1995 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM
JOHN F. KENNEDY SPACE CENTER
UNIVERSITY OF CENTRAL FLORIDA

A STUDY OF VARIOUS METHODS FOR CALCULATING LOCATIONS OF LIGHTNING EVENTS

Dr. John R. Cannon
Professor and Chair
Mathematics Department
University of Central Florida
Orlando, Florida

KSC Colleague - Carl L. Lennon
Atmospheric Science

Contract Number NASA-NGT-60002
Supplement 19

August 22, 1995

Acknowledgments

The author wishes to thank Carl Lennon, Lanna Maier, John Madura, Ray Hosler, Kari Stiles and the management of the NASA's Kennedy Space Center for the opportunity to spend an enjoyable summer conducting research on the LDAR equations. Also, the author wishes to thank William Koshak of NASA's Marshall Space Flight Center for several interesting private communications. Thank you and best wishes to all.

Abstract

This article reports on the results of numerical experiments with various techniques such as linear least squares, nonlinear least squares, statistical estimates, cluster analysis and angular filters for calculating the location of lightning events from time of arrival data for the Lightning Detection and Ranging (LDAR) System at Kennedy Space Center.

Summary

This article reports on the results of numerical experiments on finding the location of lightning events using different numerical methods. The methods include linear least squares, nonlinear least squares, statistical estimations, cluster analysis and angular filters and combinations of such techniques. The experiments involved investigations of methods for excluding fake solutions which are solutions that appear to be reasonable but are in fact several kilometers distant from the actual location. Some of the conclusions derived from the study are that bad data produces fakes, that no fool-proof method of excluding fakes was found, that a short base-line interferometer under development at Kennedy Space Center to measure the direction cosines of an event shows promise as a filter for excluding fakes. The experiments generated a number of open questions, some of which are discussed at the end of the report.

Table of Contents

I. Introduction

The Lightning Detection and Ranging (LDAR) System developed over a number of years by Carl Lennon and colleagues at the Kennedy Space Center is based on the times of arrival at seven locations of electromagnetic radiation in the 66MH frequency range emitted by electric charge movement in thunder storms. The differences in the times of arrivals are converted into differences in distances from the point of origin (x, y, z) of the radiation to the various receiving sites $\# i = 0, \cdots, 6$ located at $(x_i, y_i, 0)$, $i = 0, \cdots, 6$, where site #0 is taken to be the origin $(0, 0, 0)$ of the local rectangular coordinate system with the positive x-axis directed east, the positive y-axis directed north and the positive z-axis directed up. The geographic locations are shown in Figure 1 along with their latitude and longitude coordinates. Relative to the local coordinate system the site locations in kilometers from Site #0 are

	x_i	y_i	z_i
Site #0	0	0	0
Site #1	3.255	9.462	0
Site #2	7.466	.014	0
Site #3	5.532	-7.056	0
Site #4	-3.854	-5.792	0
Site #5	-8.424	-1.007	0
Site #6	-3.738	7.460	0

Remark: The origin $(0, 0, 0)$ is located at the top of the antenna at Site #0 and the heights of the antennas at the other sites have been adjusted taking the curvature of the earth into account to be nearly in the xy-plane. A recent survey has shown the various antenna to be within a meter or so of the xy-plane.

Denoting the time of arrival of the signal at Site $\# i$ by t_i, $i = 0, 1, \cdots, 6$, and the distance from the point of origin of the signal (x, y, z) to $(x_i, y_i, 0)$ (Site $\# i$)by

(1) $$d_i = \sqrt{(x - x_i)^2 + (y - y_i)^2 + z^2},$$

we see that

(2) $$d_i - d_j = u_{ij}, \qquad i, j = 0, \cdots, 6, i \neq j,$$

where

$$u_{ij} = c(t_i - t_j) \tag{3}$$

and

$$c = 299792.458 \text{ km/sec} \tag{4}$$

is the measured speed of light. Current practice is to use Site #0 as the trigger for the initiation of a time of arrival measurement at all sites. Transmission times to and from the other sites are taken into account in the determination of the $t_i, i = 1, \cdots, 6$. From these time measurements the six equations

$$d_i - d_0 = u_{i0}, \qquad i = 1, \cdots, 6 \tag{5}$$

which are solved for x, y and z. The equations appear to be nonlinear at first glance. However, setting

$$r = d_0 = \sqrt{x^2 + y^2 + z^2} \tag{6}$$

we find that

$$d_i = \sqrt{r^2 + r_i^2 - 2x_i x - 2y_i y} \tag{7}$$

where

$$r_i^2 = x_i^2 + y_i^2 . \tag{8}$$

Shifting d_0 to the right side of (5), squaring both sides, and employing (6), (7) and (8), we find after some elementary calculations that (x, y, z) satisfies

$$x_i\, x + y_i\, y + u_{i0} r = v_i, \qquad i = 1, \cdots, 6 \tag{9}$$

where z is contained implicitly in the unknown r and

(10) $$v_i = \left(r_i^2 - u_{io}^2\right)/2, \qquad i = 1, \cdots, 6.$$

We note here that equations (2) for $i, j \neq 0$ do not admit this simplification and remain nonlinear in x, y and r. The pairs of three equations $i = 1, 3, 5$ and $i = 2, 4, 6$ are solved for x, y and r with z determined from

(11) $$z = \sqrt{r^2 - x^2 - y^2}$$

Let $\left(x_{\text{odd}}, y_{\text{odd}}, z_{\text{odd}}\right)$ and $\left(x_{\text{even}}, y_{\text{even}}, z_{\text{even}}\right)$ denote the respective solutions.

If

(12) $$\begin{aligned} \left|x_{\text{odd}} - x_{\text{even}}\right| &< \eta, \\ \left|y_{\text{odd}} - y_{\text{even}}\right| &< \eta, \\ \left|z_{\text{odd}} - z_{\text{even}}\right| &< \eta, \end{aligned}$$

where

(13) $$\eta = \begin{cases} .35\,\text{km if } R < 7\,\text{km} \\ (.05)\,R\,\text{km if } R \geq 7\,\text{km} \end{cases}$$

and

(14) $$R = \left(x_{\text{odd}}^2 + y_{\text{odd}}^2\right)^{1/2},$$

then

(15) $$\begin{aligned} x &= \left(x_{\text{odd}} + x_{\text{even}}\right)/2, \\ y &= \left(y_{\text{odd}} + y_{\text{even}}\right)/2, \\ z &= \left(z_{\text{odd}} + z_{\text{even}}\right)/2, \end{aligned}$$

On the other hand if (12) is not satisfied, then each of the 20 combinations of three independent equations from the six equations in (9) are solved and each solution is compared with the others to find at least eight solutions which satisfy the tolerance limits in (12) with R in (14) calculated from the xy−coordinates of the solution used as the base for each comparison. Preference is given to the odd and even sites by counting them twice when they are used as the comparison base. If 8 such solutions are found to satisfy (12), then the solution (x, y, z) is determined from their average. If 8 solutions satisfying (12) cannot be found, then no solution is determined and the system proceeds to the next event. The choice of 8 stems from the combinatoric possibility of 10 good solutions if one site is down, etc. Currently, solutions are determined for 20 to 50 percent of the events observed by the LDAR system. We should note here that the even and odd combinations are nearly optimally configured to minimize the amplification of measurement errors which explains their preference in the current process. We shall expand this brief explanation in our discussion on errors in section II. To illustrate the problem the reader is referred to Figure 2 which shows the xy−plot of the ~42 solutions determined from synthetic data subjected to the addition of error samples from a uniform distribution of magnitude + or - 750 meters.

As noted in the abstract the purpose of this report is to discuss the results of numerous computer simulations which were carried out from April through August 1995. We begin with a discuss of errors in Section II and proceed to a discussion of methods considered in Section III. In Section IV, we focus on the utility of additional angular information that may become available via measurements from a short base line interferometer which is currently under development for location at the central Site #0. In Section V, we summarize conclusions and open questions for future study.

II. Measurement Errors

We begin with the system (9) which we rewrite as

$$x_i\, x + y_i\, y + u_i\, r = v_i, \qquad i = 1, \cdots, 6, \tag{16}$$

where here and in all that follows we set

$$u_i = u_{i0}. \tag{17}$$

Given any three equations $i, j, k, i \neq j \neq k,$ we can write

$$A_{ijk}\, \bar{s}_{ijk} = \bar{b}_{ijk}, \tag{18}$$

where

(19) $$A_{ijk} = \begin{pmatrix} x_i & y_i & u_i \\ x_j & y_j & u_j \\ x_k & y_k & u_k \end{pmatrix},$$

(20) $$\vec{s}_{ijk} = \begin{pmatrix} x_{ijk} \\ y_{ijk} \\ r_{ijk} \end{pmatrix},$$

and

(21) $$b_{ijk} = \begin{pmatrix} v_i \\ v_j \\ v_k \end{pmatrix}.$$

Note that we have displayed the dependence of $x, y,$ and r on i, j, k via subscripts. For perfect data, the x_{ijk}, y_{ijk}, and r_{ijk} are all equal respectively. Suppose that u_i is replaced by $u_i + \varepsilon_i$ and that v_i is replaced by $v_i - u_i\varepsilon_i - \left(\varepsilon_i^2/2\right)$ in (16) above. Then, the system (18) can be written as

(22) $$A_{ijk}\, \vec{s}_{ijk}\left(\varepsilon_i, \varepsilon_j, \varepsilon_k\right) = \vec{c}_{ijk}\left(\varepsilon_i, \varepsilon_j, \varepsilon_k\right)$$

where the solution $\vec{s}_{ijk}$ denoted in (20) above has its dependence on the errors $\varepsilon_i, \varepsilon_j,$ and ε_k displayed in (22) and

(23) $$\vec{c}_{ijk} = \vec{b}_{ijk} + \vec{f}_{ijk}\left(\varepsilon_i, \varepsilon_j, \varepsilon_k, u_i, u_j, u_k, r\right),$$

where

(24) $$\vec{f}_{ijk} = -\begin{pmatrix} \varepsilon_i r + u_i\varepsilon_i + \left(\varepsilon_i^2/2\right) \\ \varepsilon_j r + u_j\varepsilon_j + \left(\varepsilon_j^2/2\right) \\ \varepsilon_k r + u_k\varepsilon_k + \left(\varepsilon_k^2/2\right) \end{pmatrix}.$$

Subtracting (18) from (22), we see that

(25) $$\vec{s}_{ijk}(\varepsilon_i, \varepsilon_j, \varepsilon_k) = \vec{s}_{ijk}(0,0,0) + A_{ijk}^{-1} \vec{f}_{ijk}.$$

We note two things. First, for fixed measurement errors, the overall solution error

(26) $$\vec{s}_{ijk}(\varepsilon_i, \varepsilon_j, \varepsilon_k) - \vec{s}_{ijk}(0,0,0) = O(r),$$

where the Landau Order symbol is employed here:

(27) $$g = O(r) \quad \text{iff} \quad |g| \le Cr$$

where C is a positive constant, i.e.,

$$x_{ijk}(\varepsilon_i, \varepsilon_j, \varepsilon_k) - x_{ijk}(0,0,0) = O(r_{ijk}(0,0,0))$$
$$y_{ijk}(\varepsilon_i, \varepsilon_j, \varepsilon_k) - y_{ijk}(0,0,0) = O(r_{ijk}(0,0,0))$$
$$r_{ijk}(\varepsilon_i, \varepsilon_j, \varepsilon_k) - r_{ijk}(0,0,0) = O(r_{ijk}(0,0,0))$$

This effect of the increase in error with the distance of the event can be observed computationally. Second, the analysis of relative error in solutions of linear systems depends on a factor called the condition number of the matrix; i.e.,

(28) $$\frac{\|\vec{s}_{ijk}(\varepsilon_i, \varepsilon_j, \varepsilon_k) - \vec{s}_{ijk}(0,0,0)\|}{\|\vec{s}_{ijk}(0,0,0)\|} \le K(A_{ijk}) \frac{\|\vec{c}_{ijk} - \vec{b}_{ijk}\|}{\|b_{ijk}\|},$$

where $\| \;\|$ denotes some vector norm such as the maximum of the absolute values of the components,

(29) $$K(A_{ijk}) = \|A_{ijk}\| \, \|A_{ijk}^{-1}\|$$

is the condition number of the matrix A_{ijk}, and $\|A_{ijk}\|$ and $\|A_{ijk}^{-1}\|$ are the matrix norms induced by the vector norm $\| \;\|$ via the relation

(30) $$\|A_{ijk}\| = \max_{|\vec{s}|=1} \|A_{ijk}\vec{s}\|.$$

For derivations and discussions of vector norms, matrix norms, condition number, and relative error, see [18] or any other reasonable text on numerical analysis or numerical linear algebra. For the vector norm defined as the maximum absolute value of the components of the vector, the corresponding matrix norm is the maximum of the row sums of the absolute values of the entries in each row; i.e., for the matrix $A = (a_{ij})$

(31) $$\|A\| = \max_i \left\{ \sum_{j=1}^{n} |a_{ij}| \right\}$$

is induced by the vector norm

(32) $$\|\vec{s}\| = \max_i |s_i|,$$

where s_i is the i^{th} component of a vector

(33) $$\vec{s} = \begin{pmatrix} s_i \\ s_2 \\ \cdot \\ \cdot \\ \cdot \\ s_n \end{pmatrix}.$$

The point here is that the matrix A_{ijk} depends on the measurements u_i, u_j, and u_k which depend in turn upon (x, y, z). Thus, as (x, y, z) vary with fixed r, the condition number $K(A_{ijk})$ varies. See Tables 1-4. The "Y-configurations" of the odd and even numbered sites is an approximation to the optimal perfect Y-configuration which minimizes the maximum of the condition number over the $x_i, x_j, x_k, y_i, y_j, y_k, u_i, u_j$, and u_k for fixed r. Taken together the increase in error with respect to r and the variation in the condition number of A_{ijk} as (x, y, z) vary is called the Geometric Dilution of Precision (GDOP).

While GDOP is a problem for (x, y, r), it is a major problem for z. Consider the expansion of $z_{ijk}(\varepsilon_i, \varepsilon_j, \varepsilon_k)$ in its Taylor series:

$$z_{ijk}(\varepsilon_i,\varepsilon_j,\varepsilon_k)=z_{ijk}(0,0,0)+\frac{\partial z_{ijk}}{\partial\varepsilon_i}(0,0,0)\,\varepsilon_i+$$

(34)
$$\begin{aligned}&+\frac{\partial z_{ijk}}{\partial\varepsilon_j}(0,0,0)\,\varepsilon_j+\frac{\partial z_{ijk}}{\partial\varepsilon_k}(0,0,0)\,\varepsilon_k+\\&+\frac{1}{2}\frac{\partial^2_{z_{ijk}}}{\partial\varepsilon_i^2}(0,0,0)\,\varepsilon_i^2+\frac{1}{2}\frac{\partial^2_{z_{ijk}}}{\partial\varepsilon_j^2}(0,0,0)\,\varepsilon_j^z+\\&+\frac{1}{2}\frac{\partial^2_{z_{ijk}}}{\partial\varepsilon_k^2}(0,0,0)\,\varepsilon_k^2+\frac{\partial^2_{z_{ijk}}}{\partial\varepsilon_i\partial\varepsilon_j}(0,0,0)\,\varepsilon_i\,\varepsilon_j\\&+\frac{\partial^2_{z_{ijk}}}{\partial\varepsilon_i\partial\varepsilon_k}(0,0,0)\,\varepsilon_i\,\varepsilon_k+\frac{\partial^2_{z_{ijk}}}{\partial\varepsilon_j\partial\varepsilon_k}(0,0,0)\,\varepsilon_j\,\varepsilon_k+\\&+\cdots.\end{aligned}$$

As

(35)
$$z_{ijk}(\varepsilon_i,\varepsilon_j,\varepsilon_k)=\sqrt{r(\varepsilon_i,\varepsilon_j,\varepsilon_k)^2-x(\varepsilon_i,\varepsilon_j,\varepsilon_k)^2-y(\varepsilon_i,\varepsilon_j,\varepsilon_k)^2}$$

it follows that

(36)
$$\frac{\partial z_{ijk}}{\partial\varepsilon_i}=\left(r_{ijk}\frac{\partial r_{ijk}}{\partial\varepsilon_i}-x_{ijk}\frac{\partial x_{ijk}}{\partial\varepsilon_i}-y_{ijk}\frac{\partial y_{ijk}}{\partial\varepsilon_i}\right)\Big/z_{ijk}$$

and

(37)
$$\begin{aligned}\frac{\partial^2 z_{ijk}}{\partial\varepsilon_i\,\partial\varepsilon_j}=&-\frac{1}{z_{ijk}^3}\left(r_{ijk}\frac{\partial r_{ijk}}{\partial\varepsilon_j}-x_{ijk}\frac{\partial x_{ijk}}{\partial\varepsilon_j}-y_{ijk}\frac{\partial y_{ijk}}{\partial\varepsilon_j}\right)\bullet\left(r_{ijk}\frac{\partial r_{ijk}}{\partial\varepsilon_i}-x_{ijk}\frac{\partial x_{ijk}}{\partial\varepsilon_i}-y_{ijk}\frac{\partial y_{ijk}}{\partial\varepsilon_i}\right)+\\&+\frac{1}{z_{ijk}}\left(\frac{\partial r_{ijk}}{\partial\varepsilon_j}\frac{\partial r_{ijk}}{\partial\varepsilon_i}-\frac{\partial x_{ijk}}{\partial\varepsilon_j}\frac{\partial x_{ijk}}{\partial\varepsilon_i}-\frac{\partial y_{ijk}}{\partial\varepsilon_j}\frac{\partial y_{ijk}}{\partial\varepsilon_i}\right)+\\&+\frac{1}{z_{ijk}}\left(r_{ijk}\frac{\partial^2 r_{ijk}}{\partial\varepsilon_j\,\partial\varepsilon_i}-x_{ijk}\frac{\partial^2 x_{ijk}}{\partial\varepsilon_i\,\partial\varepsilon_j}-y_{ijk}\frac{\partial^2 y_{ijk}}{\partial\varepsilon_i\,\partial\varepsilon_j}\right)\quad\text{, etc.}\end{aligned}$$

From spherical coordinates, we recall that $z=r\sin\phi$ where ϕ is the angle of elevation out of the xy-plane. By differentiating the equations (22) with respect to $\varepsilon_i,\varepsilon_j,\varepsilon_k$, we observe (Not Proved-although a messy induction argument looks feasible) that

$$\frac{\partial^n x_{ijk}}{\partial \varepsilon_i^{\,l} \partial \varepsilon_j^{\,n-l}} = O\left(r_{ijk}^n\right) \tag{38}$$

and similarly for y_{ijk} and r_{ijk}. Consequently, the expansion for z_{ijk} behaves like

$$\begin{aligned} z_{ijk}\left(\varepsilon_i, \varepsilon_j, \varepsilon_k\right) = z_{ijk}(0,0,0) + \\ O\Bigg(r_{ijk}(0,0,0)\Bigg[\frac{1}{\sin\phi}\left\{\varepsilon + \varepsilon^2 + \cdots\right\} + \\ \frac{1}{\sin^3\phi}\left\{\varepsilon^2 + \varepsilon^2 + \cdots\right\} + \cdots\Bigg]\Bigg), \end{aligned} \tag{39}$$

where

$$\varepsilon = \max\left\{\left|\varepsilon_i\right|, \left|\varepsilon_j\right|, \left|\varepsilon_k\right|\right\} \tag{40}$$

has been employed to simplify the expression in (39). This behavior in the error in z can be seen computationally by fixing z and letting r increase. For fixed z as r increases ϕ and $\sin\phi$ tend to zero and together r and ϕ rapidly destroy the accuracy in z while one can still obtain reasonable estimations of x and y.

Before we turn to a discussion of the methods explored we shall conclude with a few comments on the data error versus the current solution method at Kennedy Space Center. First, the calculation of η in (13) is a recognition GDOP and builds an approximation of the error behavior into what the system accepts as a solution. Next, the second phase of the current solution practice is the recognition of down time of a single site for maintenance and repairs. Taken together both phases impose a strict filter on the data acceptable for a solution. As seen by the scatter of solution points in figures 2-9 and discussed above, the system of equations (16) is very sensitive to measurement error magnitude, distance of the event, and the elevation angle. Thus, it is not unreasonable to impose such a filter as bad data generates bad answers and the electronics is subject to various interference sources that corrupt various portions of the data. Given the sensitivity of the system to errors, it is feasible that an error in the data could occur which would fake an acceptable solution to the system. It would be interesting to investigate the various possible types of errors in the system which could produce such fakes and whether or not the error would be an occasional fake or a frequent one. We shall comment on this again.

III. Methods Considered

Many Combinations and permutations of the following general methods were considered:

A. Linear Least Squares for x, y and r.
B. Linear Programming for x, y and r.
C. Nonlinear Least Squares for $u_i, i = 1, \cdots, 6$.
D. Nonlinear Least Squares for x, y and r.
E. Statistical Estimation for x, y and r.
F. Cluster Analysis for x, y and r.

We shall discuss each of the above in turn. However, it is worth pointing out first that all of the above methods work well for good data. When one encounters bad data all methods considered fall apart. The study focused upon the possible mix of "good" data from four of the outer sites combined with the central site with "bad" data from two of the outer sites. Another way of looking at this would be the determination of x, y and z from the best four of the measurements $u_i, i = 1, \cdots, 6$. With these comments in mind, we turn to the discussion of the above methods.

A. Linear Least Squares. Linear least squares for all six equations in (16) can be written as

(41) $$\mu = \min_{x,y,r} s(x, y, r),$$

where

(42) $$S(x, y, r) = \sum_{i=1}^{6} (x_i\, x + y_i\, y + u_i\, r - v_i)^2 .$$

Linear least squares also gives the possibility of solving for an (x, y, r) using any combination of three, four, or five equations. It was observed that solutions were always generated. An example of a reasonable approach to finding a solution from the best four of the sites is to find the four sites for which the corresponding μ is the least of the 15 possible μ's obtained from considering all combinations of four equations from the six equations in (16). The (x, y, z) resulting from that best combination could be taken as the approximation to the lightning event location. In trials with synthetic data with addition of errors selected from uniform distributions ranging up to

$\pm.5$ kilometer one could always find an error combination which would produce a fake solution, i.e., the solution corresponding to the least μ would be several kilometers away from the "correct" known synthetic solution.

B. Linear Programming. The problem of minimizing the linear function

$$L(x, y, r, \eta) = \eta \tag{43}$$

subject to the constraints

$$|x_i\ x + y_i\ y + u_i\ r - v_i| \leq \eta \tag{44}$$

is a feasible linear programming problem. However, employing the simplex method with slack variables introduced to convert the constraint in equalities into equations, we could not eliminate all of the slack variables from the solution of the programming problem. Hence, we had to resort to a brute force approach of examining all solutions of combinations of three equations selected from the six equations in (16). For each such solution (x, y, r), we defined the residuals

$$\eta_i = |x_i\ x + y_i\ y + u_i\ r - v_i|, \qquad i = 1, \cdots, 6. \tag{45}$$

Then we ordered the residuals

$$w_j = \eta_{i_j} \tag{46}$$

so that

$$w_6 \leq w_5 \leq w_4 \leq w_3 \leq w_2 \leq w_1.$$

Note that $w_4 \approx w_5 \approx w_6 \approx 0$ to machine accuracy via solution of three of the equations. We considered the resulting three minimization problems. The

$$\min w_1 \tag{47}$$

yields the solution (x, y, r) to the linear programming problem (43)-(44),

(48) $$\min w_2$$

yields the solution (x, y, r) corresponding to the best fit of five of the six equations in (16)

and

(49) $$\min w_3$$

yields the solution (x, y, r) corresponding to the best fit of four of the six equations in (16). As noted above fake solutions were observed for the solutions corresponding to the minimization problems (47), (48) and (49).

C. Nonlinear Least Squares for $\bar{u}$. Considering the least squares problem (41), we considered $x = x(\bar{u})$, $y = y(\bar{u})$ and $r = r(\bar{u})$ as $\bar{u} = (u_1, u_2, u_3, u_4, u_5, u_6)$ is the basic input for the calculation of (x, y, r). We can state the problem as

(50) $$\zeta = \min_{\bar{u}} F(\bar{u}),$$

where

(51) $$F(\bar{u}) = S(x(\bar{u}), y(\bar{u}), r(\bar{u}))$$

and $(x(\bar{u}), y(\bar{u}), r(\bar{u}))$ is the solution of (41)-(42) noting that $v_i = v_i(u_i)$. In effect this can be viewed as modifying the measured data $\bar{u}$ to a nearby consistent data set. The measured data set $\bar{u}$ was taken as the input to a Newton's method iteration for the solution of the equations

(52) $$\frac{\partial F}{\partial u_i}(\bar{u}) = 0, \quad i = 1, \cdots, 6,$$

when the data is good this method works. Shifting to the question of using the best four at the $u_i, i = 1, \cdots, 6$, (50)-(52) can be modified for four of the u_i and coupled to any reasonable linear least squares or linear programming method which produces a set of u_i, u_j, u_k, u_l associated with the optimal (x, y, r) that can be taken as the initial u's for the Newton method iteration for the solution of the modified nonlinear least squares problem. When the solution to the linear method was reasonable and the Newton iteration converged, the calculation cost for the Newton method did not yield much of a gain in solution accuracy. Sometimes the linear solution was reasonable and the Newton method failed to converge. Gradient procedures to find a good starting set of u's were tried and discarded as were simple search methods. Occasionally, the Newton method would converge to a fake solution especially when the linear procedure produced one.

D. Nonlinear Least Squares for x, y and r. We recall equation (5) and consider the problem

$$v = \min_{x,y,r} G(x, y, y) \tag{53}$$

where

$$G(x, y, r) = \sum_{i=1}^{6} \left(\sqrt{r^2 + r_i^2 - 2x_i x - 2y_{iy}} - r - u_i \right)^2 \tag{54}$$

and its modifications using four and five of the measured data $u_i, i = 1, 2, \cdots, 6$. As in **C.** above, a Newton's method iteration is required to solve the equations

$$\begin{aligned} \frac{\partial G}{\partial x}(x, y, r) &= 0, \\ \frac{\partial G}{\partial y}(x, y, r) &= 0, \\ \frac{\partial G}{\partial r}(x, y, r) &= 0, \end{aligned} \tag{55}$$

and an initial reasonable (x, y, r) is needed to start the Newton iteration. The starting (x, y, r) can be obtained from any of the linear processes mentioned above. As before good data yielded convergence to a good solution, but the improvement in the solution was not significant enough to justify the computational overhead of the Newton method. An attempt was made to consider the method (53)-(55) as a filter for identifying bad data and the exclusion of fake solutions. However, convergence to fake solutions of the nonlinear least squares was observed. So, divergence of the process cannot be taken as a filter for fake solutions.

E. Statistical Estimations. Lots of solutions can be generated when the data $u_i, i = 1, \cdots, 6$, contain error. For example, three equations from six yields 20 different solutions, four equations from six yields via least squares 15 more different solutions, five from 6 yields 6 more and least squares for all six yields another for a total of 42 different solutions which can be seen to coalesce into the correct (x, y, r) as the error in the data tends to zero. There are many ways to generate an infinite supply of different solutions when the measurement error is now zero. For example, one could use weighted least squares, etc. Focusing on the x coordinate of the solution we obtain from whatever process a set of estimations of x which can be written in the form $x+\varepsilon_i, i=1,\cdots,n$, where n is the number of approximate solutions that you wish to generate and ε_i is the difference between the approximate x and the true x. For good data the average of the approximate x's is a reasonable estimation of the true x. As the data measurement error increases the accuracy of the average degenerates rapidly and the standard deviation increases quickly. Consideration of plots of the approximate solutions in the presence of various synthetic error magnitudes shows the existence of clusters of approximate solutions with one such cluster near the known solution point. For example, see Figures 2-9. So the question of identifying the best cluster of approximations and using the average of coordinates of the points in that cluster as the solution approximation is another process that arises in a natural way.

F. Cluster Analysis. An early observation in the process of identification of the best cluster was that any process of selecting points without reference back to the original measured data could turn rapidly into a random number generator as the data error increased. Reference back to the measured u_i tended to slow down the process of fake solution generation. We shall outline one attempt that had a good chance of being a reasonable process. Let the solution approximations from some process such as least squares, etc. be denoted by $(x_k, y_k, r_k), k=1,\cdots,n$ and let

(56) $$u_{ik}=\sqrt{r_k^2+r_i^2-2x_ix_k-2y_iy_k}-r_k, i=1,\cdots,6,$$

where here the i subscript refers to the Site #i coordinates while the k subscripts refer to the coordinates/output of the calculation of the kth approximate solution. Let

(57) $$J_{ik}=|u_{ik}-u_i|,$$

where u_i denotes the measurement data as defined above in (17). Let

(58) $$w_{jk}=J_{i_jk}$$

be an ordering of the J_{ik} such that for each k

(59) $$w_{1k}\geq w_{2k}\geq\cdots\geq w_{6k}.$$

Various data fit measures for each solution (x_k, y_k, r_k) can be calculated. For example,

$$\begin{aligned}
&M_6(k) = w_{1k}, \quad k = 1, \cdots, n,\\
&M_5(k) = w_{2k}, \quad k = 1, \cdots, n,\\
&M_4(k) = w_{3k}, \quad k = 1, \cdots, n,\\
&L_6^2(k) = \left(\sum_{j=1}^{6} w_{jk}^2 \right)^{1/2}, k = 1, \cdots, n,\\
&L_5^2(k) = \left(\sum_{j=1}^{5} w_{jk}^2 \right)^{1/2}, k = 1, \cdots, n,\\
&L_4^2(k) = \left(\sum_{j=1}^{4} w_{jk}^2 \right)^{1/2}, k = 1, \cdots, n,\\
&L_6^1(k) = \left(\sum_{j=1}^{6} w_{jk} \right), k = 1, \cdots, n,\\
&L_5^1(k) = \left(\sum_{j=1}^{5} w_{jk} \right), k = 1, \cdots, n, \text{ and}\\
&L_4^1(k) = \left(\sum_{j=1}^{4} w_{jk} \right), k = 1, \cdots, n,
\end{aligned} \tag{60}$$

represent respectively the l_∞ norm on the Euclidean space of vectors $\vec{u} = (u_i, \cdots, u_6)$, the l_∞ semi-norm of the 5th largest absolute value, the l_∞ semi-norm of the 4th largest absolute value, the l_2 norm, the l_2 semi-norm of the smallest to the 5th largest absolute value, the l_2 semi-norm of the smallest to the 4th largest, the l_1 norm, the l_1 semi-norm of the smallest to the 5th largest absolute value, and the l_1 semi-norm of the smallest to the 4th largest. Each norm/data fit measure was ordered large to small and say the bottom 1/5 (could be 1/4 or whatever) were selected and a tally mark placed next to the solutions contained in that position of the list representing a "good data fit" relative to that measure/norm/semi-norm. The approximate solutions with the highest number of tally marks defined the cluster. The approximate solutions in the cluster could be viewed as reasonable approximations in a broad sense. However, as the data error increased fake solutions were observed.

To filter out fake solutions some additional measurements could be very helpful.

IV. An Angular Filter

Currently a short base line interferometer is under development for location at the central Site #0. The digital sample rate is estimated at a voltage measurement per 2 nanoseconds versus a 50 nanoseconds rate for the current LDAR system. The outcome of these measurements would allow the estimation of the direction cosines of the (x, y, z) event location point/vector; i.e., an estimate of azimuth angle θ and angle of elevation ϕ would yield an estimate of

$$\left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r}\right) = (\cos\phi \cos\theta, \cos\phi \sin\theta, \sin\phi) \tag{61}$$

Assuming that these angles could be determined to within $\pm 2^{\circ}$, tests were run on synthetic data which added error into the true angles selected from a uniform distribution of $\pm 2^{\circ}$. The solid angle

$$\Omega = \{(\theta, \phi): \phi_m - 4^{\circ} \leq \theta \leq \theta_m + 4^{\circ}, \phi_m - 4^{\circ} \leq \phi \leq \phi_m + 4^{\circ}\} \tag{62}$$

was designated as the angular filter, where θ_m and ϕ_m denote, respectively, the measured azimuth angle, and the measured elevation angle. Solutions outside of Ω were discarded. Figures 2-5 show the effect of the Ω filter on the 42 least squares solutions. These figures/results show the promise of the filter while its use in practice may differ from that of eliminating approximate solutions to find a good cluster. For example, one may wish to use it to test a solution obtained from some process - say the current method in operational use. Figures 6 and 7 show a good solution cluster excluded by the Ω filter. Use of the filter may incur the cost of exclusion of some good solutions. Figures 8 and 9 show a fake solution cluster, from which the current operational method would generate a solution. Clearly Ω would designate such a solution as a fake. Also, the volume $V(\Omega, r)$ of the tetrahedron with "vertex" at $(0, 0, 0)$ and "height" r which is generated by Ω has a rate of increase $\frac{dV}{dr} \sim r^2$. Thus, for large r, the Ω filter would probably accept some fakes.

V. Conclusions and Open Questions

Conclusions:

A. Any method will work with good measurement data.

B. No fool-proof method for eliminating fake solutions was found.

C. The Ω Filter shows promise of eliminating fakes with the possible cost of excluding some acceptable solutions. This cost might be acceptable given the rate of generation of lightning events within a thunder storm.

Open Questions:

A. Which method is best? "Best" would need careful definition. No attempt was made to find a "Best" method.

B. What is the best way to utilize the Ω filter to minimize acceptance of fakes.

C. What are the possible ways that fakes can be introduced via errors in measurements? Electrical Interference, etc.?

D. The condition number of the various matrices corresponding to the Sites i, j, and k, $1 \le i < j < k \le 6$, needs to be employed in the solution process to reduce error amplification.

E. Design of other mathematical filters or measurement filters to eliminate fakes.

References

1. Cianos, N., Oetzel, G.N., and Pierce, E.T., A technique for accurately locating lightning at close ranges, *J. Appl. Meteor.*, **11,** 1120-1127, 1972.

2. Holle, R.L. and Lopez, R.E., *Overview of real-time lightning detection systems and their meteorological uses,* NOAA Technical Memorandum ERL NSSL-102, 1993.

3. Koshak, W.J., Christian, H.J., and Solakiewicz, R., *On the retrieval of lightning radio sources from time of arrival data,* to appear.

4. MacClement, W.D. and Murty, R.C., VHF direction finder studies of lightning, *J. Appl. Meteor.*, **17,** 786-795, 1978.

5. Maier, L.M. , Lennon, C.L., Britt, T. and Schaefer, S., Lightning Detection and Ranging LLDAR) System Performance Analysis, Proceedings for the Sixth Conference on Aviation Weather Systems at the 75th American Meteorological Society Meeting, Dallas, Texas, January 15-20, 1995.

6. Murty, R.C. and MacClement, W.D., VHF direction finder for lightning location, *J. Appl. Meteor.*, **12,** 1401-1405, 1973.

7. Oetzel, G.N. and Pierce, E.T., VHF technique for locating lightning, *Radio SCI.,* **4,** 199-201, 1969.

8. Poehler, H.A., *An Accuracy Analysis of the LDAR System,* NASA Contractor Report CR-154631, 1977.

9. Poehler, H.A. and Lennon, C.L., *Lightning Detection & Ranging System (LDAR) system description & performance objectives,* NASA Technical Memorandum TM-74105, 1979.

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

11. Proctor, D.E., VHF radio pictures of cloud flashes, *J. Geophys. Res.,* **86,** 4041-4071, 1981.

12. Proctor, D.E., Lightning and precipitation in a small multicellular thunderstorm, *J. Geophys. Res.*, **88,** 5421-5440, 1983.

13. Proctor, D.E., Regions where lightning flashes began, *J. Geophys. Res.*, **96,** 5099-5112, 1991.

14. Rustan, P.L., Uman, M.A., Childers, D.G., Beasley, W.H., and Lennon, C.L., Lightning source locations from VHF radiation data for a flash at Kennedy Space Center, *L. Geophys. Res.*, **85,** 4893-4903, 1980.

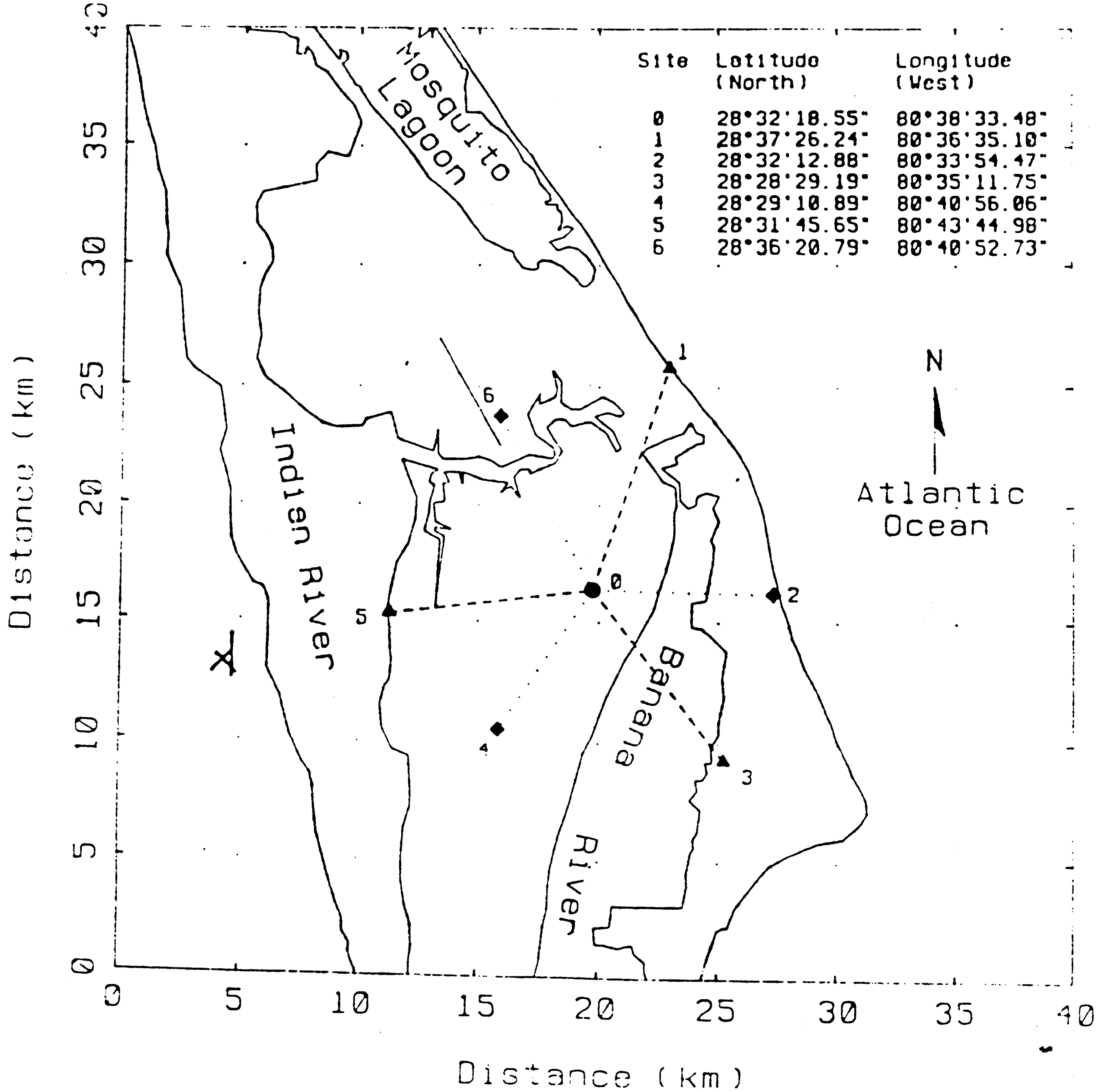
15. Taylor, W.L., A VHF technique for space-time mapping of lightning discharge processes. *J. Geophys. Res.*, **83,** 3575-3583, 1978.

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

17. Twomey, S.A., *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements,* Elsevier, New York, 1977.

18. Young, D.M., and Gregory, R.T., *A survey of Numerical Mathematics, Volumes I & II,* Addison Wesley Publishing Company, Reading Massachusetts, 1972.

FIGURE 1: LOCATIONS OF LDAR SITES.



TABLE#1:	L-00 CONDITION NUMBER FOR THE MATRIX OF (19)				
	WITH THETA= 45 DEGREES AND R=50 KM				
SITES: I, J, K	PHI=0 DEG'S	PHI=18 DEG'S	PHI=36 DEG'S	PHI=54 DEG'S	PHI=72 DEG'S
1 , 2 , 3	61.0435	55.93494	45.16827	34.93892	27.20576
1 , 2 , 4	146.0429	99.75604	52.32759	30.47398	20.38362
1 , 2 , 5	39.43026	34.02578	24.29061	16.73847	11.98991
1 , 2 , 6	52.37815	54.23213	59.37658	65.49181	66.27332
1 , 3 , 4	100.7596	76.14307	44.83299	27.66138	18.98708
1 , 3 , 5	32.64875	29.03469	21.9869	15.95798	11.84311
1 , 3 , 6	21.07241	20.74303	19.70602	17.8936	15.38053
1 , 4 , 5	178.6548	94.3098	40.25974	21.71146	14.14445
1 , 4 , 6	1203.788	177.7992	52.06078	25.43921	16.0372
1 , 5 , 6	152.1562	244.3768	349.3181	77.37132	41.66919
2 , 3 , 4	131.4395	174.2611	1930.4	144.6241	75.6265
2 , 3 , 5	70.76457	60.91602	43.5988	30.80794	25.34381
2 , 3 , 6	13.38127	13.341[illegible]	13.22082	13.02164	12.75794
2 , 4 , 5	59.92292	53.0831[illegible]	39.99109	29.05757	21.74881
2 , 4 , 6	34.35554	31.10017	24.43622	18.38688	14.06935
2 , 5 , 6	31.03037	27.61281	20.98866	15.37501	11.58672
3 , 4 , 5	35.88189	35.36891	33.79395	31.20216	30.69836
3 , 4 , 6	23.17594	22.47862	20.60585	18.11648	17.20661
3 , 5 , 6	22.03605	21.42122	19.75133	17.59359	16.77516
4 , 5 , 6	99.83394	93.51775	78.98998	63.55145	50.92098

TABLE#2: L-00 CONDITION NUMBER FOR THE MATRICES OF (19) WITH THETA=45 DEGREES AND PHI=36 DEGREES

SITES: I, J, K	R=10 KM	R=20 KM	R=30 KM	R=40 KM	R=50 KM
1 , 2 , 3	7.562704	16.79023	26.36976	35.796	45.16827
1 , 2 , 4	6.207406	17.32838	29.06532	40.72616	52.32759
1 , 2 , 5	3.800319	8.940697	14.12844	19.229	24.29061
1 , 2 , 6	13.99799	24.45105	35.80679	47.52315	59.37658
1 , 3 , 4	5.968515	15.2238	25.18018	35.03581	44.83299
1 , 3 , 5	3.606738	8.129786	12.81055	17.41429	21.9869
1 , 3 , 6	3.98372	7.696629	11.68949	15.69356	19.70602
1 , 4 , 5	5.305135	12.67618	21.85456	31.064	40.25974
1 , 4 , 6	7.163752	15.96919	28.01849	40.06938	52.06078
1 , 5 , 6	17.71435	80.49199	169.3603	260.5127	349.3181
2 , 3 , 4	37.80313	230.6574	1426.055	4403.898	1930.4
2 , 3 , 5	8.939254	17.56767	26.33672	34.99567	43.5988
2 , 3 , 6	3.54418	5.862628	8.291217	10.74992	13.22082
2 , 4 , 5	8.244946	16.25348	24.13949	32.07376	39.99109
2 , 4 , 6	4.713811	9.52872	14.48196	19.45647	24.43622
2 , 5 , 6	4.326795	8.54174	12.7215	16.86451	20.98866
3 , 4 , 5	9.347797	15.0197	21.15089	27.45198	33.79395
3 , 4 , 6	4.923535	8.49601	12.49855	16.54372	20.60585
3 , 5 , 6	4.831757	8.3426	12.12114	15.93046	19.75133
4 , 5 , 6	14.61789	29.3889	45.39818	62.07432	78.98998

TABLE#3: L-00 CONDITION NUMBER LEAST SQUARES MATRICES FROM (19)
WITH THETA=45 DEGREES AND R=50 KM

SITES: I, J, K	PHI=0 DEG'S	PHI=18 DEG'S	PHI=36 DEG'S	PHI=54 DEG'S	PHI=72 DEG'S
1, 2, 3	2987.496	2388.022	1569.014	950.2631	585.7231
1, 2, 4	15506.6	6879.296	1673.905	580.0257	265.5546
1, 2, 5	3733.733	2650.927	1215.24	599.3609	321.1228
1, 2, 6	1930.539	1949.161	2343.786	2862.446	2940.187
1, 3, 4	5998.944	3442.668	1209.967	469.6478	225.836
1, 3, 5	953.9588	719.7244	354.0826	189.247	107.3342
1, 3, 6	629.783	612.8223	560.3766	473.0363	362.5532
1, 4, 5	104203	27931.26	4473.017	1286.358	577.4056
1, 4, 6	2649161	58201.13	5091.993	1253.955	517.4218
1, 5, 6	18575.84	45718.49	84247.99	4246.766	1271.488
2, 3, 4	12363.5	20597.47	2495624	13788.87	2983.77
2, 3, 5	4428.875	3280.984	1681.046	826.7447	452.4155
2, 3, 6	377.0717	366.1686	335.1751	289.021	234.8757
2, 4, 5	3358.238	2625.204	1473.79	765.6875	421.2037
2, 4, 6	1182.481	916.8083	474.3499	265.4693	153.7965
2, 5, 6	2123.316	1679.989	969.1752	520.407	297.7738
3, 4, 5	1672.863	1557.485	1280.792	1092.266	849.8152
3, 4, 6	433.0106	409.7616	350.7537	278.268	210.0583
3, 5, 6	343.2147	327.5543	287.0623	235.659	185.2436
4, 5, 6	11024.89	9260.442	5741.057	3450.198	2245.19

TABLE#4: L-00 CONDITION NUMBER FOR LEAST SQUARES MATRICES FROM (19)
WITH THETA=45 DEGREES AND PHI=36 DEGREES

SITES: I, J, K	R=10 KM	R=20 KM	R=30 KM	R=40 KM	R=50 KM
1 , 2 , 3	37.05391	215.527	533.1044	984.2259	1569.014
1 , 2 , 4	26.6828	190.8348	523.6204	1018.672	1673.905
1 , 2 , 5	38.41716	182.0089	428.0012	772.625	1215.24
1 , 2 , 6	130.5545	395.334	849.6547	1499.456	2343.786
1 , 3 , 4	19.31481	139.8363	381.1006	738.2283	1209.967
1 , 3 , 5	10.38426	51.70143	122.7587	223.234	354.0826
1 , 3 , 6	18.69834	83.11837	194.7259	353.7285	560.3766
1 , 4 , 5	72.1908	500.4724	1389.842	2716.575	4473.017
1 , 4 , 6	68.01589	541.2159	1553.214	3074.353	5091.993
1 , 5 , 6	251.0914	4836.73	20463.48	47424.2	84247.99
2 , 3 , 4	694.2905	31356.67	1272577	1.27E+07	2495624
2 , 3 , 5	51.54088	246.6884	586.2598	1064.749	1681.046
2 , 3 , 6	12.89667	51.93632	118.5101	212.9032	335.1751
2 , 4 , 5	51.72648	229.1964	527.2562	942.3237	1473.79
2 , 4 , 6	15.03352	67.72181	160.5597	296.5604	474.3499
2 , 5 , 6	35.81037	152.6888	348.2725	620.553	969.1752
3 , 4 , 5	100.3142	259.6986	503.1655	840.332	1280.792
3 , 4 , 6	16.43427	57.55901	126.9724	224.7061	350.7537
3 , 5 , 6	15.0548	49.47289	106.3153	185.5214	287.0623
4 , 5 , 6	219.3776	835.2269	1955.032	3589.375	5741.057

FIGURE 2: ADMISSIBLE SOLUTIONS FROM 42 LEAST SQUARES SOLUTIONS WITH MAXIMUM RANDOM ERROR 750 METERS AND SOLUTION AT X=15 KM, Y= 15 KM , AND Z=20 KM

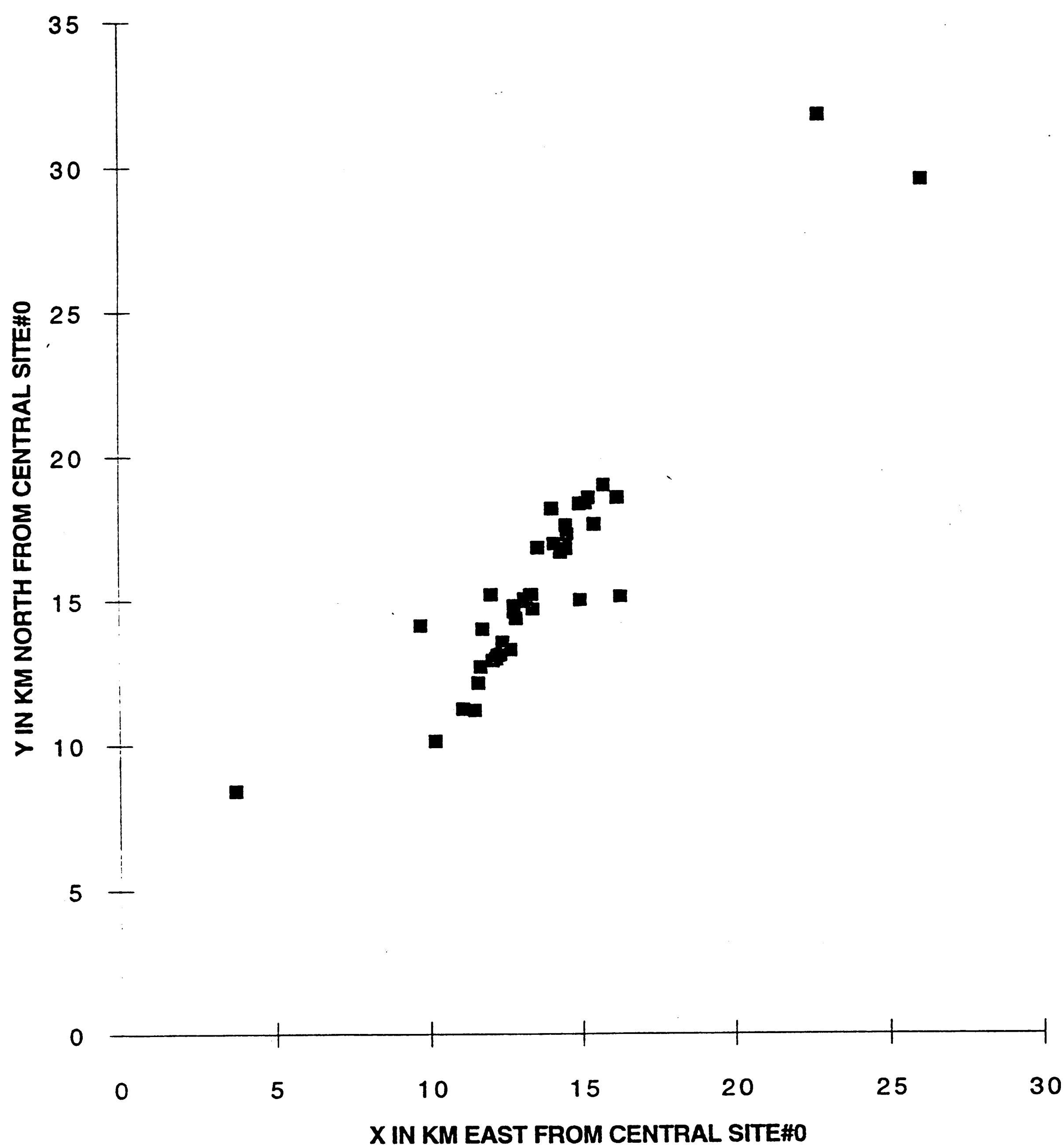


FIGURE 3: ANGLE FILTERED SOLUTIONS FROM ADMISSIBLE SOLUTIONS WITH MAXIMUM RANDOM ERROR 750 METERS AND SOLUTION AT X=15 KM, Y= 15 KM AND Z= 20 KM

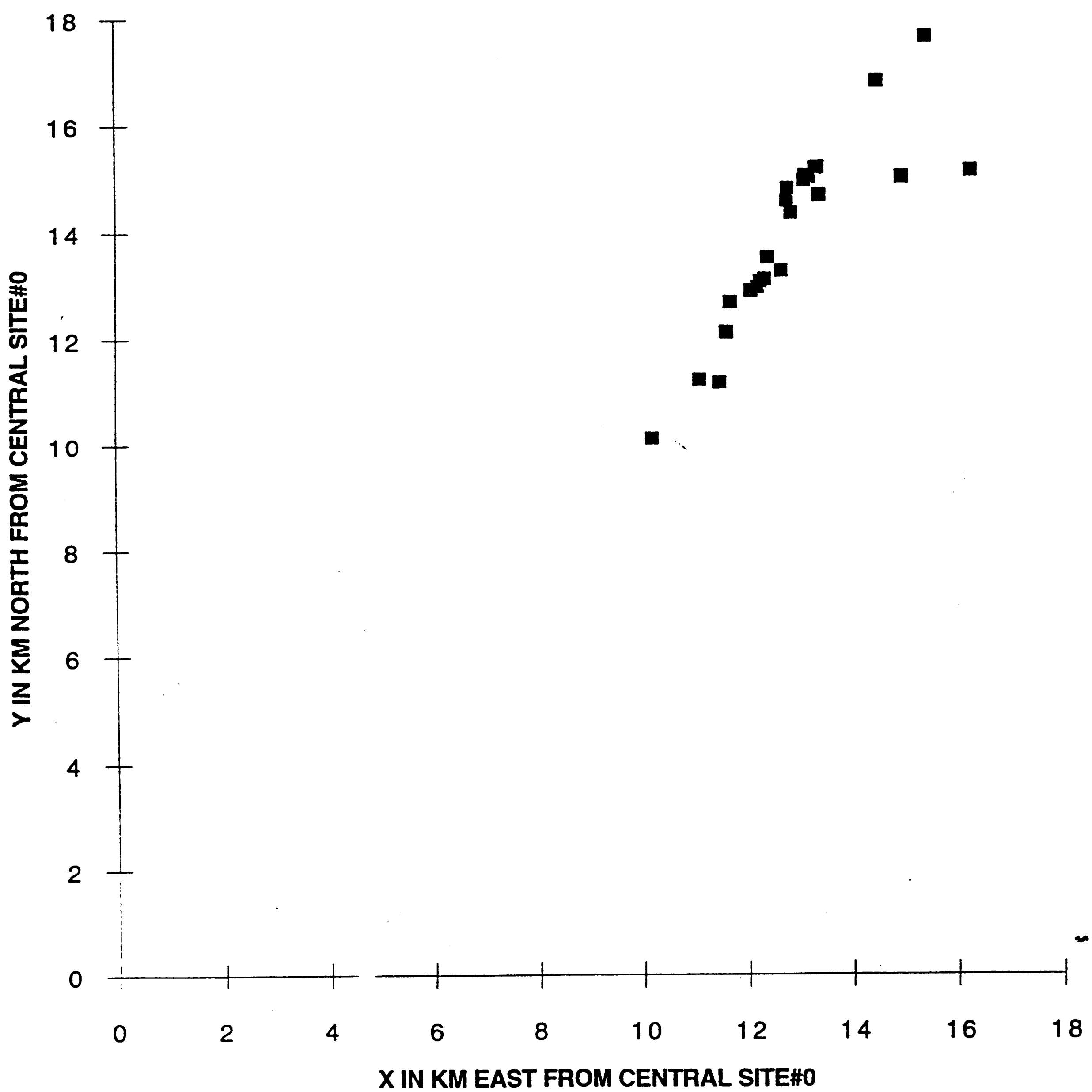


FIGURE 4: ADMISSIBLE SOLUTIONS FROM 42 LEAST SQUARES SOLUTIONS WITH MAXIMUM RANDOM ERROR 750 METERS AND SOLUTION AT X=15 KM, Y=15 KM, AND Z= 20 KM

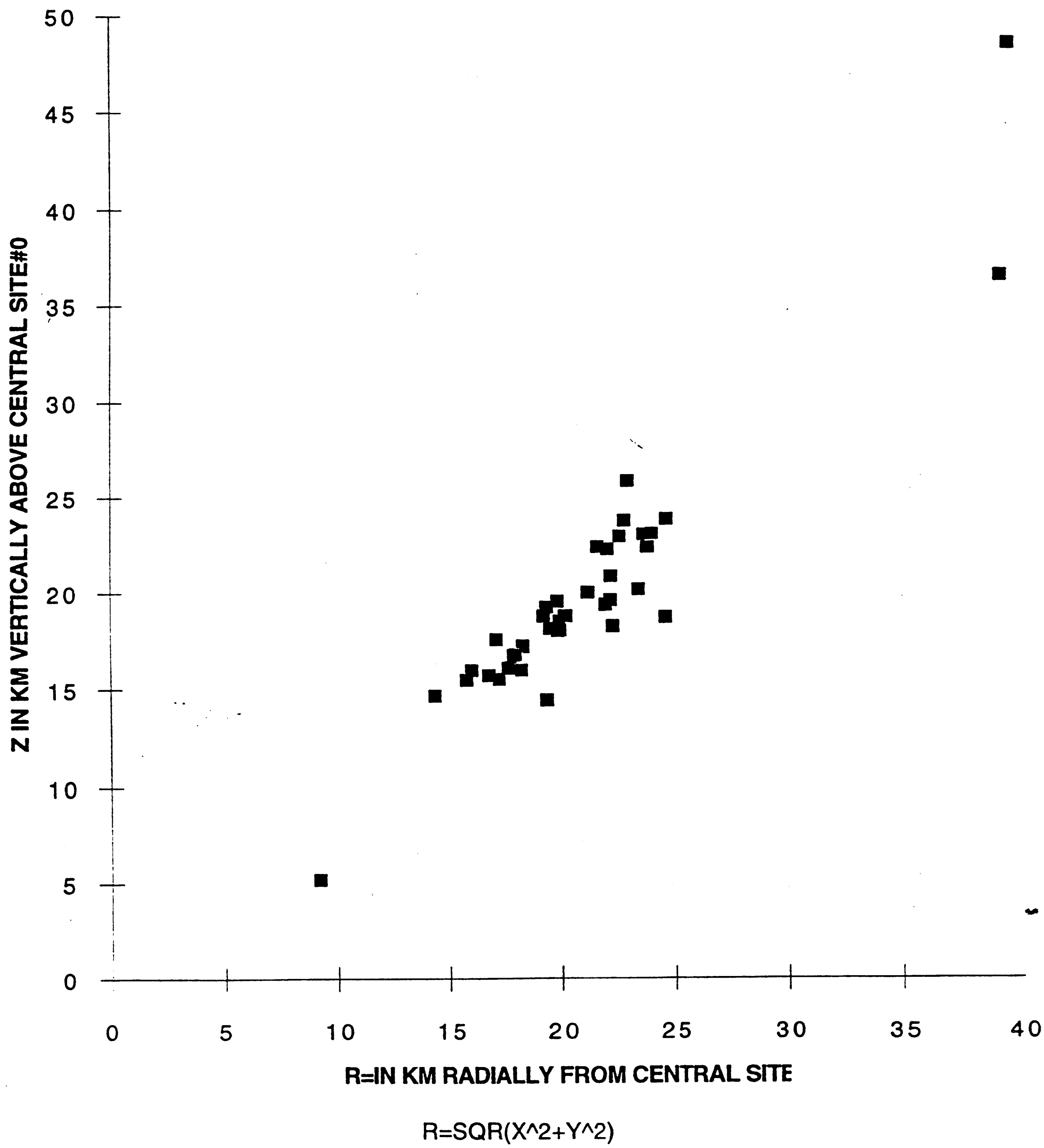


FIGURE 5: ANGLE FILTERED SOLUTIONS FROM ADMISSIBLE SOLUTIONS WITH MAXIMUM RANDOM ERROR 750 METERS AND SOLUTI0N AT X=15 KM, Y=15 KM, AND Z= 20 KM

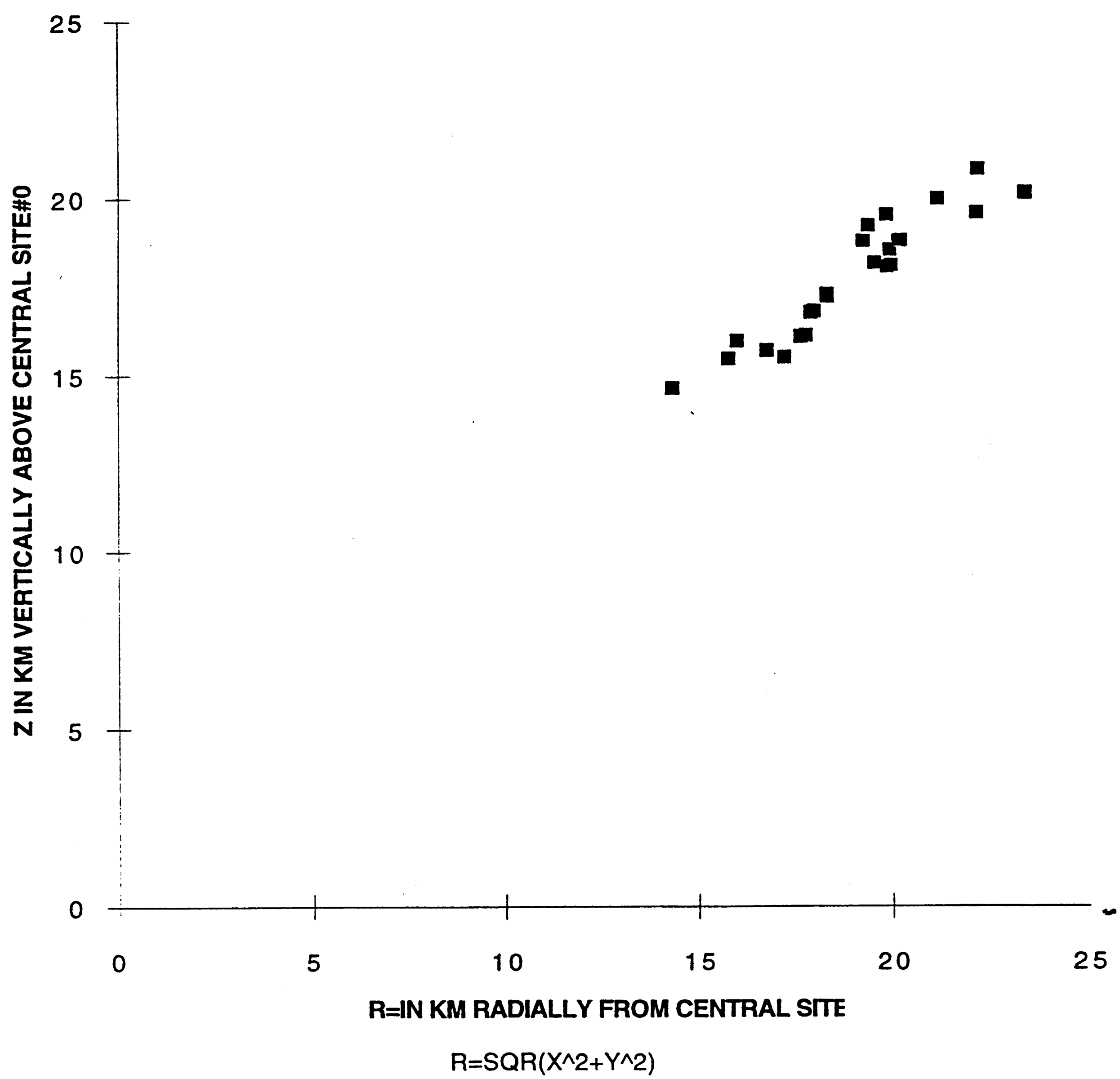


FIGURE 6: ADMISSIBLE SOLUTIONS FROM 42 LEAST SQUARES SOLUTIONS WITH SOLUTION AT X= 15 KM, Y=15 KM, AND Z=20 KM

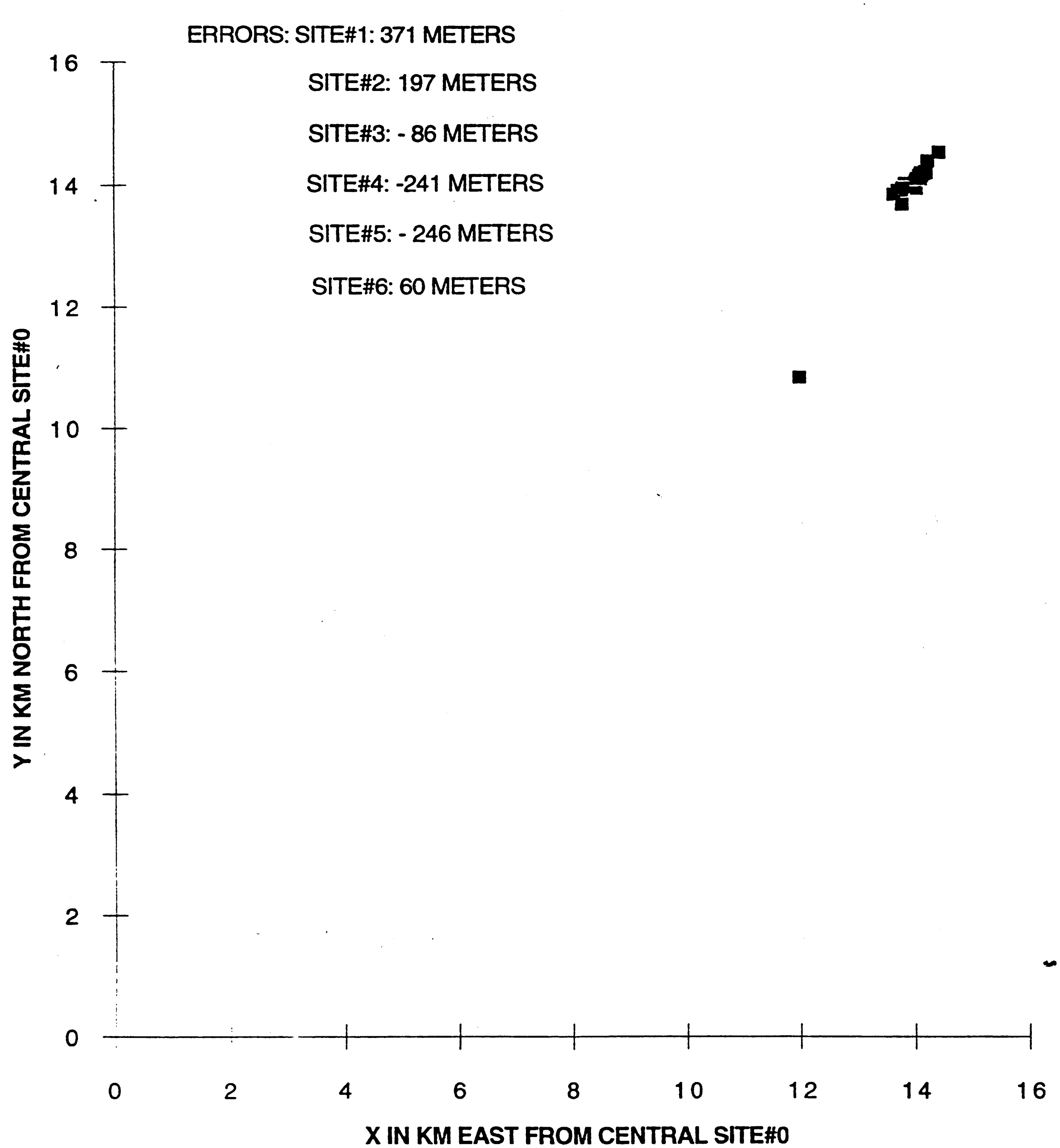


FIGURE 7: ADMISSIBLE SOLUTIONS FROM 42 LEAST SQUARES SOLUTIONS WITH SOLUTION AT X=15 KM, Y=15 KM, AND Z=20 KM

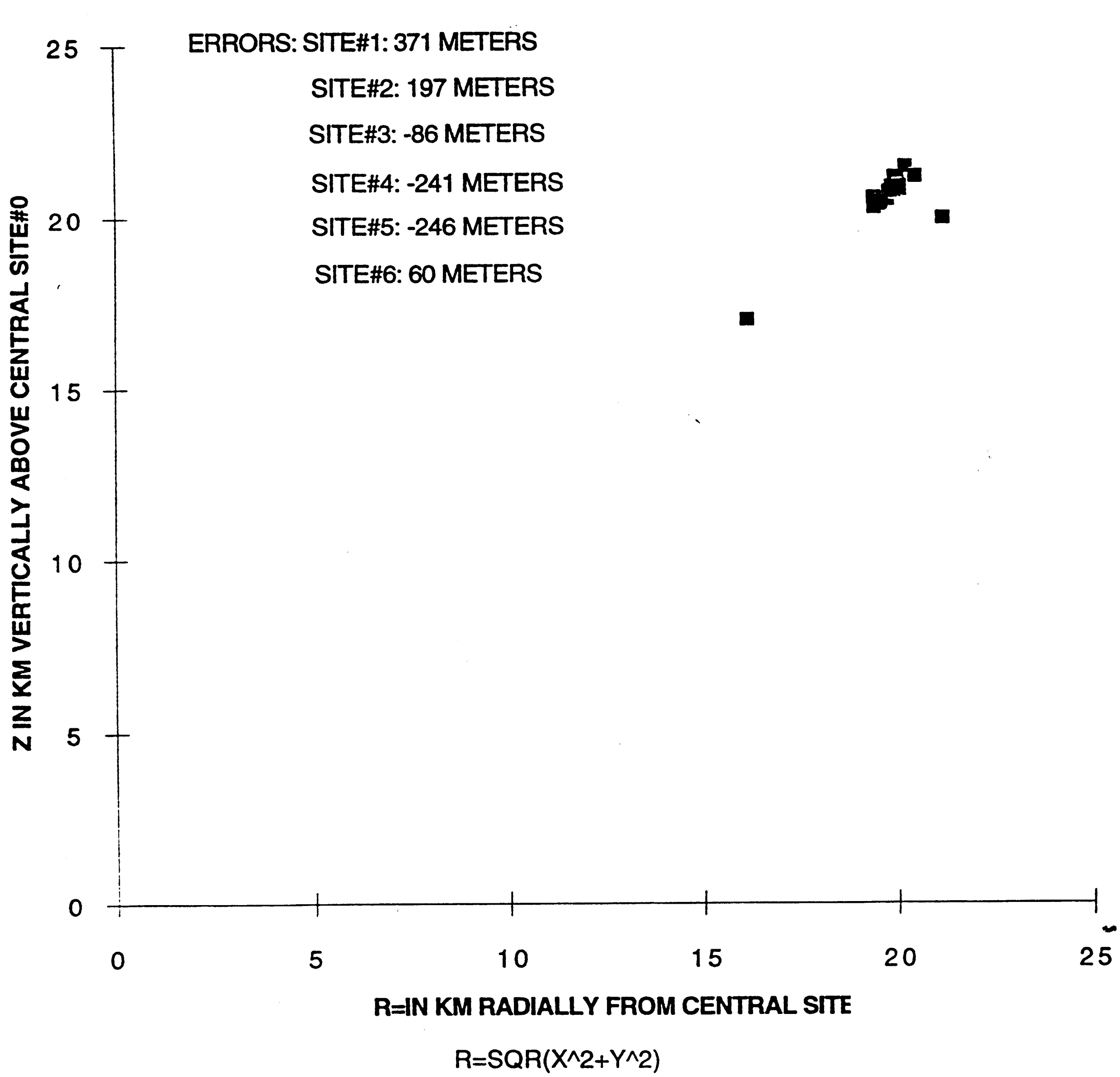


FIGURE 8: ADMISSIBLE SOLUTIONS FROM 42 LEAST SQUARES SOLUTIONS WITH SOLUTION AT X=15 KM, Y=15 KM, Z=20 KM

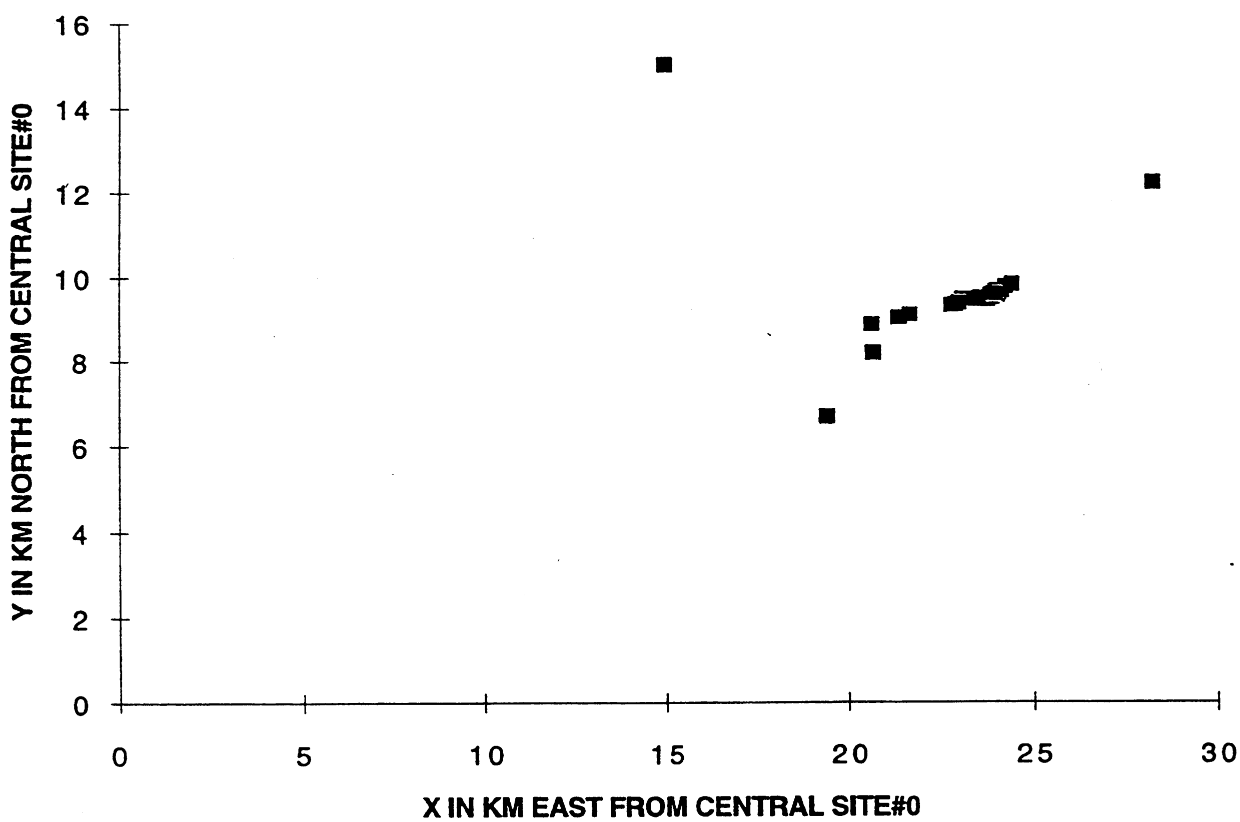


FIGURE 9: ADMISSIBLE SOLUTIONS FROM 42 LEAST SQUARES SOLUTIONS WITH SOLUTION AT X=15 KM, Y=15 KM, AND Z=20 KM