak et al., 2004; Thomas et al., 2004). Each lightning flash produces groups of VHF sources which can be clustered into flashes automatically using time-space separation thresholds and a source-to-flash clustering algorithm (e.g., Fuchs et al., 2016; MacGorman et al., 2008; McCaul et al., 2009; Thomas et al., 2003). LMA measurements have been crucial in developing scientific understanding on many fronts - from developing our understanding of lightning physics (Krehbiel et al., 2000), to evaluating model simulations (e.g., Calhoun et al., 2014; Mansell et al., 2010), to validating other measurements such as the Geostationary Lightning Mapper (Goodman et al., 2013), to relating flash location and sizes to meteorological processes (Bruning & MacGorman, 2013), to describing storm evolution and intensity (e.g., Darden et al., 2010; Emersic et al., 2011; Schultz et al., 2011).

Visualization: V. C. Chmielewski, J. Blair
Writing – original draft: V. C. Chmielewski
Writing – review & editing: V. C. Chmielewski, K. M. Calhoun

The original Oklahoma Lightning Mapping Array (OKLMA; MacGorman et al., 2008) was installed as a single cluster of stations in central Oklahoma and has remained in operation since 2003. The observational characteristics of the single array were studied by Weiss et al. (2018) and the sensitivity and accuracy modeled by Chmielewski and Bruning (2016). In April 2012, prior to the Deep Convective Clouds and Chemistry project, an additional cluster of seven stations in southwest Oklahoma was added to the OKLMA (Barth et al., 2015). All stations operate at 60–66 MHz. The two clusters of stations are centered roughly 150 km apart and have been treated as one giant network, with stations from each cluster queried to triangulate VHF sources throughout the region. In the idealized Chmielewski and Bruning (2016) model, this configuration increased the sensitivity of the network to sources located between the two clusters, as the power received by an impulsive VHF emission is likely to be above the local noise floor of some stations in each cluster (Figure S1 in Supporting Information S1). This also improved the accuracy of solutions found due to the restriction of the hyperbolic errors (e.g., Thomas et al., 2004). The sensitivity tests of Koshak et al. (2018) also predicted improved accuracy and signal to noise ratios due to timing uncertainties with spatial separation between the sensors, supporting the combined processing of the clusters.

However, adding stations also increases processing time (Hamlin, 2004). More potential matches must be considered for each VHF signal at each station and the time window for VHF radiation to traverse the network and reach each station increases. This study investigates whether this additional processing time results in a higher signal-to-noise ratio of the VHF solutions (e.g., Koshak et al., 2018) and a more sensitive network (e.g., Chmielewski & Bruning, 2016) than when the separate OKLMA clusters are treated as independent networks and the resulting VHF source solutions combined afterward. Either should result in a single, enlarged LMA domain compared to considering each network cluster independently (Figure S1 in Supporting Information S1).

2. Methods

We selected the 0000 to 0600 UTC time period on 16 June 2019 and limited the analysis domain to −100.2 to −97.0°W (approximately 75 km west of the southwest cluster and east of the central cluster), 33.83 to 36.14°N (approximately 100 km south and north of the closest cluster) and 0–20 km altitude MSL bounding typical lightning heights (shown Movies S1 and S2 in Supporting Information S1). This period (henceforth Case 1) included an isolated storm north of the network initially and a quasi-linear convective system (QLCS) which entered the southwest side of the domain by 0100 UTC and propagated eastward. The leading edge of the QLCS was over the central cluster around 0400 UTC. The system continued to produce large stratiform flashes over the domain through the end of the period (Movie S1 and S2 in Supporting Information S1), so this case included a variety of lightning activity for comparison. Each station in the OKLMA was operating with an 80 μs timing window. There were 10 stations operating in the central Oklahoma cluster and seven in the southwest Oklahoma clusfter. A supporting case from 0300 to 0700 UTC on 12 June 2019 bounded by −99.7 to −96.5°W; 34.33 to 36.64°N; 0–20 km MSL is presented for comparison focused on a single thunderstorm which traveled southeast into the domain and toward the central cluster before thunderstorms initiated over the southwest cluster (shown Movies S3 and S4 in Supporting Information S1; Case 2).

The VHF sources were located using the standard LMA time of arrival method described in Thomas et al. (2004) using LMA Analysis v10.14.5 R. This processing was performed using each of two routines. The first included all of the stations in the OKLMA, as has been done in practice since installation (referred to henceforth as unified processing). This method allows for any combination of central or southwest stations to contribute to each VHF source solution, which can include information from as many stations as recorded the signal. The second processing method located VHF source solutions from each of the two clusters separately, treating each as a fully independent network. This results in two separate, standard arrays which have performance characteristics more widely studied (e.g., Chmielewski & Bruning, 2016; Koshak et al., 2004; Thomas et al., 2004; Weiss et al., 2018) with decreasing sensitivity and accuracy with range. This could result in duplicate retrievals of VHF sources from the two arrays. The VHF solutions found by either cluster were then combined naively into a single data set to cover the full domain. This is henceforth referred to as separated processing given the separation of the initial VHF source analysis by cluster. At least six stations were required to find the time and location of each VHF source for all methods, so there was less redundant information available for each cluster in the separated method.

10.1029/2021EA002081

Table 1
Processing Characteristics for Case 1, 16 June 2019 From 0000 UTC to 0600 UTC for Each of the Unified and Separated (With and Without Duplicate Sources) Processing Routines

	Unified	Separated	Duplicates removed
Processing Time (min)	267.9	9.49	9.46
Total Count	75,900,916	57,604,574	54,821,859
Domain Count	45,635,742	42,699,242	40,098,609
≤1 χ^2	32,567,998 (71.37)	34,849,925 (81.62)	32,637,902 (81.39)
≥7 stations	22,997,719 (50.39)	22,890,198 (53.61)	21,170,467 (52.80)
Count in flashes	29,146,372 (63.9)	32,949,941 (77.2)	30,742,783 (76.6)
Number of flashes	151,503	156,201	154,199
Average area (km^2)	74.84	78.43	79.22
Median area (km^2)	17.83	19.42	19.87

Note. Average processing time per 10 min file using the same serial computing environment, total count of VHF sources with $\chi^2 \leq 5$ and ≥6 stations contributing, count of VHF sources within the analysis domain (−100.2 to −97.0 °W, 33.83 to 36.14 °N, and 0–20 km altitude MSL), count (percentage of domain count) of those with $\chi^2 \leq 1$, count (percentage) with ≥7 stations contributing, count (percentage) within the specified domain which would also be grouped into at least a 10-point flash with $\chi^2 \leq 1$. Total number of flashes with ≥10 points initiating in the analysis domain and their average and median areas.

We expect flash grouping to help normalize the network sensitivity that affects VHF source counts in the analysis domain (e.g., Weiss et al., 2018) regardless of the processing technique used. VHF sources were grouped into flashes using the open source lmatools python package (Bruning, 2015; Fuchs et al., 2016). Flashes were grouped from the: unified method, separated method (all sources) and separated method with the duplicate sources removed. For a source to be grouped into a flash it must have a reduced $\chi^2 \leq 1$ and ≥ 6 stations contributing. The space and time thresholds for grouping were 3 km and 0.15 s, respectively. The maximum expected flash duration was 3 s. Each flash must include ≥10 VHF sources. Flash characteristics including flash extent and mean or average flash area (AFA) were calculated on a 1 km grid.

2.1. Duplicate Sources

A check was performed on the separated processing to restrict the artificial increase in sensitivity if a single VHF source were mapped by each cluster of stations and double-counted. A sensitivity test of a threshold values was performed. For a baseline, the distribution of VHF source solutions in the analysis domain with reduced $\chi^2 \leq 1$ and ≥7 stations contributing which were mapped by one cluster within a three-dimensional distance of 5 km and 17 μs of the other were examined (Figure S1 in Supporting Information S1). Most of these events occurred in the central portion of the domain with less than 1 km or 1 μs of each other, with the lowest differences in the area between the two clusters. The 75th percentile values of 3.27 μs and 1.46 km was selected to identify sources which were potentially duplicated by the separated processing. In a modification of the separated processing the solution with the higher reduced χ^2 was removed. This check is not expected to capture all potential duplicate solutions, especially of less well-mapped sources, but instead to provide a baseline of how the duplication of sources in the separated processing may impact solutions and downstream analysis. The thresholds used here are based on the measured distributions of differences, but larger thresholds could significantly reduce the number of non-unique sources (Figures S2c and S2d in Supporting Information S1).

3. Results

Case 1 included tens of millions of the VHF sources from each processing method (Table 1; averaging 1.6–2.1 million per 10 min). Treating the OKLMA as two separated clusters only required 3.5% of the processing time of the unified method (Table 1). However, the unified method included ~30% more VHF sources than the separated methods, many of which are likely unphysical, appearing to the user as noisy solutions, as will be demonstrated. There were relatively few sources flagged as duplicates - only 4.8% of sources in the specified domain described previously.

The analysis domain contained more similar VHF counts than overall, with the unified method only containing ~8% more VHF sources than the separated methods. The unified method lost 13.9 million (18%) of the VHF sources to the altitude restriction alone (compared to ~0.4 million or ~1% in either separated method), suggesting that many of the VHF sources may have been nonphysical or noisy solutions above typical lightning altitudes. Many of these noisy sources occurred when the minimum number of contributing stations were involved with some combination of stations from the central and southwestern clusters (Figure S2a in Supporting Information S1), while better solutions had a larger number of contributing stations and a spread of station involvement (Figure S2b in Supporting Information S1). Two practical methods commonly used by LMA users to provide more confidence in the VHF source solutions are limiting the maximum reduced χ^2 value or increasing the minimum number of sensors contributing to a single source, each of which are examined below.

Only 71.4% of sources from the unified method had a reduced $\chi^2 \leq 1$, compared to 81%–82% from the separated methods (Table 1), again supporting that more of the sources from the unified method were influenced by something such as transient or local VHF noise resulting in poor matches. Visually, there appears to be a large number of noisy solutions with the unified method (Figure 1) while the separated method with the same >6 station and $\chi^2 \leq 1$ requirements resulted in a cleaner mapping of the flash channels (Figure 2), suggesting that the unified solutions with $\chi^2 \leq 1$ still included nonphysical solutions.

Only 50.4% of the unified sources met the criteria of ≥7 contributing stations to increase confidence in the solutions, compared to 53.4%–53.6% for the separated methods (Table 1). The unified method only contained 2% or ~0.5 million more VHF sources than the separated method with duplicates removed. Noting that the southwest cluster only contains seven stations, this means that each sensor must contribute to a southwest solution and the power received must be higher than the thresholds for most of the central sensors, so relatively low sensitivity is expected for the separated method (Figures S3 and S4 in Supporting Information S1). However, the more stringent requirement is necessary for removing the apparent noise in the unified method that is seen in Figure 1 (for ≥6 stations).

Early in the time period a large amount of radial spread occurred using the unified method toward the central cluster in the solutions associated with the isolated storm even with the seven station minimum (Movie S1 in Supporting Information S1). There still is some spread in solutions with the separated method (Movie S2 in Supporting Information S1), but not as extensively as in the unified method and more closely aligned with expected performance (e.g., Thomas et al., 2004). Additionally, the altitudes are better constrained in the separated than the unified method. This suggests that not only can the separated method significantly save on computational costs it also results in more signal and less noise.

The performance of the methods in Case 2 were generally similar. Case 2 had 60%–65% fewer sources, but still required less processing time with the separated method (~4%) and produced more sources with the unified method (~46.8%; Table S1 in Supporting Information S1). Again, the unified method produced more noise in the solutions, as shown by the lower percentage of sources with reduced $\chi^2 \leq 1$ (Table S1 in Supporting Information S1) and by the radial spread of solutions toward the central cluster and, later in the case, above and away from each cluster (Movie S3 in Supporting Information S1), patterns less egregious with the separated method (Movie S4 in Supporting Information S1).

3.1. Analysis of Duplicate Sources

The southwestern cluster observed more sources to the southwest whereas the central cluster observed more sources near the central cluster (Figure S5 in Supporting Information S1), as expected. The VHF sources tagged as duplicates were usually separated in the vertical, as expected given typically large vertical errors (e.g., Chmielewski & Bruning, 2016; Koshak et al., 2004; Thomas et al., 2004). The combination of separate solutions may provide a user some redundancy and confidence that the observed flash channels are reasonable, and create more complete propagation paths than either cluster on its own. Meanwhile the propagation path is subjectively more difficult to differentiate from noise with the unified method (Figure 1).

The inclusion of duplicate sources may or may not be desirable to an end user. Redundancy in single points can increase confidence in observations of small events with only a few sources. In which case, retaining all points is worthwhile. However, flash counts may be affected - an event with the same 5 sources observed by each cluster will produce a 10-point flash not otherwise counted. This extreme is relatively unlikely as even very loose criteria

10.1029/2021EA002081

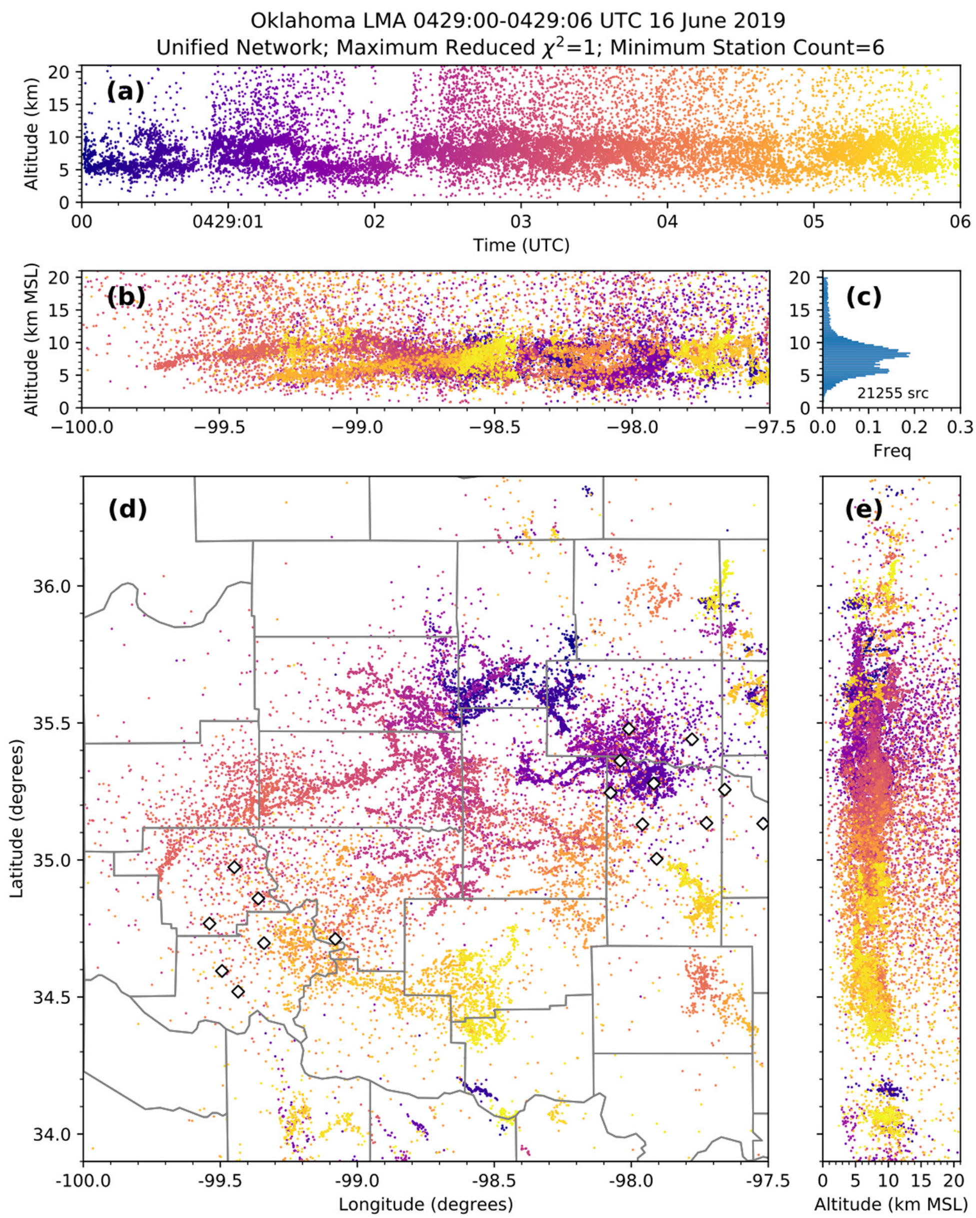

Figure 1. Very high frequency sources from the unified Oklahoma Lightning Mapping Array (OKLMA) method (colored by time) from 0429 to 0429:06 UTC on 16 June 2016 with reduced $\chi^2 \leq 1$ and ≥6 stations contributing to the source by (a) altitude and time, (b) altitude and longitude, (d) longitude and latitude and (e) latitude and altitude. (c) Shows the histogram of source by altitude. The locations of OKLMA sensors are shown by the white diamonds.

results ~10% duplicate sources. Conversely this analysis cannot determine whether sources tagged as duplicates are the same source or if they are independent or non-isotropic emissions in close proximity observed by different clusters. The peak powers recorded by each cluster may be generated by different emission sources on nearby channels or channel segments. If this effect does impact the recorded emissions, it may contribute to the noisiness of the unified method, as the assumption of a single source along a channel segment would be violated.

3.2. Flash-Level Results

Another way to restrict noise and account for network sensitivity (e.g., Weiss et al., 2018) is grouping VHF sources into flashes. Limiting VHF sources to those in flashes decreases the number of VHF sources (Table 1), but preserves the contributions from the southwest cluster in the separated routine unlike increasing the minimum

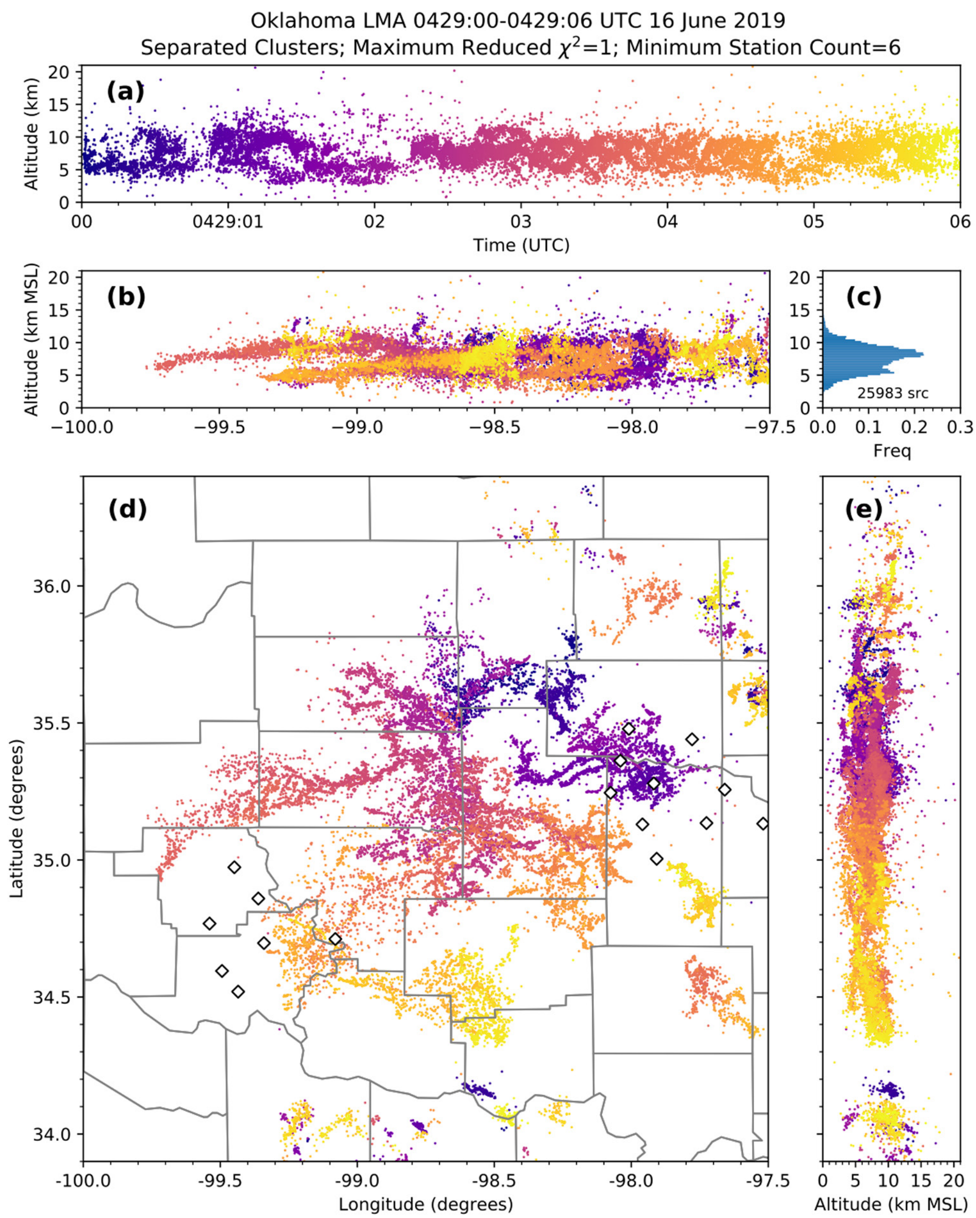

Figure 2. As in Figure 1 but with the separated method.

number of sensors. Approximately 89% of the unified sources with reduced $\chi^2 \leq 1$ were grouped into flashes, while 94% from the separated sources were, also suggesting a higher signal-to-noise ratio.

The differences propagate to flash-level analysis. Compared to the separated method, the unified method produced 3% fewer flashes while removing duplicate VHF sources reduced the flash count by 0.2%. The AFA decreased with the unified method by 4.6%, but increased negligibly (0.02%) without duplicates. The higher flash counts and sizes with the separated method suggests that while it produced more flashes, they were not solely the result of more small events reaching the 10-point criteria as flashes were also more spatially extensive.

Spatially, the 1-min AFA (Figure 3) differed but not at statistically significant levels given the large spread in flash area. The only significant differences by t-test at the 90% level were at the periphery of lightning-producing area (not shown). Differences in the composite AFA between the unified and separated (Figures 3a and 3b) methods were largest between the clusters, where the unified method produced smaller AFA. It is suspected that the less coherent channels produced by the unified method (e.g., Figure 1) restricted the extent of grouped flashes.

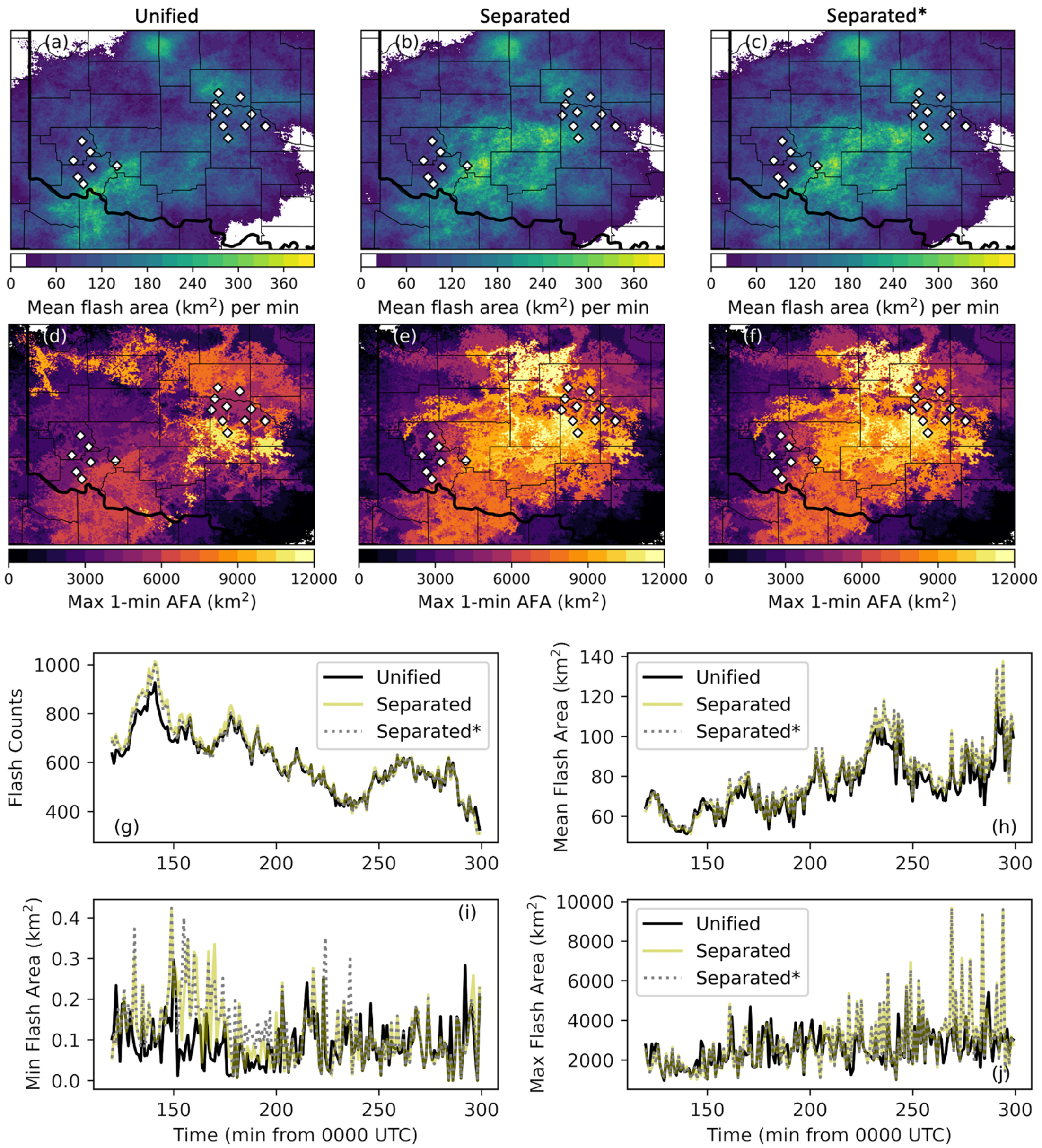

Figure 3. Mean (a)–(c) and maximum (d)–(f) of 1-min average flash area on a 1 km grid for the unified (a),(d), separated (b),(e) and separated without duplicates (c,f; marked as "separated*") methods. Time series of per-minute flash count (g), mean (h) minimum (i) and maximum (j) flash areas for each of the unified and separated methods in which the separated without duplicates is marked with an "*".

The maximum 1-min AFA produced visually striking variations between the unified and separated methods. The unified method (Figure 3d) had relatively large maxima south of the central cluster and along the northern domain. The separated method maxima (Figures 3e and 3f) were more distributed through the central domain, but the values along the northwestern edge of the domain were smaller. Noting the radial spread of sources from the unified method, the grouping may have included some noisy sources, as illustrated by the grid points containing large flash areas west of the Oklahoma-Texas border (Figure 3d).

Temporally during the passage of the QLCS (0200–0500 UTC, Figures 3g–3j), the correlations of the 1 min flash rates and AFA were extremely high between the separated method with and without duplicates ($R^2 = 0.998$ and 0.999, respectively) whereas lower correlations occurred between the unified and separated method ($R^2 = 0.835\%$ and 0.780%, respectively), suggesting some differences in trends between the two methods.

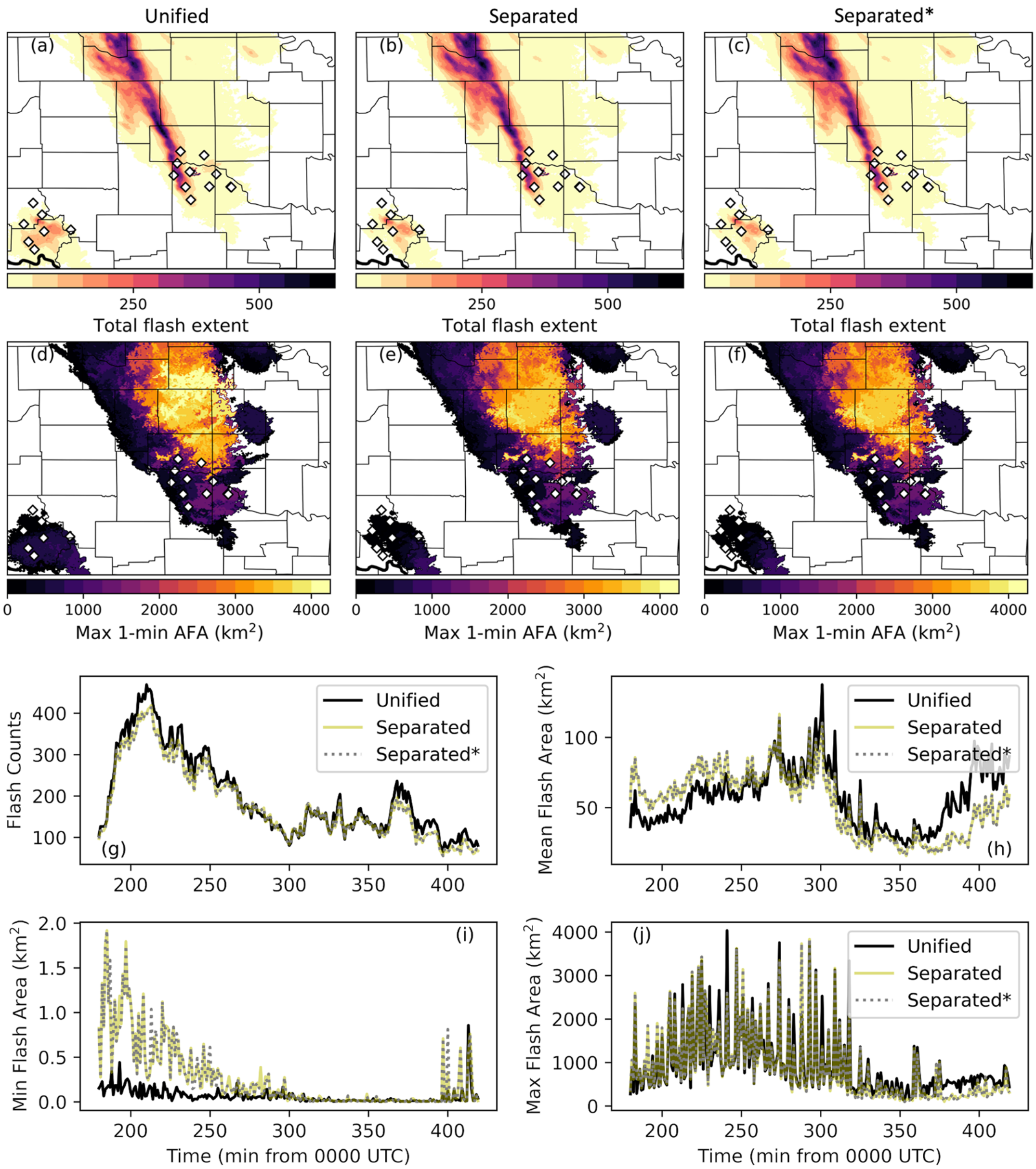

Figure 4. As in Figure 3 but for Case 2 with the total flash extent density throughout the period (a–c).

In Case 2, there were significant differences in 1-min flash extent densities over the clusters, where the radial spreading late in the period was evident in the unified method which were grouped into flashes (Figures 4a–4f). Early in the time period there were more and smaller flashes resolved with the unified method (Figures 4g–4i), which may suggest higher sensitivity on the periphery of the domain than the separated method, while the opposite occurs later in the period. Even so, correlations were high amongst the methods, with $R^2 = 0.939$ for the AFA and 0.984 for flash rates between the unified and separated methods and over 0.99 between the separated methods.

4. Conclusions

Contrary to the expectations, we found very little reason to treat both clusters of the OKLMA as one unified network in our test case with hours of widespread lightning activity. Treating each cluster separately then combining all of the VHF sources before any flash level analysis produced significantly better results. There was a 96.5%

savings in computational time by limiting the number of potential matches interrogated (e.g., Hamlin, 2004). There were more VHF sources and a higher percentage of the total VHF solutions which met basic quality control criteria when the clusters were treated independently than unified into one network, by either reduced chi^2 (81.6% vs. 71.4%) or flash grouping (77% vs. 64%) criteria, indicating both more sensitivity and a higher signal-to-noise ratio when the clusters were treated separately.

There were very few sources which could be identified as well-mapped duplicates from each cluster when the clusters were treated independently, but loosening that criteria could increase that to roughly 10% of all sources. However, we are unconvinced that this is desirable compared to the redundancy on observations and confidence in the observed signal. Such redundancy has little impact of the flash-level products with the exception of small events reaching minimum flash level criteria. Capturing these small events are useful for storm monitoring (Kuhlman et al., 2006), so extending the area of sensitivity to small events may be useful.

In Case 2 with a single storm moving into the domain the methods produced highly correlated trends in flash rates and AFA, suggesting little potential impact on storm-tracked attributes such as lightning jumps (e.g., Schultz et al., 2011). We also note that the test cases include optimal operating conditions of the OKLMA. The performance of either cluster will be negatively impacted if sensors are offline or experiencing hardware malfunctions. The widespread lightning activity in Case 1 and later in Case 2 may have added to the noisy solutions in the unified method. However, early in both cases when storms were less pervasive, the noisy solutions were still reduced by treating the clusters separately (Movies S1–S4 in Supporting Information S1).

The source of the noise in the full unified OKLMA network could be due to several things. The most obvious is local noise at the sites being incorporated into the solutions. With 17 stations, there are more possibilities for correlations of lightning VHF with local noise creating false solutions. This noise can be reduced by increasing the minimum number of stations for a solution. For example, the OKLMA often records VHF emissions from wind turbines when storms are overhead (as can be seen later in Case 2), so there are many opportunities for local noise to be enhanced during active thunderstorms. There could also be small errors in the assigned station locations or signal delays at each station compounding the errors in the solutions. Station locations were rigorously checked against GPS records and satellite data prior to this analysis, and signal delays methodically tested to account for any potential hardware degradation or replacements made since installation, so these impacts should be limited. The effect of refraction should likewise be limited in the domain, with the estimate at 200 km to be approximately 30 m (Thomas et al., 2004).

Whatever the reason is, the benefit of extending the baseline to solve the time of arrival equations is more complicated than expected from ideal theoretical conditions (e.g., Chmielewski & Bruning, 2016; Koshak et al., 2018). Perhaps if the VHF sources were more isolated, local noise sources different, or station configuration adjusted the results may be different. In the future, other sources of errors such as transient VHF noise or assumptions made for solutions such as the point source nature of VHF emissions may need to be investigated to improve the theoretical modeling of the unified processing. In practice, this analysis supports treating the two nearby clusters of the OKLMA as independent arrays to solve the time of arrival equations before combining the results prior to flash grouping. This maintains the benefit of the larger, contiguous observation domain produced by the unified routine to resolve larger flash propagation paths and simplify analysis in any one location, but reduces the processing time and increases the signal-to-noise ratio. This technique may be useful to future research with neighboring LMA clusters or in the design of field campaigns to optimally expand regional coverage of an existing LMA with temporary stations deployed as a neighboring, independent array.

Data Availability Statement

Processed Lightning Mapping Arrays data from each processing routine can be found here (https://doi.org/10.5281/zenodo.5156179).

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Acknowledgments

Funding was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement NA16OAR4320115, U.S. Department of Commerce. We thank Harald Edens, Daniel Rodeheffer, Ron Thomas, Paul Krehbiel and Mason Quick for conversations and advice surrounding reducing noise in the Oklahoma Lightning Mapping Array (OKLMA) and comparisons to other networks. We thank Timothy Lang and Eric Bruning for the development of Python-based Lightning Mapping Arrays plotting scripts. We also appreciate the efforts of Alex Fierro, Matthew Miller, Emily Luschen and Mackenzie Boylson for their efforts in maintaining the OKLMA archive.

References

Barth, M. C., Cantrell, C. A., Brune, W. H., Rutledge, S. A., Crawford, J. H., Huntrieser, H., et al. (2015). The deep convective clouds and chemistry (DC3) field campaign. *Bulletin of the American Meteorological Society*, *96*(8), 1281–1309. https://doi.org/10.1175/BAMS-D-13-00290.1

Bruning, E. C. (2015). (deeplycloudy) lmatools: lmatools-v0.5z-stable (version v0.5z). Zenodo. https://doi.org/10.5281/zenodo.32510

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

Calhoun, K. M., Mansell, E. R., MacGorman, D. R., & Dowell, D. C. (2014). Numerical simulations of lightning and storm charge of the 29-30 may 2004 geary, Oklahoma, supercell thunderstorm using enkf mobile radar data assimilation. *Monthly Weather Review*, *142*(11), 3977–3997. https://doi.org/10.1175/MWR-D-13-00403.1

Chmielewski, V. C., & Bruning, E. C. (2016). Lightning Mapping Array flash detection performance with variable receiver thresholds. *Journal of Geophysical Research: Atmospheres*, *121*(14), 8600–8614. https://doi.org/10.1002/2016JD025159

Darden, C. B., Nadler, D. J., Carcione, B. C., Blakeslee, R. J., Stano, G. T., & Buechler, D. E. (2010). Utilizing total lightning information to diagnose convective trends. *Bulletin of the American Meteorological Society*, *91*, 167–176. https://doi.org/10.1175/2009BAMS2808.1

Emersic, C., Heinselman, P. L., MacGorman, D. R., & Bruning, E. C. (2011). Lightning activity in a hail-producing storm observed with phased-array radar. *Monthly Weather Review*, *139*(6), 1809–1825. https://doi.org/10.1175/2010MWR3574.1

Fuchs, B. R., Bruning, E. C., Rutledge, S. A., Carey, L. D., Krehbiel, P. R., & Rison, W. (2016). Climatological analyses of LMA data with an open-source lightning flash-clustering algorithm. *Journal of Geophysical Research: Atmospheres*, *121*, 8625–8648. https://doi.org/10.1002/2015JD024663

Goodman, S. J., Blakeslee, R. J., Koshak, W. J., Mach, D., Bailey, J., Buechler, D., et al. (2013). The GOES-R geostationary lightning mapper (GLM). *Atmospheric Research*, *125–126*, 34–49. https://doi.org/10.1016/j.atmosres.2013.01.006

Hamlin, T. D. (2004). The New Mexico Tech lightning mapping array (Unpublished doctoral dissertation)New Mexico Instatute of Mining and Technology.

Koshak, W. J., Mach, D. M., & Bitzer, P. M. (2018). Mitigating VHF lightning source retrieval errors. *Journal of Atmospheric and Oceanic Technology*, *35*(5), 1033–1052. https://doi.org/10.1175/JTECH-D-17-0041.1

Koshak, W. J., Solakiewicz, R. J., Blakeslee, R. J., Goodman, S. J., Christian, H. J., Hall, J. M., et al. (2004). North Alabama lightning mapping array (LMA): VHF source retrieval algorithm and error analyses. *Journal of Atmospheric and Oceanic Technology*, *21*, 5432–5558. https://doi.org/10.1175/1520-0426(2004)021<0543:nalmal>2.0.co;2

Krehbiel, P. R., Thomas, R. J., Rison, W., Hamlin, T., Harlin, J., & Davis, M. (2000). GPS-based mapping system reveals lightning inside storms. *Eos, Transactions American Geophysical Union*, *81*(3), 21–32. https://doi.org/10.1029/00EO00014

Kuhlman, K. M., Ziegler, C. L., Mansell, E. R., MacGorman, D. R., & Straka, J. M. (2006). Numerically simulated electrification and lightning of the 29 June 2000 STEPS supercell storm. *Monthly Weather Review*, *134*, 2734–2757. https://doi.org/10.1175/mwr3217.1

MacGorman, D. R., David Rust, W., Schuur, T. J., Biggerstaff, M. I., Straka, J. M., Ziegler, C. L., et al. (2008). TELEX the thunderstorm electrification and lightning experiment. *Bulletin of the American Meteorological Society*, *89*(7), 997–1014. https://doi.org/10.1175/2007BAMS2352.1

Mansell, E. R., Ziegler, C. L., & Bruning, E. C. (2010). Simulated electrification of a small thunderstorm with two-moment bulk microphysics. *Journal of the Atmospheric Sciences*, *67*(1), 171–194. https://doi.org/10.1175/2009JAS2965.1

McCaul, E. W., Jr, Goodman, S. J., LaCasse, K. M., & Cecil, D. J. (2009). Forecasting lightning threat using cloud-resolving model simulations. *Weather and Forecasting*, *24*(3), 709–729. https://doi.org/10.1175/2008WAF2222152.1

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

Schultz, C. J., Petersen, W. A., & Carey, L. D. (2011). Lightning and severe weather: A comparison between total and cloud-to-ground lightning trends. *Weather and Forecasting*, *26*(5), 744–755. https://doi.org/10.1175/WAF-D-10-05026.1

Thomas, R. J., Krehbiel, P., Rison, W., Harlin, J., Hamlin, T., & Campbell, N. (2003). The LMA flash algorithm. Int. Comm. on Atmos. Electr. Versailles, France.

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

Weiss, S. A., MacGorman, D. R., Bruning, E. C., & Chmielewski, V. C. (2018). Two methods for correcting range-dependent limitations of lightning mapping arrays. *Journal of Atmospheric and Oceanic Technology*, *35*, 1273–1282. https://doi.org/10.1175/JTECH-D-17-0213.1