Mitigating VHF Lightning Source Retrieval Errors

WILLIAM J. KOSHAK
NASA Marshall Space Flight Center, Huntsville, Alabama

DOUGLAS M. MACH
Universities Space Research Association, Huntsville, Alabama

PHILLIP M. BITZER
University of Alabama in Huntsville, Huntsville, Alabama

(Manuscript received 7 March 2017, in final form 22 December 2017)

ABSTRACT

The problem of inferring the location and time of occurrence of a very high frequency (VHF) lightning source emission from Lightning Mapping Array (LMA) network time-of-arrival (TOA) measurements is closely examined in order to clarify the cause of retrieval errors and to determine how best to mitigate these errors. With regard to this inverse problem, the previous literature lacks a comprehensive discussion of the associated forward problem. Hence, the forward problem is analyzed in this study to better clarify why retrieval errors increase with increasing source horizontal range and/or decreasing source altitude. Further insight is obtained by performing carefully designed Monte Carlo inversion simulations that provide specific retrieval error plots, which in turn lead to clear recommendations for mitigating retrieval errors. Based on all of the numerical results, the following strategies are recommended for mitigating retrieval errors (when possible, and without obstructing the line of sight): expand the horizontal extent of the LMA network, maximize the vertical sensor baseline by using mountainous terrain if available, and improve TOA measurement timing accuracy. Adding sensors to the network is relatively ineffective, unless of course the addition of sensors expands the horizontal extent and/or vertical baseline of the network. It is also shown how the standard retrieval method can be generalized by considering, in addition to the regular (unpolarized) point VHF source, the polarized transient very low frequency/low frequency (VLF/LF) electric point dipole source. Multiple observations (i.e., VHF arrival time and power, and VLF/LF arrival time and electric field amplitude) are simultaneously implemented into the new generalized mathematical framework, and the potential benefits are indicated.

1. Introduction

Mathematical retrieval of a lightning very high frequency (VHF) radio source, both its location and time of occurrence, from ground-based multistation time-of-arrival (TOA) measurements (or differences in such measurements) have been widely explored (Proctor 1971; Rustan et al. 1980; Proctor 1981; Proctor et al. 1988; Koshak and Solakiewicz 1996; Thomas et al. 2001; Koshak et al. 2004; Thomas et al. 2004; Chmielewski and Bruning 2016). The retrieval techniques have been applied, for example, to TOA data derived from the Lightning Detection and Ranging (LDAR) and Lightning Mapping Array (LMA) networks; see Rustan et al. (1980) and Thomas et al. (2004), and references therein, for more details on each of these networks, respectively. Similar mathematical retrieval techniques have been applied to networks that operate in a frequency band other than the VHF (Thomson et al. 1994; Hager and Wang 1995; Betz et al. 2004; Bitzer et al. 2013). An overview of various retrieval methodologies applied to a variety of lightning detection network types is provided in Cummins and Murphy (2009).

In this paper the focus is to reexamine and more clearly identify the causes of excessive retrieval errors, and to demonstrate how some of these errors can be
mitigated. The primary focus will be given to the altitude retrieval, since it is the most difficult source parameter to retrieve when the source horizontal range (i.e., the horizontal distance $D$, shown in Fig. 1) is very large and/or when the source altitude is very low (Koshak and Solakiewicz 1996; Koshak et al. 2004; Thomas et al. 2004; Chmielewski and Bruning 2016). Although the issues of large retrieval errors are difficult to deal with, it is hoped that this paper will motivate the use, or at least consideration, of the more aggressive mitigation steps suggested here to overcome these retrieval errors. While there have been simple geometric analyses and prior modeling work (see references above) that provide useful guidance on the design of LMAs to minimize error, there have been no rigorous systematic studies of retrieval errors such as provided in this paper; that is, we begin with a rigorous and modern statement of the inverse problem, provide a formal forward problem analysis that appropriately generalizes results to handle arbitrary signal-to-noise conditions, and then perform systematic error simulations that do not conflate error sources.

Owing to the widespread use of LMA networks, the primary focus is on retrievals that are based on LMA data. LMA hardware and data processing technologies will likely continue to improve in the future and lead to ever more accurate TOA measurements. There also exists flexibility in how LMA networks can be deployed (in terms of the number of sensors, network geometry, network extent, and to some degree sensor altitude). In addition, two close LMA networks could potentially be merged into a single larger network. With all of these possible adjustments, the findings in this work provide both insight and recommendations.

The writing is organized as follows. Section 2 gives an overview of the basic inverse problem and provides the standard retrieval method employed for inferring the four unknown parameters (3D location and time of occurrence) of a VHF point source. However, as alluded to above, the solution to the inverse problem is flawed (i.e., results in large retrieval errors) under certain conditions. To better understand these difficulties, section 3 examines the forward problem in order to quantify how sensitive the measurements are to small changes in the source parameters. With some context provided in section 4 on the maximum line of sight (LOS) possible between sensor and source, results for a baseline simulation are provided in section 5. Sections 6–9 illustrate how alterations in key LMA network characteristics (i.e., number of sensors, horizontal network extent, sensor altitude, and magnitude of measurement error) can result in beneficial reductions in retrieval errors relative to the baseline run. In addition, the appendix introduces a new generalized retrieval method that involves retrieving both the standard (unpolarized) VHF point source and the associated collocated transient polarized very low frequency/low frequency (VLF/LF) electric point dipole source of the lightning channel. Section 10 provides a summary.

2. The inverse problem

This section provides a brief overview of the standard method for retrieving the unknown location and time of occurrence ($X, Y, Z, T$) of a VHF lightning point source emission. Additional details can be found in Koshak et al. (2004) and Thomas et al. (2004).

The VHF radio wave emission from the point source to the observation point is assumed to follow a straightline path model; that is, the model arrival time $t_i$ of the wave at the $i$th LMA sensor is given by $t_i = t + R/c$. Here, $R_i = R(x, y, z) = [(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2]^{1/2}$ is the distance between the model point source at location $(x, y, z)$ and the $i$th LMA sensor observation point at $(x_i, y_i, z_i)$. $t$ is the time of occurrence of the model point source, and $c$ is the speed of light. The LMA network is assumed to have $i = 1, \ldots, m$ sensors.

The $i$th LMA sensor observes the arrival time $t_i = t + \epsilon_i$, where $\epsilon_i$ is the measurement error; $\epsilon_i$ is typically assumed to be a normally distributed random variable with a mean of zero and a standard deviation $\sigma_i$. Moreover, the standard deviation is usually assumed to be the same at each site; for example, $\sigma_i = \sigma \sim 50$ ns for the New Mexico Institute of Mining and Technology LMA (Thomas et al. 2004) and about 23 ns for the West Texas LMA (Chmielewski and Bruning 2016). Upgrades to the various LMA networks occur on a continual basis, so these reported values are certainly not fixed. Finally, note that $\epsilon_i$ technically also includes modeling error, since the transit equation $t_i = t + R/c$ does not hold exactly (i.e., propagation paths are not exactly a straight line, propagation speed is slightly less than $c$, and the VHF source is not exactly a point source), but these effects are so small that they have been neglected in all LMA studies to date.

The most widely used retrieval algorithm involves searching for a model VHF point source $(x, y, z, t)$ that minimizes a reduced chi-squared goodness-of-fit function, given by

$$
\chi^2 = \frac{1}{m-4} \sum_{i=1}^{m} \{ \frac{t_i - [t + R(x, y, z)/c]}{\sigma_i} \}^2 .
$$

(1)

Here, the number of degrees of freedom is $\nu = m - 4$. The search is initialized by a linear retrieval method introduced in Koshak and Solakiewicz (1996) and
FIG. 1. (a) Basic geometry between site and source. (b) Geometry associated with a vertical displacement of source. (c) Geometry for a horizontal displacement of source.
further analyzed in Koshak et al. (2004). From this initialization the search proceeds using the iterative non-linear Levenberg–Marquardt algorithm (Marquardt 1963) in an attempt to find the minimum of $\chi^2$. The result of the Levenberg–Marquardt iterations is to hone in on a final solution given by $(x_i, y_i, z_i, t_i)$. It is possible that this solution is not the global minimum of $\chi^2$ as intended if the iterations terminate in a local relative minimum of the $\chi^2$ hypersurface. Generally speaking, external constraints (i.e., additional independent measurements and/or mathematical constraints that exert a valid physical constraint on the solution process) can be added to (1) in an attempt to improve the retrieval and/or extract more information; for example, see the appendix for a detailed example of how (1) can be generalized.

Even before a retrieval can be considered, it is imperative that each arrival time (from the set of $m$ arrival times to be inverted) is matched with the same VHF source. That is, within the 80–100-μs measurement window within which the LMA usually operates, the peak radio wave amplitude is used to identify the arrival time at that site. However, the VHF source that produces a peak at one site might not be the same source that produces a peak at another site. So, one has to successfully match arrival times from the various sensors that are truly produced from the same source. This matching process involves examining many combinations of arrival times and associated solutions, which is only practical using the numerically quick Koshak–Solakiewicz linear retrieval method mentioned above (see Thomas et al. 2004 for additional details).

In summary, the mathematical inversion problem is to determine, as best as possible, the four parameters $(X, Y, Z, T)$ of the unknown VHF source from the $m$ arrival time observations $(\tau_1, \ldots, \tau_m)$. The difference between the retrieved solution $(x_i, y_i, z_i, t_i)$ and the unknown source $(X, Y, Z, T)$ represents the retrieval errors. Many factors (i.e., the measurement errors $\epsilon_i$, the number of sensors $m$, the geometry of the network, the horizontal extent of the network, the altitude of sensors, the location of the source relative to the network, the specific retrieval algorithm employed, and the success in correctly matching arrival times) are important in assessing and understanding the retrieval errors.

3. Information content

Fundamental to understanding an inverse problem is to examine in detail the associated forward problem. By performing the forward problem over and over, one obtains an idea of how sensitive the measurements are to changes in the source parameters. Specifically, the forward problem here involves determining how changes in any of the four source parameters $(x, y, z, t)$ affect each arrival time $t_i$. Measurement systems that are largely insensitive to fluctuations in the source provide little information about source attributes. Such low information content systems correspond to ill-posed inverse problems and result in large retrieval errors, unless external constraints can be applied to better resolve the unknown source. External constraints typically consist of mathematical constraints that impose additional physical restrictions on the type of retrieved solution, but they can also be in the form of additional independent measurements that are distinct from the initial set of observations employed.

As a simple example of a forward problem analysis, consider the effect of increasing the source altitude $z$ by $\delta z = 0.1$ km. This results in a change $\delta t_i$ at the $i$th LMA sensor. The sensor needs to have enough sensitivity to detect the change $\delta t_i$. However, even if it does, the sensor measurement error $\epsilon_i$ might be large enough to obscure this change, where $\delta t_i = \delta t_i + \epsilon_i$ is the sensor-detected change. Hence, the sensitivity of a measurement and the measurement error together set fundamental limits on what level of source detail can be accurately retrieved. Poor sensor sensitivity and/or large measurement error result in solution ambiguity and unacceptably large retrieval errors.

To better quantify these complications, consider a source that activates at time $t$ at an altitude $z_1$ and that has a horizontal distance of $D_i$ from the $i$th sensor; see the geometry in Figs. 1a and 1b. The arrival time resulting from this source is $t + R(x, y, z_1)/c$. By increasing the source altitude to $z_2$, the arrival time becomes $t + R(x, y, z_2)/c$. The difference in these arrival times is $\delta t_i = [R(x, y, z_2) - R(x, y, z_1)]/c$. Expressing the slant distances in terms of horizontal and vertical distances allows one to define a signal-to-noise ratio (SNR) for the vertical source displacement as $\eta_i$, given by

$$\eta_i = \frac{\delta t_i}{\sigma} = \sqrt{D_i^2 + (z_i - z_2)^2} - \sqrt{D_i^2 + (z_i - z_1)^2}. \quad (2)$$

A plot of the SNR $\eta_i$ is given in Fig. 2a. It clarifies why it is difficult to accurately retrieve $z$ of a distant and/or low-altitude source. In effect, the sensor is less sensitive to vertical source displacements when the source altitude is low and/or when the source horizontal range is large. Three cases were considered. One source has its altitude moved from 2 to 3 km (red curve), another from 6 to 7 km (blue curve), and the third from 10 to 11 km (black curve). In all three cases, the SNR decreases with horizontal range, and the lower-altitude sources are associated with smaller SNR than higher-altitude sources. As expected, it is also found that the SNR decreases when the magnitude of the vertical displacement is...
decreased [i.e., in the limit as \( z_1 \) approaches \( z_2 \), the numerator in (2) approaches zero].

Another way to understand the drop in SNR with range is as follows. The elevation angle \( \alpha_1 \) of the source at \( z_1 \) is related to the geometry in Fig. 1b by

\[
R_i(x, y, z_1) = \frac{D_i}{\cos \alpha_1}.
\]

Similarly,

\[
R_i(x, y, z_2) = \frac{D_i}{\cos \alpha_2}.
\]

Hence, as \( D_i \) gets large, the cosine of each elevation angle approaches unity

\[
\Rightarrow R_i(x, y, z_1) \sim R_i(x, y, z_2) \sim D_i \Rightarrow \delta t_i \sim 0 \Rightarrow \text{SNR} \sim 0.
\]

One can also examine sensitivity to horizontal displacements in the source. Applying the law of cosines to the base triangle in Fig. 1c and applying the trigonometric identity for the cosine of the difference between angles, the \( D_{iu} \) between the sensor and the displaced source can be determined. Then, applying the Pythagorean theorem allows one to determine the slant distance \( R_{iu} \). The SNR \( \eta_2 \) for a horizontal displacement a distance \( \rho \) in the direction \( \phi \) is then

\[
\eta_2 = \frac{\Delta t_i}{\sigma} = \frac{\sqrt{D_i^2 + 2D_i \rho \cos \phi + \rho^2 + (z_i - z)^2} - \sqrt{D_i^2 + (z_i - z)^2}}{cc\sigma}.
\]

The plot in Fig. 2b illustrates the sensor sensitivity to different directions of horizontal displacement of the source. As expected, forward \((\phi = 0)\) and backward \((\phi = \pi)\) horizontal displacements produce the largest change in arrival time at a sensor. The smallest change occurs for an orthogonal displacement \((\phi = \pi/2)\); that is, since only one sensor is considered, a single sensor is virtually “blind” to lateral displacements of the source. As is evident from (3), all changes increase with the magnitude of the displacement \( \rho \). Unlike for vertical displacements, there is no decay in the SNR with horizontal range. This is easy to see for the case \( \phi = 0 \), where from Fig. 1c one has \( R_i = D_i / \cos \alpha \) and \( R_{iu} = (D_i + \rho) / \cos \alpha_i \). So, as \( D_i \) increases, the cosine of each elevation angle approaches unity, and the slant distances become \( R_i \sim D_i, R_{iu} \sim D_i + \rho \); that is, they do not approach the same value as in the vertical displacement discussed above but rather differ by an amount \( \rho \).

Note that the plots in Fig. 2 (as well as in Figs. 3 and 4) do not account for the eventual loss of signal at the sensor. That is, the transient electromagnetic wave emitted by the source attenuates with range from the source. Hence, as \( D_i \) increases eventually there will not be a detectable signal at the sensor (and hence no computable arrival time). For example, with respect to Fig. 2b, a particularly weak source and/or low-gain sensor might result in the sensor not registering any signal \((\eta_2 = 0)\) at a distance of, say, 150 km.
FIG. 3. SNR values $\eta_x$, $\eta_y$, and $\eta_z$ associated with the difference in TOA obtained from two vertically separated sensors (located at the origin) when the source is displaced by 1 km in the (top) $x$, (middle) $y$, and (bottom) $z$ directions. Sensor vertical separation is (left) 0.5 and (right) 4 km. The SNR values are so small that they had to be amplified by the multiplication factors (1000 or 5) shown.
FIG. 4. As in Fig. 3, except that now the sensors are horizontally separated in the $x$ direction by (left) 50 and (right) 100 km. No amplification of the SNR values is necessary [i.e., actual SNR values are shown: (top) $\eta_x$, (middle) $\eta_y$, and (bottom) $\eta_z$].
A change in an arrival time at a particular sensor can also occur by changing the source time of occurrence $t$, where $(x, y, z)$ are held fixed. This implies an SNR $\eta_t$, given by

$$\eta_t = \frac{|\delta t|}{\sigma} = \frac{|\delta t|}{\sigma}.$$  \hspace{1cm} (4)

This effect is straightforward and simply says that changes in the source time show up directly as changes in sensor-detected arrival time.

In general, and as shown in (2)–(4), a vertical displacement, a horizontal displacement, and a direct displacement, a horizontal displacement, and a direct difference is to the vertical displacement. Hence, whereas $\eta_1$ tells one how sensitive an arrival time value is to a vertical displacement in the source, $\eta_z$ tells one how sensitive an arrival time difference is to the vertical displacement.

To avoid mixing horizontal and vertical effects, the SNR values for two network geometries are examined: a network consisting of two vertically separated sensors (Fig. 3) and a network consisting of two horizontally separated sensors (Fig. 4). A value of 1 km was used for $\delta x$, $\delta y$, and $\delta z$, and the source $z$ was taken as 6 km. In Fig. 3, the vertical network has a sensor at $(0, 0, 0)$ km and one directly above at $(0, 0, d_z)$ km; the values $d_v = 0.5$ km (left column) and $d_v = 4$ km (right column) were used. The (amplified) SNR values are plotted as a function of horizontal source location $(x, y)$. Specifically, the (amplified) values of $\eta_1$ (top row), $\eta_y$ (middle row), $\eta_z$ (bottom row) are provided. The vertical networks have poor sensitivity, so the SNR values were amplified in order to illustrate the pattern of sensitivity using the same color scale (i.e., otherwise, the plots would be mostly all red except in a region very close to the plot origin). Note that the sensitivity, although very small, increases for a larger vertical baseline $d_v$. The two sensors in the horizontal network (Fig. 4) were separated by a distance $d_h$ and were located at $(−d_h/2, 0, 0)$ km and $(d_h/2, 0, 0)$ km; the values $d_h = 0$ km (left column) and $d_h = 100$ km (right column) were used. The values $\eta_1$ (top row), $\eta_y$ (middle row), and $\eta_z$ (bottom row) are provided, and they did not require amplification. Again, note that the sensitivity improves as the horizontal baseline $d_h$ increases. Overall, even though a network consisting of two sensors (and employing time differences) can be very sensitive to some fluctuations of a distant source, such sensitivities can quickly vanish for other source azimuths; for example, the top two rows in Fig. 4 show “network blind spots” (intrusion of a smaller SNR denoted by the red regions). Fortunately, for a network composed of many sensors, the blind spots of one pair of sensors is often removed by good sensitivity of another pair of sensors.

Determining how much information is contained in a set of $(t_1, \ldots, t_m)$ measurements, where $m > 2$, is more complicated, but insight can be gained by performing systematic retrieval simulations (see sections 5–9). Since LMA networks typically have relatively small vertical baselines in comparison with their horizontal baselines, it is usually more difficult to retrieve $z$ than the other source parameters. As such, the focus of the simulations in sections 5–9 is on mitigating the $z$ retrieval errors.
4. Line of sight

Before providing retrieval error results, it is important to first understand what the maximum horizontal distance between an LMA sensor and a VHF source can be before it is obstructed by Earth’s curvature. Ignoring atmospheric refraction, the geometry for the LOS between an LMA sensor at altitude $z_i$ and a source at $z$ is shown in Fig. 5a. Applying the Pythagorean theorem to the right triangles shown allows one to easily solve for the distances $d_1$ and $d_2$. Viewing the LMA sensor altitude to be fixed at $z_i$, the maximum LOS distance $D_i^*$ as a function of source altitude is given by

$$D_i^*(z) = \sqrt{r_e^2 + 2z_r z_i + r_e^2 + 2r_e z}.$$  \hfill (6)

As an illustration, a value of $z_i = 218.6 \text{ m}$ (the altitude above sea level of the Alabama A&M University site of the North Alabama LMA network) is assumed. The geodetic latitude of this site is just under 34.9°, which corresponds to a geocentric latitude of just above 34.7°. The 1984 World Geodetic System (WGS-84) Earth ellipsoid model can be used to compute Earth’s radius $r_e$ at the geocentric latitude of the site. A value of $r_e = 6371.1763 \text{ km}$ is obtained, and the associated plot of $D_i^*(z)$ is provided in Fig. 5b. This shows that a source, say, 400 km away in horizontal range from the site has to be at least 9.45 km in altitude in order to have a clear LOS to the site. If the source is at a lower altitude, then Earth’s curvature obstructs the LOS.

Such physical limits need to be kept in mind when interpreting the spatial distribution of retrieval errors presented below, and in other investigations. Generally speaking, the plots to follow in this study extend out to the 400-km range in the $x$ and $y$ directions. Hence, it should be understood that the source technically needs to meet its minimum height value to attain unobstructed LOS. In some of the simulations provided in this study, low source altitudes are tested to provide the reader a more comprehensive understanding of retrieval errors. Again, it should be understood that the tests are valid only up to a certain horizontal range (even though a particular plot might exceed this valid range).

5. Baseline run

Figure 6 summarizes source altitude retrieval errors obtained from a Monte Carlo simulation. In the simulation, a known source with a fixed altitude is tested at a particular $(x, y)$ location. The activation time of the source is set to $t = 0 \text{s}$. The arrival times are created at each fictitious LMA sensor; the simulated LMA network has nine sensors arranged on a $3 \times 3$ square grid. The network dimensions are $50 \text{ km} \times 50 \text{ km}$, so the distance between adjacent sensors in the $x$ and $y$ directions is 25 km. Note that using a square grid eliminates systematic azimuthal biases in the retrieval error patterns that would occur with an irregular network site geometry. The sensor altitudes are fixed; each sensor altitude was defined by randomly selecting an altitude in the (uniformly distributed) range 0–0.5 km. This resulted in a specific sensor altitude range of 0.085–0.491 km, which is realistic; for example, the North Alabama LMA network had sensor altitudes in the range 0.1720–0.5211 km (Koshak et al. 2004). The arrival time at each site is calculated from the known source, and errors (randomly picked from a normal distribution with mean zero and $\sigma = 50 \text{ ns}$) are added to these values.

The simulated measurements are then inverted using the standard approach described in section 2 (i.e., the Koshak–Solakiewicz linear retrieval initialization, followed by the Levenberg–Marquardt algorithm search). A total of 100 such simulated retrievals are performed at each $(x, y)$ known source test point, and the process is repeated for many other test points across a wide (400 km $\times$ 400 km) domain. Figure 6 shows the mean altitude retrieval error at each test location. Note in this baseline run that the altitude retrieval errors increase for lower and more distant sources, as expected from the SNR analysis given above.

6. Adding more measurements

Normally the first reaction toward combating large retrieval errors is to consider deploying more sensors, which unfortunately can be expensive. Indeed, deciding how many measurements to make is fundamental in any inverse problem. Given measurement errors, a set of $m$ measurements do not necessarily provide $m$ independent pieces of information about the unknown source. For example, consider just two LMA sensors. If the measurement errors are sufficiently large and the two LMA sensors are located sufficiently close to each other, then the two sensors would not provide any independent (i.e., distinct) pieces of information about the unknown source. Therefore, one must be careful in deciding if the costs of an additional sensor have any real return on investment. This is exactly where simulations of the type presented here can be used. In particular, see the network-specific simulation tool developed in Chmielewski and Bruning (2016).

Figure 7 quantifies the benefits of adding additional sensors. Coupled with the baseline run (Fig. 6, upper-right plot), the simulation series covers $m = 9, 16, 25, 36$, and 49 sensors. The assumed standard deviation in measurement error is fixed at $\sigma = 50 \text{ ns}$. To avoid improvements as a result of an increasing network size, note that the network size is fixed at 50 km. However,
there is a slight random modulation in retrieval error as a result of varying sensor altitudes (which are still randomly varied from 0 to 0.5 km). Overall, there are clear benefits to increasing the number of sensors. However, one would have to agree that only a marginal reduction in retrieval error is achieved for significant additional cost (including increased maintenance costs).

7. Expanding network horizontal extent

The SNR analysis provided in section 3 (Fig. 2a and the third row of Fig. 4) directly indicates that one should expect improvements in the altitude retrieval by simply expanding the horizontal extent of the network, because such an expansion increases the magnitude of the SNR (note

---

**Fig. 5.** (a) Geometry for the LOS calculation. Refraction is ignored. Sensor/source heights are exaggerated for clarity. (b) Maximum LOS distance between the A&M site and a VHF source of altitude \( z \).
that with respect to Fig. 2a, the increase in SNR is accomplished by the fact that the horizontal network expansion decreases the $D_i$ between the source and the sensor, for those sensors that expand toward the source).

Indeed, the Monte Carlo simulations provided in Fig. 8 clearly demonstrate the beneficial effect of increasing the network size. The network size is defined as the distance along one side of the square network. Note here that the number of sensors is kept fixed. A network size of 50 km has already been provided in the baseline run shown in the upper-right plot of Fig. 6, and the network sizes (100, 200, 300, 400 km) are shown in Fig. 8 to complete the series.

Note that the regions of small retrieval error directly over each site are consistent with the fact that the SNR increases dramatically over a site as given in Fig. 2a. The reduction in altitude retrieval errors is significant, and these results emphasize the importance of expanding the network as far as practical. Although the plots are omitted for brevity, note that expanding the horizontal extent of the network also significantly reduces the horizontal location retrieval errors because (as seen in the first two rows of Fig. 4) horizontal expansion improves the SNR. In summary, far more benefit is achieved by simply expanding the horizontal extent of the network than adding many sensors as in Fig. 7; that is, an increase in network size from 50 to only
100 km produces about as much improvement in the retrieval of distant sources as adding 40 sensors!

LMA networks continue to proliferate, and the simulation results presented here emphasize the importance of coalescing networks into larger size networks whenever possible. Although it takes longer for maintenance crews to service distant sites, this issue can be resolved by delegating these work activities to partners that are closer to the distant sites.

8. Increasing a sensor altitude

In this section the reduction in altitude retrieval error as a result of increasing the altitude of just one sensor is systematically examined. For example, some networks are deployed near mountainous terrain, and so one wonders if there are any benefits in placing one of the network sensors in that mountainous terrain. Clearly, one would not wish to do so if the mountains adversely obstruct LOS in any way. For a square network geometry, one could envision placing one corner sensor in the mountains and the remaining sensors in the adjacent valley so that storm detection would be unobstructed throughout that valley. This is the case simulated here.

Note that the benefit of increasing the vertical separation between sensors has been discussed and demonstrated in Koshak et al. (2004), but only one
(less realistic, less systematic) simulation was performed using some rather large sensor altitudes and an irregular network geometry. Nonetheless, the study and section 4.3 of Thomas et al. (2004) give detailed substantiation as to why it is more difficult to accurately retrieve source altitude when the sensor altitudes are similar.

Figure 9 demonstrates the improvement when the upper-right corner site, located at \((x = 25 \text{ km}, y = 25 \text{ km})\), takes on the altitude values \((0.5, 1.0, 2.0, 4.0) \text{ km}\). The upper-right corner site in the baseline run (Fig. 6, upper-right plot) had a sensor altitude of \(0.085 \text{ km}\) (the lowest sensor in the baseline run). So, conjoining these figures gives an altitude series \((0.085, 0.5, 1.0, 2.0, 4.0) \text{ km}\). All the other sites have fixed altitude values identical to what was used in the baseline run (i.e., their altitudes are all below 0.5 km).

Overall, the reduction in altitude retrieval errors seen in Fig. 9 is significant, and these results emphasize the importance of taking advantage of mountainous topography when possible. Note that a sensor altitude of 4 km is 13,120 ft \((1 \text{ ft} = 0.305 \text{ m})\), which is still below the highest summit in the contiguous United States (i.e., Mount Whitney at 14,505 ft). Mounting a sensor on a tower or a skyscraper would also offer some improvement (all else being the same, including no additional sources of noise from the building or tower), but obviously not as much as
provided by taller mountaintops. Finally, the patterns of retrieval error for horizontal distance and source time of occurrence showed little change, that is, only secondary distortions in the error pattern (the plots are omitted for brevity).

9. Reducing measurement errors

Clearly, one expects source retrieval errors to decrease as the $\sigma$ in measurement errors decreases. In fact, as $\sigma$ approaches zero, the linear retrieval method (Koshak and Solakiewicz 1996; Koshak et al. 2004) approaches perfect retrieval results in all four source parameters ($x, y, z, t$), to within computer machine precision, and the Levenberg–Marquardt algorithm is then not even needed. However, as mentioned in section 2, $\epsilon_i$ (and hence $\sigma$) also technically depends on modeling error. Hence, the value of $\sigma$ actually approaches zero only in the limit of perfect measurements and a perfect model.

Figure 10 quantifies the benefits of achieving smaller measurement errors. From comparing Fig. 10 with the results of Fig. 7, it is clear that having a few very accurate sensors is superior to having many less accurate sensors.

10. Summary

This study has revisited the details of inferring the location and time of occurrence of a VHF lightning

---

**Fig. 9.** The benefit of increasing the altitude of the upper-right sensor located at $x = 25$ km, $y = 25$ km. The source altitude is 7 km.
source emission from lightning mapping array network TOA measurements in order to better clarify the cause of retrieval errors and to determine how best to mitigate these errors. A cardinal rule of any inverse problem is to first thoroughly inspect the associated forward problem; that is, by seeing how well changes in the location and time of occurrence of a simulated VHF source manifest themselves in the associated simulated TOA measurements, relative to the measurement noise level. This provides vital insight. For example, if one changes the source location in a certain direction across a certain distance, but the associated change in the TOA values is smaller than the simulated TOA measurement errors, then there is fundamentally no hope of retrieving this level of detail about the source location (unless one adds better independent constraints to the problem, such as new/better measurements or new and meaningful mathematical/physical constraints).

Therefore, after providing a rigorous statement of the inverse problem in section 2, we have devoted section 3 to investigating the associated forward problem. Considering just one sensor and one VHF source, we have derived the fundamental SNR formulas [(2)–(4)] for the simulated measurements that are associated with vertical, horizontal, and temporal changes in the simulated VHF source, respectively. This connection is important since the SNR values at or below unity represent fundamental limits on the level of retrievable source detail. Figure 2 shows the SNR values for the vertical and horizontal source displacements. Vertical

---

**Fig. 10.** The benefit of reducing the standard deviation of the measurement error. The source altitude is 7 km.
displacements at low altitude and/or at large horizontal distances provide very little signal relative to noise. Transverse source displacements are more difficult to sense than longitudinal displacements. We have also gone one step further in the analysis by looking at the fundamental information content associated with a two-sensor system when the TOA values are differenced. Variation in the associated SNR is derived in (5), and the SNR plots are provided in Figs. 3 and 4 for vertically and horizontally separated sensors, respectively. The complexity and geometrical beauty of these plots are a reminder that one can extract different types of information based on how the basic TOA observations are combined, which is an “art form,” and deserves additional probing to fully optimize retrievals.

To provide clear recommendations for mitigating retrieval errors, sections 5–9 have been devoted to performing carefully designed Monte Carlo inversion simulations that provide specific retrieval error plots. Some simulations have been performed in previous studies (e.g., Koshak et al. 2004; Thomas et al. 2004; Bitzer et al. 2013; Chmielewski and Bruning 2016) but have several diverse features that make it difficult to clearly determine the overall results (i.e., variable network geometries, different approaches for plotting results, variable inversion methodologies, differences in assumed TOA measurement errors, conflating of multiple effects, and missing tests). This diversity also makes it difficult to determine what adjustments to the LMA network are most important in reducing retrieval errors. Therefore, this paper has performed all the simulations in a highly controlled way in order to avoid all this diversity so that a clear list of recommendations can be obtained. The standard Levenberg–Marquardt algorithm (with parameter initialization performed using the linear retrieval method introduced and applied in Koshak and Solakiewicz (1996) and Koshak et al. (2004)) was employed to carry out the simulated inversions. LOS computations are provided in section 4 to place logical bounds on the retrieval error plots. A baseline inversion run consisting of a square nine-station network geometry was first examined. The effect of adding more measurements without expanding the network (section 6) was then tested. The effects of expanding the nine-station network (section 7), increasing the altitude of one sensor in the nine-station network (section 8), and then decreasing TOA measurement errors (section 9) were then each independently tested.

Based on all of the numerical results, the following recommendations for mitigating source retrieval errors (particularly source altitude) are encouraged:

1. Expanding the horizontal extent of an existing set of sensors in a network will result in substantial reductions in source retrieval errors, as quantified in this study. Such expansions should be logically limited however to avoid any LOS issues, or issues associated with the limited propagation range of VHF signals (which can hamper source detection for those sensors expanded away from the source and can also complicate VHF source matching among the sensors).
2. In the deployment of an LMA network, placing a sensor in mountainous terrain will increase the vertical sensor baseline (which directly improves source altitude retrievals, as quantified in this study). However, this recommendation should be avoided if adverse LOS issues occur as a result of such an action.
3. It is beneficial to add sensors to the network in order to reduce retrieval errors. However, one should recognize that the relative return on such an investment is limited if the horizontal and/or vertical network baseline is not improved by such additions, as quantified in this study.
4. Technological improvements that reduce sensor measurement errors are strongly encourage and will result in associated reductions in retrieval errors as quantified in this study.

Note that this paper has involved a mathematically rigorous approach that confirms the basic guidance provided in previous simple geometric analyses and prior modeling work, but it has provided additional clarity and detail (i.e., formal forward problem analyses are provided for both vertical and horizontal source displacements and are generalized to arbitrary SNR values; conflating the effects of network expansion with the addition of sensors is avoided; and sensitivity simulations were performed very carefully in order to improve the overall ease of intercomparing the effects of adjusting certain LMA characteristics one at a time and for making sure that the complications/details of network geometry do not obscure the comparisons). In addition, this is the first paper to explicitly demonstrate (and stress the potential benefits of) more actively using topography to reduce the altitude retrieval errors.

Finally, it has been shown that the standard retrieval method (section 2) can be generalized (see appendix). In this generalization, the VHF point source model of section 2 is replaced by a source composed of an unpolarized VHF point source collocated with a transient VLF/LF electric point dipole source. This follows the notion that natural radiation sources have, in general, both an unpolarized and polarized component, as is quantified by the Stokes vector. The VHF breakdown radiation is assumed to have a mostly random (unpolarized) nature, whereas the radiated electric field in the VLF/LF is assumed to be mostly polarized, since it is associated with organized unidirectional current flow.
in the channel (P. Krehbiel, Institute of Mining and Technology, 2016, personal communications). To retrieve the generalized source, the chi-squared function in section 2 was generalized to include not only the LMA VHF TOA observations, but also 1) LMA VHF received power observations, 2) VLF/LF TOA observations, and 3) VLF/LF electric field amplitude observations. The VLF/LF observations are made with a ground-based flat-plate electric field antenna network that complements the LMA network. The ratio of the number of measurements to the number of unknowns is better in the generalized approach than in the standard retrieval method, and even more so for distant sources where the radiation term dominates in the VLF/LF field equation. However, much additional work is needed to test the practicality of the generalized method (e.g., via simulated retrievals and actual data inversions) so that retrieval accuracy is specifically characterized.

Acknowledgments. This research has been supported by NASA Headquarters as part of the NASA Marshall Space Flight Center Science Innovation Fund, under the direction of the former and current MSFC Science and Technology Office Chief Scientists Drs. James Spann and Gary Jedlovec, respectively. The authors are grateful to Dr. Michael Lapointe of the National Space Science and Technology Office Chief Scientists Drs. James Spann and Gary Jedlovec, respectively. The authors are grateful to Dr. Michael Lapointe of the National Space Science and Technology Center for his helpful guidance and encouragement throughout the term of this effort. This research has also been supported by the NASA ROSES-2014 NNH14ZDA001N-WEATHER program led by Dr. Ramesh Kakar of NASA Headquarters. In addition, we give thanks to Dr. Paul Krehbiel of New Mexico Tech and Dr. Kenneth Cummins of the University of Arizona for their helpful discussions on radio frequency lightning measurements.

APPENDIX

Generalizing the Standard Retrieval Method

The lightning point source model employed in section 2 involves only the random (i.e., unpolarized) VHF radiation originating from the highly random breakdown processes that are associated with the initiation of a lightning discharge. However, when currents eventually organize and propagate into a given direction (the direction of channel propagation), a more polarized VLF/LF radio emission can be considered and modeled using an electric point dipole source. Hence, in general, one can consider superposing the two radiation types: the VHF point source emission (assumed to be mostly unpolarized radiation associated with random breakdown) and the VLF/LF electric point dipole source emission (assumed to be mostly polarized radiation from organized channel current flow). This picture is consistent with the notion that natural radiation sources are, in general, partly unpolarized and partly polarized, as quantified by the Stokes vector for that radiation. Since the desire is to map the lightning channel out in space and time, short enough time intervals are considered so that the two components (the VHF point source and the VLF/LF electric point dipole) of the source are assumed to occur at the same time, t, and are collocated at r.

To retrieve this more robust two-component source, the generalized method involves using three additional measurement types beyond the standard VHF TOA observations discussed in section 2. First, in order to retrieve information about the VLF/LF electric dipole source, the generalized approach uses measurements of the vertical VLF/LF electric field amplitude as obtained from a ground-based network of mF flat-plate electric field antennas (EFAs). Second, TOA measurements of the VLF/LF wave obtained from the EFA network are used and are implemented exactly as in Bitzer et al. (2013). Third, the VHF LMA received power \( P_i = 10^{\frac{5}{10} \Omega} \) at each ith LMA site (in watts) is used, where \( \Omega_i \) is the received power (dBW).

Note that these three additional sets of measurements have not been directly applied within the mathematical retrieval process (i.e., in conjunction with the VHF TOA observations) to improve source retrievals. The additional data provide additional constraints on the location and time of occurrence of the source, and more information/details about the source itself. Hence, in order to fully optimize source retrievals, there is a desire to supplement the standard VHF TOA observations with these additional measurements. As shown below, such a desire leads to a generalization of the \( \chi^2 \) in (1) of the main text.

A full analysis of the generalized method presented here is beyond the scope of this appendix. In this “first step” consideration, the focus is to derive the basic form of the generalized retrieval method. Follow-on analyses/evaluations (e.g., retrieval simulation tests, actual data inversions) by the broader lightning research community are of course necessary, and are often protracted, before adequate retrieval results can be fully demonstrated and proven.

The VLF/LF emission can be viewed as being produced by a transient current surge across a finite but relatively short distance scale within the propagating lightning channel. That is, the VLF/LF emission is the result of a dipole moment change (represented by a point dipole vector \( p \)) produced by a channel current \( I(t') \) flowing through a very short channel segment \( ds \) in a brief interval of time \( dt' \) (where \( t' \) is a dummy time variable that is employed to avoid any confusion with the variable \( t \) for time of occurrence that was introduced in section 2). One could also consider the current to be a
function of distance along the channel segment, but again the length of this segment $\delta s$ is taken to be small enough to ignore such variations. For the very short time interval $\delta t$ considered, the orientation of the point dipole can be taken as fixed, and it can be viewed as being in the direction of the propagating lightning channel during that brief period. The magnitude of the point dipole moment varies during the period $\delta t$; that is, it can be written as $p(t)$. Note that some studies have probed lightning currents using dipole models, but the retrieval algorithms employed were significantly different from what is presented here and were restricted to investigating compact intracloud discharges (CIDs) that are oriented vertically (Nag et al. 2010; Nag and Rakov 2010a, b).

The electromagnetic fields at the origin of a Cartesian coordinate system as a result of a time-dependent, arbitrarily oriented electric point dipole source $p$ located at position $r$ above the $(x, y)$ Earth-conducting plane have been derived (Panyukov 1996; He et al. 2000; Popov and He 2000). Again, even though the dipole moment magnitude varies in time, its orientation, described by the spherical coordinate angles ($\lambda$, $\Theta$) is assumed to be fixed during the brief interval $\delta t$. Hence, the functional dependence of the source can be written as $p = p(t)\rho(\lambda, \Theta)$. In addition, the expression for the electric field at the origin must be generalized to the arbitrary EFA locations $r_{ej}$, where $j = 1, \ldots, m$, as shown in Fig. A1. But, because the fields are evaluated on the conducting-Earth plane (no topography in the model), $z_{ej} = 0$ must hold exactly (strictly speaking), or at least approximately. To avoid redundant TOA observations in the VHF and VLF/LF, one should not place an EFA sensor at each LMA site. But, each network should cover roughly the same geographical region and each should have a sufficient number of sensors to afford accurate retrievals.

With these provisions, (3) of He et al. (2000) for the vertical electric field $E_z$ can be generalized to the vertical electric field $E_{zj}$ at the $j$th EFA as

$$E_{zj}(t) = \frac{1}{2\pi\varepsilon_0 c^2 R_{ej}} \times \left\{ \left( \beta_j - \frac{2}{3}\cos\Theta \right) [\hat{p}] + \frac{3c\beta_j}{R_{ej}} [p] + \frac{3c^2\beta_j}{R_{ej}^2} [p] \right\}. \quad (A1)$$

The three terms within the curly brackets are, from left to right, the radiation, intermediate, and electrostatic terms. The square brackets indicate an evaluation at the retarded time $t_{ret} = t - R_{ej}/c$, for example, $[p] = p(t' - R_{ej}/c) = p(t_{ret})$. The dot and double dot above the variable $p$ indicate the first and second time derivatives, respectively. The geometrical variable $\beta_j$ is given by

$$\beta_j = \sin\theta_j \cos\lambda_j \sin\Theta \cos(\lambda_j - \Lambda) - \sin^2\theta_j \cos\Theta + \frac{2}{3}\cos\Theta, \quad (A2)$$

and a trigonometric identity gives $\cos(\lambda_j - \Lambda) = \cos\lambda_j \cos\Lambda + \sin\lambda_j \sin\Lambda$. The spatially dependent terms are
\[
\sin \theta_j = \frac{D_{ej}}{R_{ej}}, \quad \cos \theta_j = \frac{z - z_{ej}}{R_{ej}},
\]
\[
\sin \lambda_j = \frac{y - y_{ej}}{D_{ej}}, \quad \cos \lambda_j = \frac{x - x_{ej}}{D_{ej}}.
\]

(A3)

Here, the distances are provided in Fig. A1 (again, with \(z_{ej} \approx 0\) understood) and can be written as
\[
D_{ej} = \sqrt{(x_{ej} - x)^2 + (y_{ej} - y)^2}
\]
\[
R_{ej} = \sqrt{(x_{ej} - x)^2 + (y_{ej} - y)^2 + (z_{ej} - z)^2}.
\]

(A4)

Identifying \(\tau_{ej}\) as the measured TOA of the VLF/LF wave at the \(j\)th EFA, one can define the associated VLF/LF amplitude measurement \(a_j\) at the EFA sensor as
\[
a_j = 2\pi e_o c^2 E_{ej} (\tau_{ej}).
\]

(A5)

A mathematical model \(\mu_j\) for describing the amplitude measurement can be obtained from (A1). If \(\delta t'\) is not short in duration, then (A1) shows that the shape of the curves \(p(t')\), \(p'(t')\) and \(p''(t')\) affect when \(E_{ej}(t')\) will reach its peak. However, given a particular dipole orientation and considering \(\delta t'\) to be very short, the point dipole source creates a strong delta function–like VLF/LF radiation pulse that propagates along distance \(R_{ej}\) to excite the EFA. Similar to section 2, the transit equation for an arrival time of the VLF/LF wave is \(t_{ej} = t + R_{ej}/c\), where \(\tau_{ej} = R_{ej} + e_{ej}\) and \(e_{ej}\) represents the measurement error in the arrival time. Hence, in this situation, the time of occurrence \(t\) of the point dipole moment source is just the retarded time \(\tau_{ej}\). Therefore, the square brackets in the weights ([\(\bar{p}\), \([p]\), \([\bar{p}]\)]) can be removed in favor of writing these weights explicitly in terms of \(t\), that is, \([\bar{p}] = \bar{p}(t) = w_1\), \([p] = \bar{p}(t) = w_2\), and \([\bar{p}] = p(t) = w_3\). Since \(t\) is a constant, the weights should be viewed as just three variables, not as three functions. Hence, the model \(\mu_j\) is just a function of eight unknowns \((x, y, z, \Lambda, \Theta, w_1, w_2, w_3)\). With the abbreviation \(w = (w_1, w_2, w_3)\), the model \(\mu_j\) can be written as
\[
\mu_j = \mu_j(r, \Lambda, \Theta, w)
\]
\[
= \frac{1}{R_{ej}} \left[ w_1 \left( \beta_j - 2 \cos \theta \right) + w_2 \frac{3e \beta_j}{R_{ej}} + w_3 \frac{3c^2 \beta_j}{R_{ej}} \right].
\]

(A6)

With the preceding comments, one can construct chi-squared functions for both the VHF and VLF/LF datasets as
\[
\chi^2_{\text{VHF}} = \sum_{i=1}^{m} \frac{[\tau_i - (t + R/c)]^2}{\sigma_i^2} + \sum_{i=1}^{m} \frac{[P_i - P_i A/(4\pi R^2)]^2}{\sigma_i^2},
\]
\[
\chi^2_{\text{VLF/LF}} = \sum_{j=1}^{n} \frac{[\tau_j - (t + R_{ej}/c)]^2}{\sigma_{aj}^2} + \sum_{j=1}^{n} \frac{(a_j - \mu_j)^2}{\sigma_{aj}^2}.
\]

(A7)

Here, \(P_i\) is the VHF power (in watts) measured at the \(i\)th LMA sensor, \(A\) is the area of an LMA antenna, and \(P_i\) is the VHF power (in watts) emitted by the VHF point source. The variables \(\sigma_{pi}\), \(\sigma_{ej}\), and \(\sigma_{aj}\) represent the standard deviation of the measurement errors in VHF power \(P_i\), VLF/LF TOA \(\tau_{ej}\), and VLF/LF amplitude \(a_j\), respectively. Finally, the generalized \(\chi^2\) appropriate for this generalized approach is
\[
\chi^2 = \chi^2(r, t, P, \Lambda, \Theta, w) = \chi^2_{\text{VHF}} + \chi^2_{\text{VLF/LF}}.
\]

(A8)

The retrieval solution can be found in the usual way by employing a numerical method (e.g., the Levenberg–Marquardt algorithm) to minimize the cost function in (A8).

Note that Thomas et al. (2004) conjectured that received powers cannot be usefully applied in the retrieval process because of local radiation effects (interference, attenuation) and because of unknown lightning source radiation patterns. However, we consider this conclusion to be premature, since a formal quantitative information content analysis of the received power data has not yet been conducted. At this stage we believe it to be an open question, and we would frankly be surprised if the power contained zero information about the VHF source, despite the legitimate concerns noted by Thomas et al. (2004).

Whereas the standard retrieval method (section 2) involved \(m\) arrival time measurements for constraining 4 unknowns \((x, y, z, t)\), the generalized method involves \(2(m + m_e)\) measurements for constraining 10 unknowns \((x, y, z, t, P, \Lambda, \Theta, w_1, w_2, w_3)\). If one sets \(m_e = m\), then this would result in \(4m\) measurements constraining 10 unknowns. Hence, the standard retrieval using just VHF TOA data would imply a “constraining ratio” of \(m/4 = 0.25m\), whereas the generalized method has a larger (i.e., better) constraining ratio of \(4m/10 = 0.4m\). In this case the reduced chi-squared associated with (A8) is \(\chi^2/\nu\), where \(\nu = 4m - 10\). So, for example, with \(m = 10\), the degrees of freedom in the generalized method are \(4(10) - 10 = 30\), as compared with \(m = 4\) in the standard method. In addition, the generalized method offers the possibility of extracting more details about the lightning source, namely, the parameters \((P, \Lambda, \Theta, w)\), which, in particular, can provide insight into lightning channel energetics.
Moreover, when a source is far from the networks, the generalized method offers a significant advantage over the standard method; that is, for distant sources, the radiation term in (A1) dominates so that weights \( w_3 \) and \( w_5 \) need not be retrieved. Hence, the constraining ratio increases from 0.4m to 4m/8 = 0.5m. This theoretical finding is important to note because, as has been clearly demonstrated in the forward problem analyses and simulations in the main text, distant sources are very difficult to retrieve accurately. If wave dispersion in the VLF/LF relative to the VHF is sufficiently large, then distinct propagation velocities for these two frequency regimes would be required (and the assumption of a unique source time for the two-component source remains in force).

Overall, the concept of the generalized method introduced here offers the possibility of further mitigating retrieval errors in the variables (\( x, y, z, t \)), and it has the added benefit of potentially providing more information about the details of the source. However, in order to handle the complicating effects mentioned above (i.e., interference, attenuation, source radiation pattern, and wave dispersion), further modifications to the power and electric field amplitude modeling terms in (A7) might be required before practical retrievals can be fully realized. On the other hand, it might be found that the terms in (A7) as written are adequate for obtaining practical retrievals. This cannot be decided at this early stage, and it will require future experimental, numerical, and theoretical tests to fully resolve.

REFERENCES


