JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. D12, PAGES 25,721–25,729, DECEMBER 20, 1995

An analysis of errors in the location, current, and velocity of lightning

William W. Hager and Dongxing Wang

Department of Mathematics, University of Florida, Gainesville

Abstract. Recently, Le Vine and Willett obtained a correction to the transmission line model for computing radiation from lightning. In this paper we examine procedures, based on the formula of Le Vine and Willett, for obtaining the location, velocity, and current of lightning discharges from multiple station measurements such as those being implemented by Thomson. In the appendix we apply the algorithms of this paper to recently obtained data for an observed in-cloud pulse. We show that the propagation velocity of the pulse is 0.61 c, while the channel length is 118.7 ± 9.2 m. In numerical simulations of lightning discharges, we observed in an earlier paper that before each flash, there was a channel formation phase where electric breakdown occurred along short channels in the cloud with length on the order of 100 m. Hence the length of the channel associated with this pulse is consistent with the length of the breakdown regions during the channel formation phase observed in the numerical simulations.

1. Introduction

Utilizing the transmission line model [see *Uman and McLain*, 1969, 1970] for computing radiation from lightning, *Le Vine and Willett* [1992] obtain the following formula relating the electric field $\vec{E}$ at an observer to the location of a lightning channel and the velocity $\vec{v}$ of propagation:

$$\vec{E}(t) = -\frac{c\mu_0}{4\pi R}\left[\frac{\vec{v}-(\vec{v}\cdot\hat{R})\hat{R}}{c-\vec{v}\cdot\hat{R}}\right](I(t)-I(t+t_a-t_b-t_c)). \quad (1)$$

Here $\vec{R}$ is the vector from the observer to the midpoint of the lightning channel, R is the length of $\vec{R}$, $\hat{R}$ is the unit vector $\vec{R}/R$, t_a is the time for light to travel from the start of the channel to the observer, t_b is the time for light to propagate from the end of the channel to the observer, and t_c is the time for the current wave to propagate the length of the channel. In deriving (1), terms of order $1/R^2$ were neglected, and it was assumed that the length of the channel is much smaller than R.

This formula can also be applied to a long channel by partitioning it into short segments, applying (1) to each segment, and adding the contributions to the electric field due to each segment. In particular, let s be a parameter that measures distance down the channel ($s = 0$ corresponds to the start of the channel), let $\vec{R}(s)$ denote the vector from the observer to location s on the channel, and define

$$\Phi(s) = \frac{c\mu_0}{4\pi R(s)}\left[\frac{\vec{v}-(\vec{v}\cdot\hat{R}(s))\hat{R}(s)}{c-\vec{v}\cdot\hat{R}(s)}\right].$$

Then for a straight channel, we have

$$\vec{E}(t) = \Phi(L)I(t+t_a-t_b-t_c) - \Phi(0)I(t) - \Lambda(I)(t) \quad (2)$$

$$\text{where } \Lambda(I)(t) = \int_0^L I(t+t_a-(s/v)-R(s)/c)\Phi'(s)ds,$$

Paper number 95JD02527.
0148-0227/95/95JD-02527$05.00

and v is the length of $\vec{v}$. The formula (2) also applies to a curved channel without sharp bends if $\vec{v}$ in the definition of Φ is replaced by $\vec{v}(s)$. When the channel contains sharp bends, additional terms enter (2) corresponding to the jump in Φ across each bend. Both the $\Phi(L)$ and $\Phi(0)$ terms in (2) are $O(1/R)$ while the Λ term is $O(L/R^2)$ for large t and $O(t/R^2)$ for small t.

In Thomson's system for multiple station measurements of lightning, described by *Thomson et al.* [1994], the vertical component of the electric field is measured. In this paper we examine the problem of determining the location of the lightning channel, the current profile I, and the velocity of propagation for a straight channel using measurements of $\vec{E}_z(t)$ at several stations. The location of the start of the lightning channel is obtained by analyzing the time-of-arrival of the electric field at the stations. A minimal variance estimate is obtained by solving a nonlinear least squares problem, while a linear approach involving averaged equations yields almost as much accuracy as the nonlinear approach, but with much less computational effort. An estimate for the propagation velocity and initial current is obtained by solving a bilinear least squares problem by computing the roots of a related 10th degree polynomial. It is observed that the solution to this bilinear least squares problem is very sensitive to measurement errors, and that tiny measurement errors can lead to huge errors in the velocity. Even though the global solution to this least squares problem may be completely meaningless, one of the local minimizers was always a good approximation to the physical velocity. The physically correct local minimizer can be determined using a least squares fit of the data over a long time interval. This long time fit algorithm utilizes the initial guesses for the velocity obtained from the roots of the 10th degree polynomial, a finite element approximation to the current profile, a linear least squares system to determine the coefficients in the finite element approximation, and an optimization algorithm to find the length of the channel and the propagation velocity.

The analysis in this paper is mainly relevant to in-cloud lightning. For work related to cloud to ground strokes, see *Thottappillil and Uman* [1993, 1994], and *Rachidi and Thottappillil* [1993]. In these papers the authors compare the current in triggered lightning experiments to the far-field radiation for many models, including the transmission line model.

Recently, Thomson provided us with the digitized version of Figure 8 [*Thomson et al.*, 1994]. This digitized data allowed us to apply algorithms of this paper to an observed in-cloud pulse. The results of this analysis are reported in the appendix. It was determined that the propagation velocity of this in-cloud pulse is 0.61 c, while the channel length is 118.7 m with a standard deviation of 9.2 m. This channel is much shorter than the typical channel associated with lightning. In numerical simulations of lightning discharges reported by *Hager et al.* [1989], we observed that each flash was often preceded by a channel formation phase during which time there were a series of short discharges in the cloud along channels with length on the order of 100 m. The channel length of the pulse reported in Figure 8 of *Thomson et al.* [1994] is consistent with the length of these short discharges during the channel formation phase in the simulations.

2. Location

Let (x,y,z) denote the location of the start of the lightning channel, let (x_i,y_i,z_i) denote the location of station i, and let t_i denote the arrival time at station i of the radiation that first started to propagate at time t. (Throughout this paper, we place a subscript i on any observable to denote its value at station i). The equations describing the location of the start of the lightning channel are

$$\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2} = c(t_i-t), \tag{3}$$

$i = 1, \cdots, m$, where m is the number of stations. Although in theory the location and time of the lightning can be determined from the measurement of t_i at four different stations, there will be errors in the measurement of the t_i, so it is better to obtain data at more than four stations, and then compute the location and time of the lightning that, in some sense, satisfies all the equations as closely as possible.

Let e_i denote the error associated with the ith equation in (3):

$$e_i = \sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2} - c(t_i-t).$$

One common method to solve an overdetermined system of equations is to compute the solution that minimizes the least squares error; that is, we replace the system (3) by the least squares problem

$$\text{minimize} \sum_{i=1}^{m} e_i^2. \tag{4}$$

(When different stations have different variances for their timing errors, each e_i^2 in (4) should be divided by the corresponding variance to obtain a location estimate with minimal variance). The solution to the nonlinear least squares problem (4) can be obtained using a variety of algorithms for unconstrained optimization (for a survey, see section 2 and 10 by *Hager et al.* [1993]). These algorithms are all iterative techniques that start with some given estimate for the solution, and successively improve that estimate.

We now discuss techniques for obtaining a good approximation to a solution to (4). Historically, the location problem has been approached in the following way: Each equation in (3) is squared to obtain the relations:

$$(x-x_i)^2 + (y-y_i)^2 + (z-z_i)^2 = c^2(t_i-t)^2, \tag{5}$$

$i = 1, \cdots, m$. Then one of the stations, say the station corresponding to $i = 1$, is chosen as the "central station," and the equation for the central station is subtracted from the other equations to obtain a linear system of $m - 1$ equations in the unknowns x, y, z, and t. For some of the early treatments of the location problem [see *Lewis et al.*, 1960; *Proctor*, 1971], more recent ideas on minimizing errors are being developed by *Koshak and Christian* [1994]. If $m \geq 5$, then there are at least as many equations as unknowns, and in principle, we can solve this overdetermined system in a least squares sense. In practice, there are important special cases that need to be considered; for example, if $z_i = 0$ for each i, then the coefficient of z in the linear system is zero. In this case, the linear system determines x, y, and t, while z is obtained by substituting the computed values of x, y, and t in (5).

There are two fundamental problems with the central station approach. First, the equation

$$(x-x_1)^2 + (y-y_1)^2 + (z-z_1)^2 = c^2(t_1-t)^2$$

associated with the central station contains measurement errors. Hence when we subtract this equation from the others, the measurement errors associated with the central station are combined with the errors of the other stations. Second, if $z_i = 0$ for each i and we solve for z using the equation

$$z = \sqrt{c^2(t_1-t)^2 - (x-x_1)^2 - (y-y_1)^2}, \tag{6}$$

then the error in z may be huge, even though x, y, and t are determined relatively accurately. To see this second error, let us consider the equation

$$z = \sqrt{\tau^2 - \sigma^2}$$

where τ and σ are given scalars used to compute z. Of course, τ and σ correspond to the temporal and spatial terms of (6) respectively. To first order, the relative change $\Delta z/z$ corresponding to the change $\Delta\tau$ in τ can be expressed:

$$\frac{\Delta z}{z} \approx \frac{\tau^2}{z^2}\frac{\Delta\tau}{\tau}.$$

In other words, the relative change in z is τ^2/z^2 times the relative change in τ. If τ is near σ, then z is much smaller than τ and the error amplification factor τ^2/z^2 can be large.

We propose the following methodology for dealing with both these errors. Although it may be impossible to satisfy all the equations in (5) (due to measurement errors), we would like to compute a solution that in some sense averages out the error; more precisely, we would like the computed solution to satisfy the equation

$$\frac{1}{m}\sum_{j=1}^{m}(x_j-x)^2 + (y_j-y)^2 + (z_j-z)^2 \approx \frac{1}{m}\sum_{j=1}^{m} c^2(t_j-t)^2. \tag{7}$$

After subtracting the equation average (7) from the overdetermined system (5), we obtain a linear system of the form

$$A\vec{w} = \vec{b}, \quad \vec{w} = \begin{bmatrix} x \\ y \\ z \\ ct \end{bmatrix}, \tag{8}$$

where A is a $m \times 4$ matrix whose i^{th} row is

$$2\sum_{j=1}^{m}[x_j - x_i,\ y_j - y_i,\ z_j - z_i,\ c(t_i - t_j)],$$

Table 1. Location (meters) of Thomson's Stations

Station	x	y
1	0.00	0.00
2	−10081.53	2302.47
3	3683.20	7238.87
4	3531.41	−8592.09
5	−1593.04	−10151.32

and $\vec{b}$ is a vector whose i^{th} component is

$$b_i = \sum_{j=1}^{m} [x_j^2 - x_i^2 + y_j^2 - y_i^2 + z_j^2 - z_i^2 + c^2(t_i^2 - t_j^2)].$$

When the columns of A are linearly independent, the least squares solution to the overdetermined linear system $A\vec{w} = \vec{b}$ can be expressed $\vec{w} = (A^TA)^{-1}A^T\vec{b}$ (see chapter 5 of *Hager* [1988]). Note that when $m = 4$, the system $A\vec{w} = \vec{b}$ constitutes four equations in four unknowns and typically, there is a unique solution that satisfies all the equations exactly, and the errors e_i are zero. (Of course, the location of the lightning is still in error, even though the e_i are zero, due to errors in the measurement of the t_i.)

When the columns of the A matrix in (8) are linearly dependent, this strategy for obtaining an estimate of the location and time of lightning must be modified since the matrix A^TA is singular. Let us again discuss the important singular situation where all the stations are at the same elevation, say $z_i = 0$ for each i. In this case the column of A corresponding to z is entirely zero, and the least squares solution to (8) will only determine x, y, and t. We again use (5) to solve for the unknown value of z using the previously computed estimates for x, y, and t; however, instead of using a fixed station, like $i = 1$ to solve for z, we use the value of i for which the travel time $t_i - t$ is smallest. For this particular i, the error amplification factor τ^2/z^2 will be smallest.

As an illustration of these techniques, we consider the following numerical experiment: The location of Thomson's field detection stations at the Kennedy Space Center (KSC; as of 1994) appear in Table 1 (the stations are essentially at sea level, and station 1, near the shuttle landing facility, is chosen to be the origin of the coordinate system; the x axis is east/west while the y axis is north/south). We suppose that lightning occurs at elevations varying between 1 km and 12 km, and that the standard deviation of the error in timing measurements is 30 ns. We consider a uniform rectangular grid at each elevation inside the box -20 km $\le x \le 10$ km and -15 km $\le y \le 15$ km, with 500 m separating adjacent grid points in the x and the y directions. At each grid point, we consider 100 instances of lightning emanating from this point, and we perturb the arrival times at each station using a normal distribution with standard deviation 30 ns. For each instance of lightning, we determine the relative error in location corresponding to the central station approach, to the average equation approach, and to the nonlinear least squares approach (4). We solve the nonlinear least squares problem (4) in the following way: Using the central station approach, we obtain a guess for the solution to (4), and using several iterations of the conjugate gradient routine of *Hager* [1989], we refine this guess. In Table 2 we give the average location error at each elevation.

Table 2. Average Error (meters) in Location for 30 ns Timing Error

Elevation	Error: Central Station	Error: Average Equation	Error: Least Squares (4)
1 km	152.8	105.9	98.4
2 km	87.7	63.8	59.9
3 km	72.7	55.0	52.0
4 km	64.5	55.2	51.3
5 km	57.1	48.6	46.4
6 km	58.4	52.1	49.5
7 km	66.3	58.8	56.0
8 km	63.2	56.9	54.8
9 km	56.9	53.6	51.4
10 km	66.3	62.9	61.0
11 km	65.0	62.0	60.5
12 km	69.5	63.2	62.5

On the basis of the errors reported in Table 2, it appears that the average equation approach is almost as accurate as the nonlinear least squares approach while the central station approach has error between 8.7% greater (near 10 km) and 50% greater (near 1 km) than the error in the nonlinear least squares approach. In a certain sense, the solution to the least squares problem (4) is the best possible estimate one can obtain for the location. More precisely, for a linear problem, it can be shown [see *Strang*, 1986] that the linear unbiased estimate for the solution with smallest variance is the one that minimizes the least squares expression (4) (as noted earlier, each of the error terms e_i^2 needs to be weighted by the inverse of the variance of measurement error when different stations have different timing errors). Although the location problem is not a linear problem, one can perform a local linearization at any given point, and study the problem of obtaining location estimates at this point with minimal variance, assuming small measurement errors. According to the linear theory, the solution to the least squares problem (4) provides the estimate for location with smallest variance. Hence, for small errors in timing measurements, the

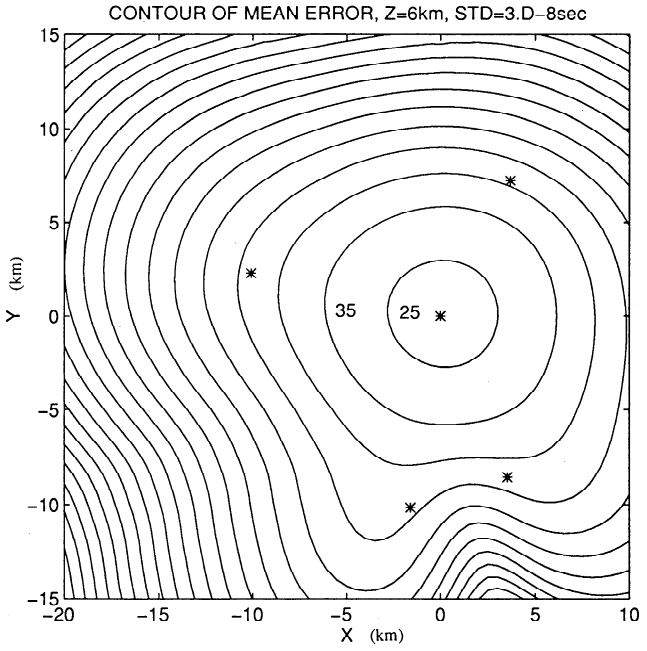

Figure 1. Mean error (meters) at Kennedy Space Center (KSC), altitude = 6 km, 10 m between contours

solution to (4) is the best possible location estimate for the nonlinear problem as well.

In a second series of numerical experiments, we determine the average location error for the nonlinear least squares approach at each grid point. In Figure 1, we plot the contours of constant average error for $z = 6$ km. In contour plots like this one, it is observed that for lightning near the ground, there are regions above each station where the location of the lightning can be determined relatively accurately, however, between the stations, the errors can be 10 times larger. As the altitude increases, the pockets of low error above each station coalesce together and eventually the error profile resembles a bowl with its low point near the centroid of the stations.

3. Velocity

Let us consider a value of t close enough to zero that the first and last term on the right side of (2) are negligible compared to the $\Phi(0)$ term. Hence at this value of t, we have the approximation

$$\vec{E}(t) \approx -\frac{c\mu_0}{4\pi R_a}\left[\frac{\vec{v} - (\vec{v}\cdot\hat{R}_a)\hat{R}_a}{c - \vec{v}\cdot\hat{R}_a}\right] I(t). \tag{9}$$

In this section we consider the problem of obtaining an estimate for the propagation velocity $\vec{v}$ and the initial current profile utilizing (9). Treating (9) as an equality, and evaluating at the time $t = \delta t$, we obtain the relation

$$E_i(\delta t) = -\frac{c\mu_0}{2\pi R_{a_i}}\left[\frac{\vec{v}_z - (\vec{v}\cdot\hat{R}_{a_i})(\hat{R}_{a_i})_z}{c - \vec{v}\cdot\hat{R}_{a_i}}\right] I(\delta t), \tag{10}$$

$i = 1, \cdots, m$, where E_i is the vertical electric field at station i. After multiplying through by the denominator in (10) and rearranging, we obtain an equation in the form

$$(P + IQ)\vec{v} = \vec{r}, \tag{11}$$

where the scalar I corresponds to $I(\delta t)$ in (10), and where row i of $P = 2\pi\vec{R}_{a_i}E_i/c\mu_0$, row i of $Q = \hat{R}_{a_i}(\hat{R}_{a_i})_z - (0, 0, 1)$, and $r_i = 2\pi R_{a_i}E_i/\mu_0$.

For any given I, the least squares solution to (11) can be written

$$v = (C^TC)^{-1}C^T\vec{r}, \quad C = P + IQ.$$

For this choice of I, the least squares error is

$$(C\vec{v} - \vec{r})^T(C\vec{v} - \vec{r}) = \vec{r}^T\vec{r} - \vec{r}^TC(C^TC)^{-1}C^T\vec{r}. \tag{12}$$

We now choose I to minimize the least squares error (12). Of course, when minimizing (12), we can neglect the $\vec{r}^T\vec{r}$ term since it is independent of I. Hence minimizing (12) is equivalent to maximizing $g(I)$ defined by

$$g(I) = \vec{r}^TC(I)(C(I)^TC(I))^{-1}C(I)^T\vec{r}, \quad C(I) = P + IQ.$$

Since $C(I)$ is a linear function of I, g reduces to the ratio of two polynomials. Since $C(I)^TC(I)$ is 3×3, g is typically the ratio of two polynomials of degree 6. One method for determining the value of I, say $I_{\max}$, that maximizes g is the following: Among the roots (typically 10 altogether) of $g'(I) = 0$, $I_{\max}$ is the real root for which the corresponding value of g is largest. The roots of the 10th degree polynomial are easily computed using the Jenkins-Traub algorithm as implemented by *Hager* [1987]. Since the derivative of a function vanishes at a local maximum, this procedure for finding the maximum value of g will work if and only if the maximum value of g is not attained as I tends to $+\infty$ or to $-\infty$. Although one can construct examples where the maximum of g is attained as I tends to $+\infty$ or to $-\infty$, it is highly unlikely that this situation will occur in practice. After determining the current $I_{\max}$ associated with the least squares error, the corresponding estimate for the velocity is

$$\vec{v} = (C(I_{\max})^TC(I_{\max}))^{-1}C(I_{\max})^T\vec{r}. \tag{13}$$

When the measured electric field is polluted with noise, the following idea can be used to improve the estimate for the velocity: Instead of evaluating (9) at the time $t = \delta t$, we integrate (9) over the interval $[0, \delta t]$. If the noise has zero mean, then the error in the integral of the electric field is often much smaller than the error in the value of the electric field at a specific instant of time. The calculation proceeds exactly as described above; the end result is an estimate for the integral of I coupled with a more accurate estimate for the velocity itself.

Even after taking into account noise in the data, much care is needed when using this procedure to estimate the propagation velocity and the initial current. To illustrate how totally erroneous results can be obtained from high quality data, and to explain more precisely how to interpret this least squares estimate for the velocity, let us consider a lightning channel that starts at the location $x = 1000$ m, $y = 1000$ m, and $z = 7000$ m relative to the Kennedy Space Center grid described in section 2. Suppose that the velocity of propagation is 0.5 c, the channel points in the direction of the unit vector (1,2,2)/3, and the current profile (for $t \geq 0$) is

$$I(t) = e^{-at} - e^{-bt}, \quad a = 2\times10^4, \quad b = 3.5\times10^6. \tag{14}$$

This is the same current profile studied by *Le Vine and Willett* [1992]. We consider the time interval $[0, 10^{-6}]$ during which time the pulse will travel a distance of 150 m. We suppose that the channel is longer than 150 m so that the $\Phi(L)$ term in (2) is zero on the time interval $[0, 10^{-6}]$. Using the formula (2), we evaluate the electric field at each station. We then integrate the electric field over the interval $[0, 10^{-6}]$ and estimate the propagation velocity using the least squares procedure described above. The integrated electric field at the five stations is a multiple of the following vector:

$$(1.432, -2.108, -3.822, -1.404, -1.278)$$

The computed least squares estimate for the velocity vector is

$$(0.5102, 1.0153, 1.0079) \times 10^8 \text{ m/s},$$

which agrees well with the actual velocity vector (0.5, 1.0, 1.0) $\times 10^8$ m/s: the error is due to the omission from (10) of the integral term in (2).

Now let us make tiny changes in the electric field values, and recompute the velocity. More precisely, we consider perturbations in the electric field values of the form $\pm k_i(0.01)$ where k_i is an integer, corresponding to station i, with values between -3 and 3. Altogether, there are $7^5 = 16807$ different perturbations of the electric field as the k_i take on all possible values in the range $[-3,3]$. We find that with 7250 of these values of the electric field, the error in the computed velocity is

more that 100%. In studying the computations, it can be seen that the huge error is due to the following phenomenon: The function g, that we maximize to obtain an estimate for the current, has many local maximizers, and as we change the electric field, these local maxima move around; for a small change in the electric field, a totally different local maximizer can produce the smallest least squares error, and this local maximizer can be far away from the local maximizer associated with the actual velocity. Despite the fact that the global solution to the least squares problem may be far from the desired physical value, we found that for every one of the 16807 perturbations in the electric field, there was always a local maximizer of g for which the associated velocity was within 10% accuracy of the true velocity. In other words, even though the global solution to the least squares problem can be far away from the desired physical value, there is still a local solution to the least squares problem that is near the correct physical value.

In order to determine which of these local maximizers of g is the physically correct one, we need to combine the procedure of this section with the procedure of the following section where we fit the model to the data over a long time interval, in contrast to the short-interval fit considered in this section. Using the long-interval fit, nonphysical local maximizers can be excluded. On the other hand, the local maxima obtained by the short-interval method provide excellent starting guesses for the long-interval approach. In summary, the least squares function g that we maximize in order to estimate the velocity has many local maximizers, and the local maximizer with the smallest least squares error may not be the one that best approximates the actual velocity. One needs to evaluate all the local maximizers of g and use them as starting guesses in the long-interval fit of the next section in order to exclude the nonphysical maxima.

4. Current Profile

In this section we consider the problem of fitting the model (2) to the observed electric field on the time interval $[0, T]$ where T is chosen large enough that the electric field and current are negligible for $t > T$. Throughout this section, it is assumed that the start of the channel is known (see section 2), and that we wish to determine the unknown channel end, propagation velocity, and current profile. Observe that if the channel end and the propagation velocity are known, then the unknown current profile enters the model (2) linearly. If the current profile is discretized, then (2) essentially provides a linear system of equations that can be solved for the unknown current profile. After solving for I, we can determine the associated least squares error, Ω, by subtracting the right side of (2) from the left side, squaring, and integrating over the interval $[0, T]$. This least squares error Ω is a function of L and $\vec{v}$. The channel end and propagation velocity are chosen to minimize $\Omega(L, \vec{v})$ over L and $\vec{v}$. There are a variety of methods that can be used to carry out this minimization [see *Hager et al.*, 1993]. In our numerical experiments, we simply used a coordinate descent technique where one cycles through the variables, holding all the variables fixed except for one variable which is minimized over; the minimization of this single variable was accomplished by a bisection type procedure [see *Hager and Ghosh*, 1990].

For our numerical experiments, we used a finite element discretization [*Strang and Fix*, 1973] of the current profile. More precisely, we approximated I by the finite series:

$$I(t) = \sum_{j=1}^{n} \gamma_j \phi_j(t),$$

Table 3. Estimates for Channel Length (L) and for Propagation Velocity ($\vec{v}$)

Description	L, m	$\vec{v}$, m/s
True values	1200.0	$(0.500, 1.000, 1.000) \times 10^8$
350 intervals	1188.7	$(0.507, 1.016, 1.010) \times 10^8$
700 intervals	1198.2	$(0.502, 1.005, 1.003) \times 10^8$
350 intervals, 5% noise	1188.2	$(0.511, 1.012, 1.008) \times 10^8$
350 intervals, 20% noise	1205.6	$(0.524, 1.024, 1.017) \times 10^8$

where the coefficients γ_j are scalars and the basis functions ϕ_j are piecewise linear functions that vanished outside the interval $[0, T]$. The typical basis function has unit height and the following shape: _____/_____ . Defining the function ψ_{ij} by

$$\psi_{ij}(t) = \Phi_i(L)\phi_j(t - \Delta t_i) - \Phi_i(0)\phi_j(t) - \Lambda_i(\phi_j)(t),$$

where $\Delta t_i = t_{b_i} + t_c - t_{a_i}$, the coefficients γ_j are obtained by solving the linear system $A\vec{\gamma} = \vec{b}$ where

$$b_j = \sum_{i=1}^{m} \langle \psi_{ij}, E_i \rangle \quad \text{and} \quad a_{jk} = \sum_{i=1}^{m} \langle \psi_{ij}, \psi_{ik} \rangle.$$

Here $\langle f, g \rangle$ stands for the integral of $f(t)g(t)$ over the interval $[0, T]$.

As a small illustration, suppose that lightning occurs within the Kennedy Space Center grid described in section 2, with channel starting at $x = 1000$ m, $y = 1000$ m, and $z = 7000$ m, and ending at $x = 1400$ m, $y = 1800$ m, and $z = 7800$ m. Thus the length of the channel is 1.2 km. In addition, suppose that the velocity of propagation is 0.5 c and the current profile is given by (14). We evaluate the electric field at Thomson's stations using the transmission line model (2). Using these values for the electric field, we reconstruct the location of the lightning channel, the propagation velocity, and the current profile by minimizing the least squares error $\Omega(L, \vec{v})$.

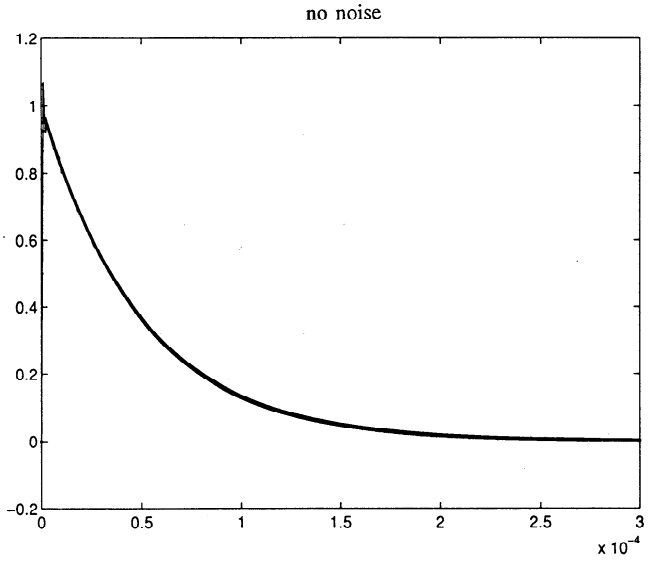

Figure 2. Computed and actual current profile, 350 basis functions.

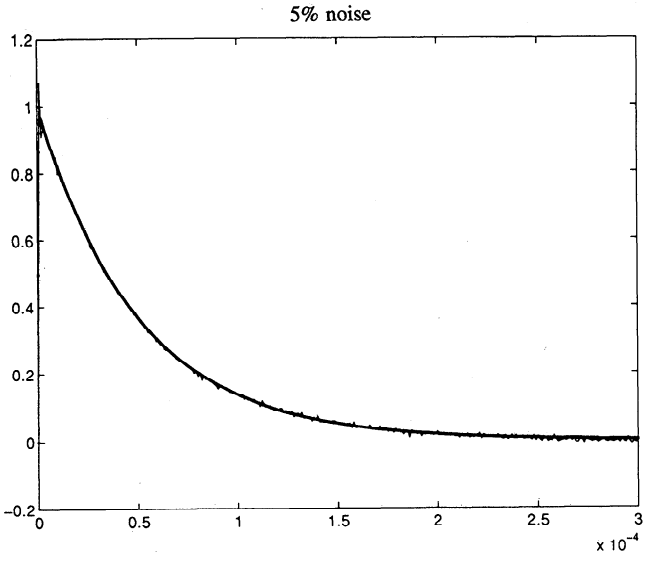

Figure 3. Computed and actual current profile, 5% noise, 350 basis functions.

The time interval $[0, T]$ that we consider in the least squares approximation is [0, .0004 s], and the current profile I is approximated using 350 piecewise linear, continuous basis functions. The least squares fit to the data obtained by minimizing the function $\Omega(L, \vec{v})$ appears in Table 3. The errors in the channel length and the propagation velocity are related to the discretization of the current profile and to the truncation time $T = 0.0004$ s. The true current can only be approximated by piecewise linear functions and the true current is still slightly positive for $t > 0.0004$ s. In our case, most of the error is related to the current discretization. If the time interval is partitioned into 700 subintervals instead of 350 subintervals, we obtain the improved estimates for L and for $\vec{v}$ shown in Table 3. As seen in Figure 2, the computed current profile also agrees quite well with the actual current profile. The two curves agree to within the width of a line except near $t = 0$ where the computed current overshoots slightly. When the number of basis functions is increased from 350 to 700, the overshoot disappears completely.

To model noisy data, we made random perturbations in the sampled electric fields at each station where the magnitude of the perturbation was bounded by some fraction (either 0.05 or 0.20) times the maximum electric field detected at all the stations during the time interval $[0, T]$. In Figures 3 and 4, we show the current profile computed with 5% and 20% random noise. The corresponding estimates for the channel length and for the propagation velocity also appear in Table 3. Observe that even with 20% noise, the errors in the channel length and propagation velocity are still only a few percent.

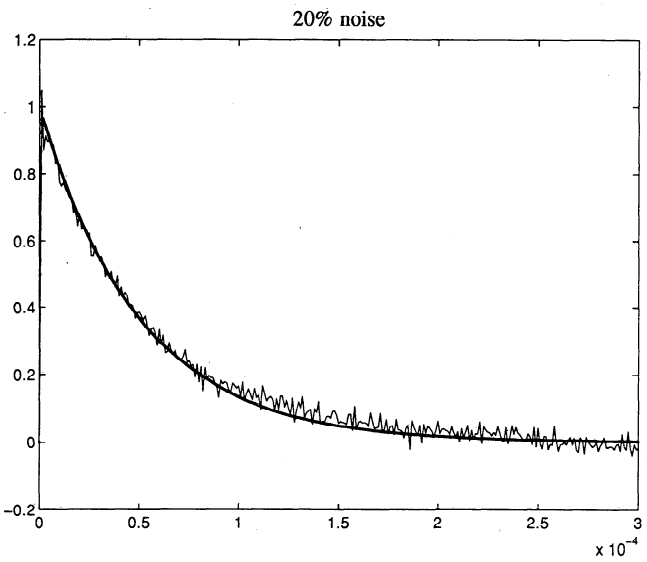

Figure 4. Computed and actual current profile, 20%% noise, 350 basis functions.

Appendix: An Observed Pulse

In this appendix we examine the pulse reported in Figure 8 of [*Thomson et al.*, 1994]. Part (a) of this figure is reprinted in Figure A1 here. Observe that Thomson presents normalized values of the derivative electric field. Differentiating the transmission line equation, we see that the derivative electric field is related to the derivative current. Table A1 gives the relative arrival times for the pulse of Figure A1. Using the nonlinear least squares algorithm of section 2, we obtain the following coordinates for the location (meters) of the start of the pulse relative to KSC coordinates:

$$\begin{aligned} x &= 8918 \pm 54, \\ y &= -13899 \pm 65, \\ z &= 5223 \pm 49. \end{aligned} \tag{A1}$$

The tolerances 54, 65, and 49 in (A1) denote the standard deviation associated with a 30 ns standard deviation in timing measurements. Column 3 of Table A1 gives the corresponding residuals e_i (see section 2). Observe that the 2-norm of the residual (that is the square root of the sum of the squares of the e_i) is about 4.0 m, while for a 30 ns timing error, the average 2-norm of the residual for a pulse emanating from the location (A1) is about 7.2 m. Hence if the residual corresponding to this pulse is typical, then the timing error may be closer to 20 ns rather than the 30 ns employed for the numerical experiments of section 2.

It turns out that the pulse of Figure A1 corresponds to a rather short channel, so to simplify the analysis, we work with the formula (1) describing a small channel segment. It follows from (1) that if $\bar{E}_i'$ denotes the normalized derivative vertical electric field at station i, we have

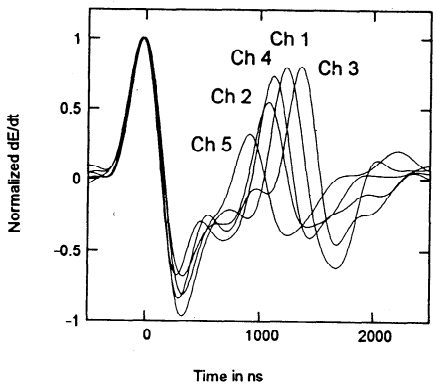

Figure A1. Normalized dE/dt as it appears in *Thomson et al.* [1994, Figure 8].

Table A1. Arrival Times (μs) and Residuals e_i (meters) for Each Station

Station	t_i	$\lvert e_i \rvert$
1	36.309375	2.9
2	63.624000	2.0
3	53.233625	0.6
4	9.222875	1.4
5	19.651000	1.1

$$\bar{E}_i' = \bar{I}'(t - \Delta t_i) - \bar{I}'(t), \tag{A2}$$

where $\bar{I}'$ is the normalized derivative current, and

$$\Delta t_i = t_{b_i} + t_c - t_{a_i}.$$

We will use (A2) to compute both the unknown $\bar{I}'$ and the time shift Δt_i. From the time shifts, we obtain the end of the channel, and by an integration of $\bar{I}'$, we obtain the normalized current profile.

To begin, let us examine finite difference relations of the form

$$y(t) - y(t-\Delta t) = z(t). \tag{A3}$$

Of course, y and z here play the role of $\bar{I}'$ and $-\bar{E}_i'$ in (A2). If $y(t) = 0$ for $t \le 0$, then y is completely determined by (A3). In particular, on the interval $[0, \Delta t]$, $y(t) = z(t)$; on the interval $[\Delta t, 2\Delta t]$, $y(t) = z(t) + z(t-\Delta t)$. And in general, we have

$$y(t) = \sum_{j=0}^{\lfloor t/\Delta t \rfloor} z(t-j\Delta t) \tag{A4}$$

where $\lfloor t/\Delta t \rfloor$ is the largest integer $\le t/\Delta t$.

For our physical problem, we know that $y(t)$ tends to zero as t tends to infinity, which gives us the condition:

$$\sum_{j=0}^{\infty} z(s + j\Delta t) = 0 \text{ for each } s \text{ between 0 and } \Delta t. \tag{A5}$$

The condition (A5) is gotten from (A4) by reversing the order of the sum, and by letting t tend to infinity. Since z is known, we can think of (A5) as a condition for ruling out unsuitable shifts Δt. Observe that if Δt satisfies (A5), then so does $\Delta t/2$. To see this, split the sum corresponding to $\Delta t/2$ into 2 separate sums associated with the even terms and with the odd terms; each of these separate sums vanishes from (A5). In general, if Δt is a solution of (A5), then so is $\Delta t/k$ for any integer k. The physically relevant Δt is the largest one satisfying (A5). (This can be established by taking Laplace transforms, by examining the situations where zeros cancel poles, and by observing that the Laplace transform of an absolutely integrable function does not have any poles on the imaginary axis).

Returning to the data of Figure A1, suppose that measurements are taken at the times $0 = t_0, t_1, t_2, \cdots$ where $t_{i+1} - t_i = \delta$ is the sampling time interval. Let us define the following function M:

$$M(\Delta t) = \sum_{0 \le t_i \le \Delta t} \frac{\delta}{\Delta t} \left| \sum_j f(t_i + j\Delta t) \right|,$$

where the index j of the inner sum is restricted so that $t_i + j\Delta t$ lies within the time window corresponding to the data. Motivated by (A5), we wish to determine the largest Δt where M is relatively small. The M's associated with the data of Figure A1 appear in Figure A2. The largest local minima are listed in the second column of Table A2.

Analogous to the procedure of section 4, we now consider a least squares problem:

$$\text{minimize} \quad \int_0^T \sum_{i=1}^{m} (\bar{I}'(t) - \bar{I}'(t-\Delta t_i) - \bar{E}_i')^2 dt.$$

This least squares problem is more complex than the one studied in section 4 since we now minimize both over I and over the time shifts Δt_i. Using the time shifts of Table A2 as a

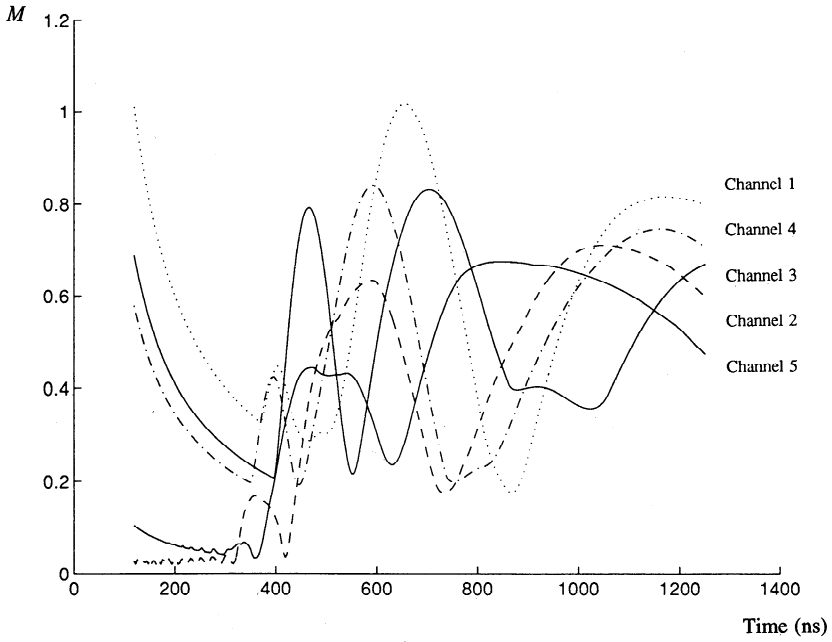

Figure A2. M associated with each station.

Table A2. Time Shifts (ns) and Residuals e_i (meters) for Each Station

Station Number	Local Minima of M	Δt_i from optimization	$\lvert e_i \rvert$
1	859	863.18	1.0
2	734	750.65	0.7
3	1022	952.34	0.2
4	753	790.14	0.5
5	631	639.54	0.4

starting guess, and applying the conjugate gradient code of [Hager, 1989], we obtain the shifts appearing in column 3 of Table A2.

Having determined the time shifts, let us consider the following equation:

$$t_{b_i} + t_c = \Delta t_i + t_{a_i}. \tag{A6}$$

The arrival time t_{a_i} is the time for light to travel from the location (A1), where the pulse starts, to each station; hence, the right side of equation (A6) is known. On the left side of (A6), the unknowns are t_{b_i}, the time for light to propagate from the end of the channel to each station and t_c, the time for the current pulse to propagate the length of the channel. Thus (A6) has precisely the same structure as the equation discussed in section 2. Applying the algorithm of section 2, we obtain the following location (meters) for the end of the lightning channel:

$$\begin{aligned} x &= x_0 - 54.2 \pm 10.9, \\ y &= y_0 - 102.9 \pm 13.2, \\ z &= z_0 + 23.5 \pm 9.8. \end{aligned} \tag{A7}$$

where (x_0, y_0, z_0) is the true location of the start of the lightning channel and $\Delta x = -54.2$, $\Delta y = -102.9$, and $\Delta z = 23.5$ is the relative translation from the true start to the end of the channel. The length of the channel is 118.7 ± 9.2 m. Our best estimate for the start of the channel and the associated standard deviation appear in (A1). The fourth column of Table A2 gives the residual error e_i of section 2. Observe that the residuals associated with the relative location of the channel end are much smaller than the residuals associated with the channel start. The standard deviations in the relative translation appearing in (A7) are based on the observed residuals e_i appearing in Table A2.

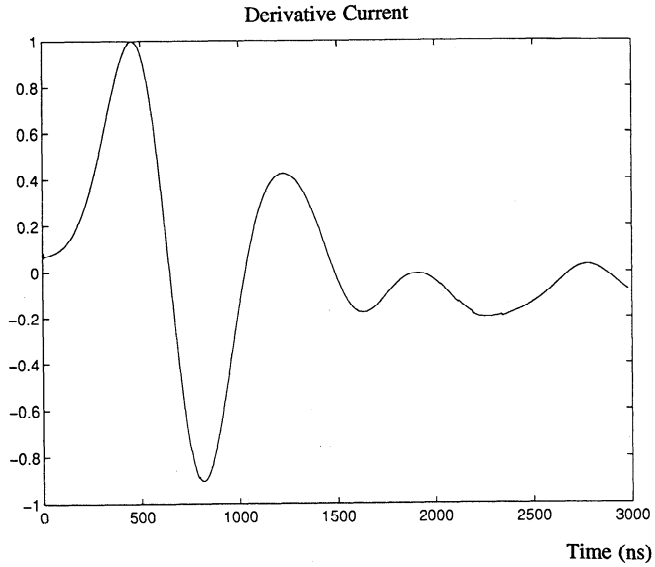

Figure A3. Derivative current profile.

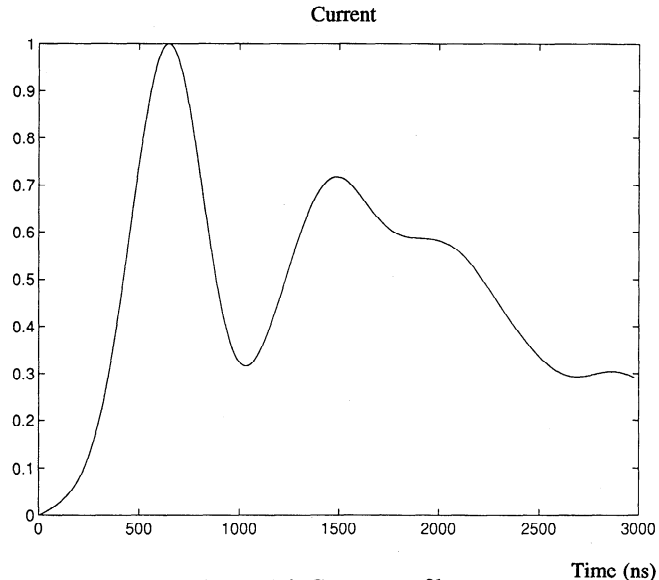

Figure A4. Current profile.

Now, we divide the channel length by the propagation time t_c, computed by the algorithm of section 2, to obtain the propagation velocity of the pulse:

$$\text{velocity} = 0.61 \text{ c}$$

Observe that the propagation velocity of this in-cloud pulse is significantly faster than the typical velocity measured for return strokes beneath clouds; see *Uman* [1987] for a discussion of return stroke speed along with references. For example *Idone and Orville* [1982] find that the first return stroke within 1.3 km of ground has mean velocity 0.32 c, while the mean speed of subsequent strokes is 0.4 c.

Finally, Figure A3 plots the profile for the derivative current, while Figure A4 plots the current itself for the 3 μs time window. The derivative current was obtained through the optimization process described above, while the current was obtained by integrating the derivative.

Acknowledgments. The author wishes to thank the reviewers for their careful reading of the manuscript and for their comments and suggestions. The author also greatfully acknowledges the support of NSF through a collaborative research project in the geosciences, bringing together a mathematician, W. W. Hager, and an experimentalist, E. M. Thomson. At the beginning of this project, E. M. Thomson asked W. W. Hager to develop algorithms to determine the peak current and the propagation velocity of return strokes by combining the transmission line model with the electric field measurements he would obtain from a system being installed at the Kennedy Space Center. The algorithms in this paper were developed in response to E. M. Thomson's request. Thomson's experimental system is described in the recent paper [*Thomson et al.*, 1994]. On March 14, 1995, E. M. Thomson provided W. W. Hager with digitized data corresponding to Figure 8 in [*Thomson et al.*, 1994]. This digitized data is used in the appendix to investigate some of the algorithms in this paper.

References

Hager, W. W., The generalized fast Fourier transform and the Jenkins-Traub algorithm, *Numerische Mathematik, 50*, 253–261, 1987.

Hager, W. W., *Applied Numerical Linear Algebra*, 424 pp., Prentice-Hall, Englewood Cliffs, N.J., 1988.

Hager, W. W., A derivative-based bracketing scheme for univariate minimization, *Comput. Math. Appl., 18*, 779–795, 1989.

Hager, W. W., and N. Ghosh, A derivative-free bracketing scheme for univariate minimization, *Comput. Math. Appl., 20*, 23–34, 1990.

Hager, W. W., J. S. Nisbet, J. R. Kasha, and W.-C. Shann, Simulations of electric fields within a thunderstorm, *J. Atmos. Sci., 46*, 3542–3558, 1989.

Hager, W. W., R. Horst, and P. M. Pardalos, Mathematical programming – a computational perspective, in *Handbook of Statistics, vol. 9, Computational Statistics*, edited by C. R. Rao, pp. 201–278, North-Holland, New York, 1993.

Idone, V. P., and R. E. Orville, Lightning return stroke velocities in the Thunderstorm Research International Program (TRIP), *J. Geophys. Res., 87*, 4903–4915, 1982.

Koshak, W. J., and H. J. Christian, A non-hyperbolic solution to the problem of mapping lightning sources using VHF time-of-arrival measurements, 1994 Fall AGU meeting, Abstract A21D-6.

Le Vine, D. M., and J. C. Willett, Comment on the transmission-line model for computing radiation from lightning, *J. Geophys. Res., 97*, 2601–2610, 1992.

Lewis, E. A., R. B. Harvey and J. E. Rasmussen, Hyperbolic direction finding with sferics of transatlantic origin, *J. Geophys. Res., 65*, 1879–1905, 1960.

Proctor, D. E., A hyperbolic system for obtaining VHF radio pictures of lightning, *J. Geophys. Res., 76*, 1478–1489, 1971.

Rachidi, F., and R. Thottappillil, Determination of lightning currents from far electromagnetic fields, *J. Geophys. Res., 98*, 18,315–18,321, 1993.

Strang, G., *Introduction to Applied Mathematics*, 758 pp., Wellesley-Cambridge Press, Wellesley, Mass., 1986.

Strang, G., and G. Fix, *An Analysis of the Finite Element Method*, 306 pp., Prentice-Hall, Englewood Cliffs, N.J., 1973.

Thomson, E. M., P. J. Medelius, and S. Davis, System for locating the sources of wideband dE/dt from lightning, *J. Geophys. Res., 99*, 22,793–22,802, 1994.

Thottappillil, R., and M. A. Uman, Comparison of lightning return-stroke models, *J. Geophys. Res., 98*, 22,903–22,914, 1993.

Thottappillil, R., and M. A. Uman, Lightning return stroke model with height-variable discharge time constant, *J. Geophys. Res., 99*, 22,773–22,780, 1994.

Uman, M. A., *The Lightning Discharge*, 377 pp., Harcourt Brace, Orlando, Fla., 1987.

Uman, M. A., and D. K. McLain, Magnetic field of lightning return stroke, *J. Geophys. Res., 74*, 6899–6910, 1969.

Uman, M. A., and D. K. McLain, Radiation field and current of the lightning stepped leader, *J. Geophys. Res., 75*, 1058–1066, 1970.

W. W. Hager and D. Wang, Department of Mathematics, University of Florida, Gainesville, FL 32611. (e-mail: hager@math.ufl.edu)

(Received February 2, 1995; revised August 10, 1995; accepted August 14, 1995.)