Bulletin of the Seismological Society of America, Vol. 76, No. 6, pp. 1699–1717, December 1986

APPRAISING EARTHQUAKE HYPOCENTER LOCATION ERRORS: A COMPLETE, PRACTICAL APPROACH FOR SINGLE-EVENT LOCATIONS

BY GARY L. PAVLIS

ABSTRACT

For conventional single-event, nonlinear, least-squares hypocentral estimates, I show that the total error is expressible as a linear combination of three terms: (1) measurement error; (2) modeling errors caused by inadequacy of the travel-time tables; and (3) a nonlinear term. Errors in calculating travel-time partial derivatives are shown to have no effect, provided a stable solution can be found. This is in contrast to linear problems where errors in calculating matrix elements can distort the solution drastically. The error appraisal technique developed here examines each of the three error terms independently. The first can be analyzed by standard confidence ellipsoids with critical values based on measurement error statistics. The second can cause conventional error ellipsoid calculations that derive a critical value from an estimate based on rms residuals, to give misleading results. I introduce an alternative extremal bound procedure for appraising such errors. Travel-time modeling errors are bounded as the product of ray arc length and an estimate of the nominal scale of slowness errors along the ray path. These are used to derive an upper bound on systematic errors in each hypocentral coordinate based on a novel bounding criteria. Finally, I show that, for errors of a reasonable scale, the nonlinear error term can be estimated adequately using a second-order approximation. Given an upper bound on the total location error, bounds on the travel-time error induced by nonlinearity can be calculated from the spectral norm of the Hessian for each measured arrival time. The systematic errors in each hypocentral coordinate due to nonlinearity can then be bounded using the same criteria used for constructing modeling error bounds. This overall procedure is complete because it allows one to independently appraise the relative importance of all sources of hypocentral errors. It is practical because the required computational effort is small.

INTRODUCTION

Some of the most significant advances in earth science that can be attributed to seismology have resulted from interpretations of spatial patterns of earthquake hypocenters. For large scale observations, such as the fact that earthquakes occur mainly along plate boundaries, the issue of earthquake location errors is of minor importance. As we try to extract more information from seismicity patterns, however, we soon face the issue of how precise earthquake locations really are.

It is well known that earthquake locations are subject to two quite different types of errors: relative location scatter and systematic biases (e.g., Douglas, 1967; Dewey, 1971, 1972; Jordan and Sverdrup, 1981). The first is caused by measurement errors and short range fluctuations in the seismic velocity structure of the earth (Leaver, 1984) and can properly be treated from a purely statistical point of view. The latter is caused largely by inadequate knowledge of even the long wavelength velocity structure of the earth (e.g., Jordan and Sverdrup, 1981; Pavlis and Hokanson, 1985). I argue below that it is not proper to consider such errors as resulting from a zero mean random process. Conventional error ellipsoids (Flinn, 1965) implicitly make this assumption (Jordan and Sverdrup, 1981), which I show here can yield very misleading results. In this paper, I introduce a new, alternative method for bounding

systematic mislocation errors based on a component-wise bounding theory. Unlike simple matrix vector norm bounding criteria, the bounds constructed by this new approach do not tend to be overly pessimistic, provided one can specify a reasonable upper bound on the scale of slowness anomalies within the network.

Most conventional error estimates are based on a linear approximation to a set of nonlinear equations [techniques discussed by Tarantola and Valette (1982) and Rowlett and Forsyth (1984) are exceptions]. The analysis I give here indicates that nonlinearity can be viewed as a different type of systematic bias whose effect is virtually inseparable from that due to modeling errors. On the positive side, however, calculations presented here suggest the influence of nonlinearity is probably small for most locations and is largely controlled by modeling error biases. In any case, it is shown that such errors can be bounded using a second-order approximation and a component-wise bounding theory similar to that used for bounding the influence of modeling errors.

The net result of this paper is a procedure for appraising location errors that I claim is complete and practical. It is complete because it gives relatively independent measures for all three sources of hypocentral error: (1) measurement errors; (2) velocity model errors; and (3) nonlinearity. The procedure is practical because the error estimates described are relatively easy to calculate and require few assumptions. This is in contrast to the complete error analysis of Tarantola and Valette (1982) which is complete but not necessarily practical because it is computationally demanding and requires *a priori* assumptions about the form of the uncertainty in the velocity structure of the earth (Thurber, 1986).

Fundamental Equations and Definitions

The data always used for earthquake locations are a set of m arrival times measured from one or more phases recorded by a network of seismometers. I will denote this vector of measured arrival times as $\langle \boldsymbol{t} \rangle \in R^m$. It is related to reality as follows

$$\langle \boldsymbol{t} \rangle = \boldsymbol{t} + \tau \mathbf{1} + \bar{e} \tag{1}$$

where

$\boldsymbol{t}$ = true travel-time vector;
τ = origin time ($\mathbf{1} \in R^m$ denotes a vector of all ones); and
$\bar{\boldsymbol{e}}$ = measurement error.

With actual data, $\boldsymbol{t}$ is unknowable. Instead, we must always rely on a mathematical model of $\boldsymbol{t}$ calculated from some estimate of the earth's seismic velocity structure. I use $\boldsymbol{t}_{\text{model}}(\hat{\boldsymbol{x}})$ to symbolize this vector of travel times. $\boldsymbol{t}_{\text{model}}$ is a function of the estimated *spatial* coordinates of the hypocenter, $\hat{\boldsymbol{x}}$. A complete specification of the hypocenter, of course, also requires an estimate of τ, $\hat{\tau}$, as well. For convenience, this 4-vector will be symbolized as $\hat{\boldsymbol{h}}$. With these definitions, I define the residual function

$$\bar{\boldsymbol{r}}(\hat{\boldsymbol{h}}) = \langle \boldsymbol{t} \rangle - \boldsymbol{t}_{\text{model}}(\hat{\boldsymbol{x}}) - \hat{\tau}\mathbf{1}. \tag{2}$$

Using equation (1), equation (2) becomes

$$\bar{\boldsymbol{r}}(\hat{\boldsymbol{h}}) = \bar{\boldsymbol{e}}_{\text{model}}(\hat{\boldsymbol{x}}) + \Delta\hat{\tau}\mathbf{1} + \bar{\boldsymbol{e}} \tag{3}$$

where $\bar{\boldsymbol{e}}_{\text{model}}(\hat{\boldsymbol{x}}) = \boldsymbol{t} - \boldsymbol{t}_{\text{model}}(\hat{\boldsymbol{x}})$ and $\Delta\tau = \tau - \hat{\tau}$.

All computerized earthquake location methods minimize $\| \bar{\boldsymbol{r}}(\hat{\boldsymbol{h}}) \|$, where $\| \ \|$ denotes a vector norm. The details of how this proceeds vary greatly, but all methods are the same in one respect. All are iterative procedures that calculate a series of corrections, $\delta\boldsymbol{h}_k$, to estimate $\boldsymbol{h}$ as

$$\hat{\boldsymbol{h}} = \boldsymbol{h}_0 + \sum_{k=1}^{n_{\text{iter}}} \delta\boldsymbol{h}_k \tag{4}$$

where n_{iter} is the total number of corrections required to obtain a stable solution from an initial guess, $\boldsymbol{h}_0$, of the hypocenter. By stable I mean specifically that the sequence $\delta\boldsymbol{h}_k$ is convergent such that

$$\| \delta\boldsymbol{h}_k \| < \epsilon, \tag{5}$$

where ϵ is some appropriately small number. The $\delta\boldsymbol{h}_k$ are always calculated as

$$\delta\boldsymbol{h}_k = \boldsymbol{A}^*_{k-1} \bar{\boldsymbol{r}}(\hat{\boldsymbol{h}}_{k-1}) \tag{6}$$

with $\hat{\boldsymbol{h}}_{k-1} = \boldsymbol{h}_0 + \sum_{l=1}^{k-1} \delta\boldsymbol{h}_1$. $\boldsymbol{A}^*_{k-1} \in R^{4\times m}$ is a generalized inverse. With the exception of the nonlinear method described by Thurber (1985), $\boldsymbol{A}^*_{k-1}$ is always calculated directly from the matrix of partial derivatives $\bar{\boldsymbol{A}}$ with components

$$\bar{A}_{ij} = \frac{\partial}{\partial h_j} (t_{\text{model}})_i, \qquad (j = 1, 2, 3), \quad \text{and}$$

$$\bar{A}_{i4} = 1(h_4 = \tau). \tag{7}$$

There are essentially as many variations in how $\boldsymbol{A}_k{}^*$ is calculated as there are location programs. Fortunately, we do not have to present a different error analysis for every possibility that exists. The reason is that every existing location method can be equated to a weighted least-squares problem. For our purposes, this means specifically that there exists a positive definite matrix, $\boldsymbol{W} \in R^{m\times m}$ for which the solution to the usual equations of condition

$$\bar{\boldsymbol{A}}\delta\boldsymbol{h} = \bar{\boldsymbol{r}}(\hat{\boldsymbol{h}}) \tag{8}$$

satisfies

$$\delta\boldsymbol{h} = \boldsymbol{A}^+\boldsymbol{r} < \epsilon \tag{9}$$

where

$$\boldsymbol{A} = \boldsymbol{W}\bar{\boldsymbol{A}}, \quad \boldsymbol{r} = \boldsymbol{W}\bar{\boldsymbol{r}}, \tag{10}$$

ϵ is as in equation (5), and $\boldsymbol{A}^+$ denotes the pseudoinverse of $\boldsymbol{A}$ (e.g., Lawson and Hanson, 1974, pp. 36–40). Equation (9) means that $\hat{\boldsymbol{h}}$ is a stable solution in the least-squares sense to the weighted equations of condition, $\boldsymbol{W}\bar{\boldsymbol{A}}\delta\boldsymbol{h} = \boldsymbol{W}\bar{\boldsymbol{r}}$. With standard least-squares procedures like HYPO71 (Lee and Lahr, 1975), the corre-

spondence is obvious. With procedures using damped least squares (e.g., Herrmann, 1979) or the recently published application of Newton's method by Thurber (1985), the correspondence is not so obvious. However, equation (9) still holds because the purpose of both algorithms is to promote convergence of the sequence defined in equation (4). Both seek a solution that minimizes $\| \boldsymbol{r} \| = [\bar{\boldsymbol{r}}^T \boldsymbol{W}^2 \bar{\boldsymbol{r}}]^{1/2}$. As long as the location is well constrained, both methods will converge to the same solution as an equivalent least-squares procedure. Furthermore, even a procedure which seeks to minimize $L1$ ($\| \bar{\boldsymbol{r}} \|_1 = \sum_{i=1}^{m} | \bar{r}_i |$) can be cast in this form (Anderson, 1982). We are also forced, on the other hand, to make three fundamental assumptions.

1. $\hat{\boldsymbol{h}}$ is well constrained. That is, we assume $\boldsymbol{A}^+$ is not singular.
2. $\hat{\boldsymbol{h}}$ is the global minimum of $\| \boldsymbol{r} \|$, not a local minima.
3. $\hat{\boldsymbol{h}}$ is not enormously different from the true hypocenter, $\boldsymbol{h}$.

The consequences of nos. 2 and 3 are identical, but the way they can arise is different. What I mean by "enormously different" is stated specifically below.

Analysis of Hypocentral Errors

Second-order theory. $\hat{\boldsymbol{h}}$ inevitably contains errors that arise from a number of factors. To examine this, expand each component of the residual vector in a Taylor series as follows (Lee and Stewart, 1981, p. 124)

$$\bar{r}_i(\hat{\boldsymbol{h}} + \delta \boldsymbol{h}) = \bar{r}_i(\hat{\boldsymbol{h}}) + \sum_{j=1}^{4} \bar{A}_{ij} \delta h_j + \bar{n}_i \tag{11}$$

where $\bar{A}_{ij}$ is as defined in equation (7), and $\bar{n}_i$ is the sum of all second and higher order terms of the expansion. The m equations of (11) can be summarized in matrix form as

$$\bar{\boldsymbol{r}}(\hat{\boldsymbol{h}} + \delta \boldsymbol{h}) = \bar{\boldsymbol{r}}(\hat{\boldsymbol{h}}) + \bar{\boldsymbol{A}} \delta \boldsymbol{h} + \bar{\boldsymbol{n}} \tag{12}$$

where the correspondence of different terms is obvious. Two facts about $\bar{\boldsymbol{A}}$ and $\bar{\boldsymbol{n}}$ need to be emphasized: (1) they are calculated from a mathematical model of reality, and (2) they are themselves nonlinear functions of $\boldsymbol{h}$. In equation (12), both are evaluated at $\hat{\boldsymbol{h}}$.

We shall put equation (12) to use by assuming $\delta \boldsymbol{h}$ is the difference between $\hat{\boldsymbol{h}}$ and reality (i.e., $\boldsymbol{h} = \hat{\boldsymbol{h}} + \delta \boldsymbol{h}$). If we then apply the weighting matrices defined in equations (8) to (10), we obtain the similar form

$$\boldsymbol{r} = \hat{\boldsymbol{r}} + \boldsymbol{A} \delta \boldsymbol{h} + \boldsymbol{n} \tag{13}$$

where $\boldsymbol{n} = \boldsymbol{W} \bar{\boldsymbol{n}}$. $\boldsymbol{r}$ and $\hat{\boldsymbol{r}}$ are a shorthand used to denote $\boldsymbol{r}(\boldsymbol{h})$ and $\boldsymbol{r}(\hat{\boldsymbol{h}})$, respectively. From equation (9), we see

$$\boldsymbol{A}^+ \boldsymbol{r} - \delta \boldsymbol{h} + \boldsymbol{A}^+ \boldsymbol{n} = \boldsymbol{A}^+ \hat{\boldsymbol{r}} < \epsilon, \tag{14}$$

since we require $\boldsymbol{A}^+ \boldsymbol{A} = \boldsymbol{I}$. However, from equation (5), $\| \boldsymbol{A}^+ \hat{\boldsymbol{r}} \| \approx 0$. Hence, using equation (3), equation (14) can be written as

$$\delta \boldsymbol{h} \approx \boldsymbol{A}^+ (\boldsymbol{e}_{\text{model}} + \boldsymbol{n} + \boldsymbol{e}) \tag{15}$$

within the tolerance parameter specified by ϵ. (Note $\boldsymbol{e}_{\text{model}}$ is evaluated at $\boldsymbol{h}$, not $\hat{\boldsymbol{h}}$.)

A conventional analysis would assume $\boldsymbol{n} \approx 0$, yielding classical results involving error ellipsoids (Flinn, 1965). One of the main points of this paper, however, is to consider the importance of $\boldsymbol{n}$. To do so, we need a method for calculating it. Application of equation (5.68) of Lee and Stewart (1981) to equation (11) yields

$$\bar{n}_i = \tfrac{1}{2}\delta\boldsymbol{h}^T\boldsymbol{H}_i\delta\boldsymbol{h} + \cdots \tag{16}$$

where $\boldsymbol{H}_i$ is the Hessian matrix. Analytic forms for $\boldsymbol{H}_i$ for a constant velocity medium and a two-layer model are given in a recent paper by Thurber (1985). I use Thurber's results below to approximate $\bar{n}_i$ to second order as

$$\bar{n}_i \approx \tfrac{1}{2}\delta\boldsymbol{h}^T\boldsymbol{H}_i\delta\boldsymbol{h} = (\bar{n}_2)_i. \tag{17}$$

Equation (15) then becomes

$$\delta\boldsymbol{h} \approx \boldsymbol{A}^+(\boldsymbol{e}_{\text{model}} + \boldsymbol{n}_2 + \boldsymbol{e}). \tag{18}$$

There are two levels of approximation in equation (18): (1) the convergence criterion ϵ, and (2) the second-order approximation of $\boldsymbol{n}$. (The latter is emphasized with the 2 subscript on $\boldsymbol{n}$.)

Equation (18) is the focal point of this paper. The first step is clearly to investigate the limits of this approximation. This is done in the following section using computer simulations.

Computer simulations. In this study, I chose to consider the location precision of earthquakes in the vicinity of the rupture zone of the 1984 Morgan Hill, California, earthquake (Cockerham and Eaton, 1984). A crude approximation to the velocity structure in this area was used to generate travel times for a series of synthetic events. This velocity model consisted of two constant velocity quarter-spaces joined along a vertical plane striking north 31.5° west and passing through the point 37°16′ north latitude by 121°40′ west longitude (Figure 1). Synthetic arrival times for every station in the U.S. Geological Survey Central California Network (CALNET) and all University of California Berkeley stations within 100 km of this point (121 stations) were calculated from this model for the set of events shown in Figure 1 using two different quarter-space models. Velocities for these two models are given in Table 1. The measurement error vector $\bar{\boldsymbol{e}}$ in equation (1) was simulated by using a random number generator to produce random samples from a normal distribution with zero mean and a variance of 0.05 sec. The same $\bar{\boldsymbol{e}}$ was then added to each synthetic event arrival time vector. The advantage of this is that it allows comparison of errors induced by $\bar{\boldsymbol{e}}$ as a function of position. On the other hand, it gives a privileged position to a set of random numbers. However, repetition of these results with different random vectors indicates the results are not very dependent on the exact choice of $\bar{\boldsymbol{e}}$, and the one presented here is representative.

All events were located using a simple, damped least-squares procedure similar to that described by Herrmann (1979), using travel times calculated from a constant velocity medium with a velocity of 5.6 km/sec. Figures 2 and 3 show the resulting location estimates. All location estimates show an eastward bias caused by approximating the quarter-space model with a constant velocity medium. The scale of this bias is ≈10 km for model A events and ≈2 km for model B events. Figure 4 can be used to examine the validity of the second-order approximation in the context of these two different error scales. This is summarized here by examining only the

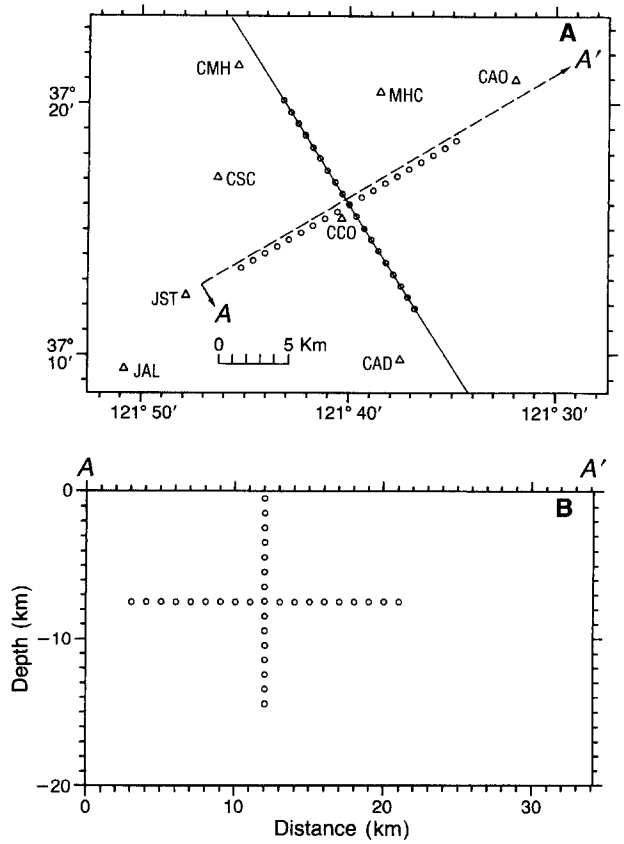

FIG. 1. Actual location of synthetic events examined in this study. (A) Map view. (B) Cross-section. Station locations are marked with triangles and their three character name code. The cross-section AA' is as shown in the map view. The diagonal line running from northwest to southeast marks the boundary of the quarter-spaces used to generate the synthetic data arrival times (Table 1).

TABLE 1
SYNTHETIC QUARTER-SPACE MODEL PARAMETERS

Model	Southwest Quarter-Space Velocity (km/sec)	Northeast Quarter-Space Velocity (km/sec)
A	5.0	6.2
B	5.5	5.7

norm of the first- and second-order prediction error

$$E_1 = \| \delta \boldsymbol{h} - \boldsymbol{A}^+(\boldsymbol{e}_{\text{model}} + \boldsymbol{e}) \|$$

and

$$E_2 = \| \delta \boldsymbol{h} - \boldsymbol{A}^+(\boldsymbol{e}_{\text{model}} + \boldsymbol{n}_2 + \boldsymbol{e}) \| .$$

The 4-vector norm is calculated for any 4-vector $\boldsymbol{a}$ as

$$\| \boldsymbol{a} \| = [a_1^2 + a_2^2 + a_3^2 + v_0^2 a_4^2]^{1/2} \tag{19}$$

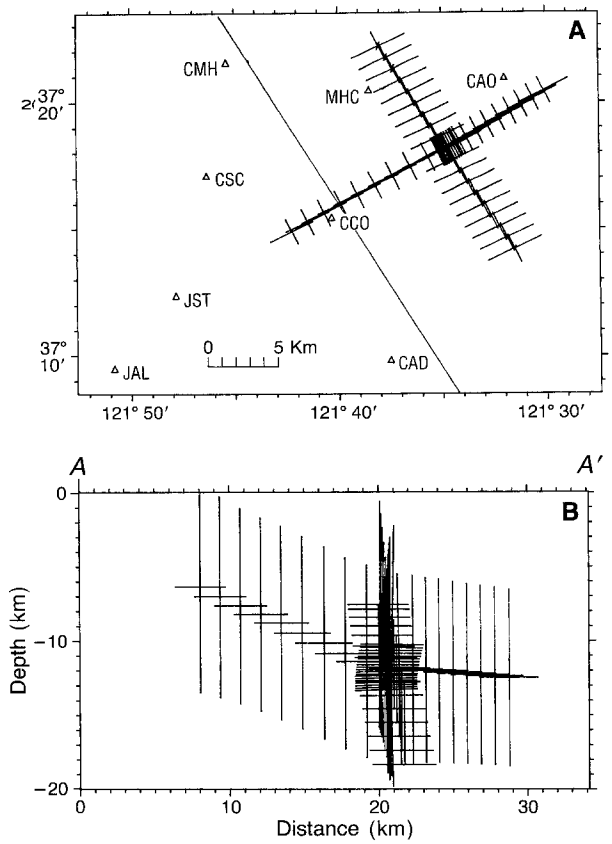

FIG. 2. Estimated locations of synthetic events from model A using 5.6 km/sec constant velocity medium. For comparison, the map view (A) and cross-section (B) frames are identical to those shown in Figure 1. Estimated locations are at the centers of crossing lines which are the projections of the major axes of the conventional 95 per cent confidence ellipsoids. Critical values for these ellipsoids are based on an F statistic as originally advocated by Flinn (1966).

where v_0 is the constant velocity medium velocity. v_0 is used to convert origin time errors into an equivalent length scale so all components of $\boldsymbol{a}$ are at least measured in the same units.

For model A, the first-order errors are extremely large. This indicates that in the presence of such a large bias, conventional first-order analysis techniques can be expected to give very misleading results. Inclusion of the second-order term improves the picture by a factor of about 4, but E_2 is not negligibly small. E_2 is still almost two orders of magnitude larger than the convergence parameter ϵ which was 10 m for all tests. On the other hand, for model B, E_2 is generally of the same size as ϵ. The only exceptions are the three shallowest sources along the vertical arm of the jack-shaped test pattern. This failure is to be expected, however, because the depths of these three events are poorly constrained and required auxiliary constraints to obtain a stable solution. This is not accounted for in this theory. Furthermore, Figure 2 shows that these events are located very close to station CCO. Since the second-order term varies as R^{-1}, where R is the hypocentral distance from a given station (Thurber, 1985), one expects the third-order term to vary as R^{-2}. Thus, higher order terms would only be important for very short hypocentral distances, which is consistent with these results.

I conclude from the results presented in Figure 4 that the second-order approx-

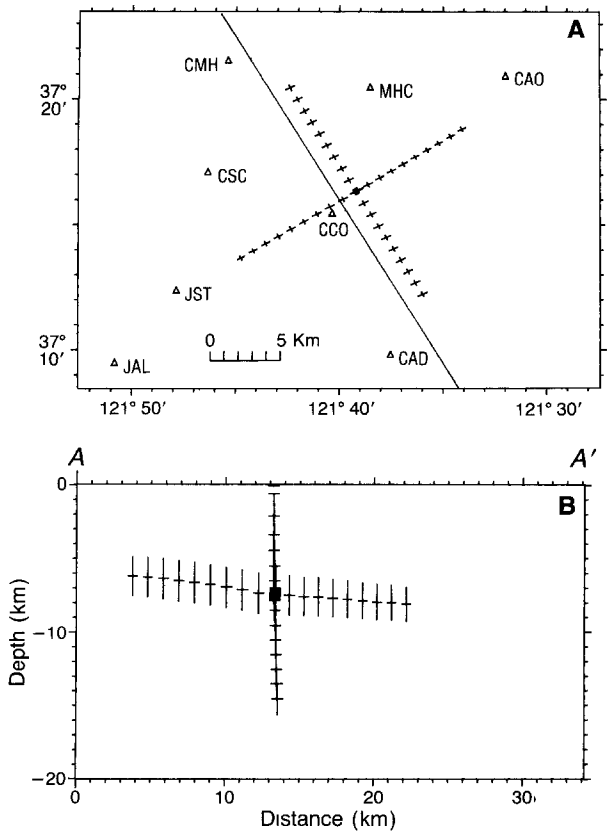

FIG. 3. Same as Figure 2, but here synthetic arrival times were calculated from model B.

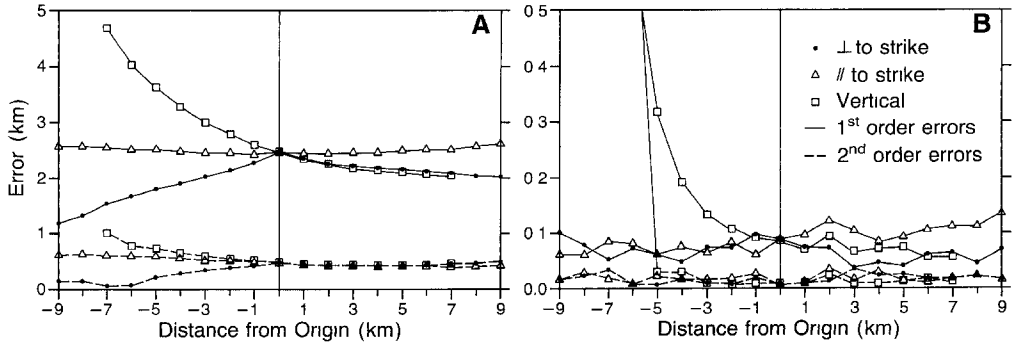

FIG. 4. Prediction errors of location estimates shown in Figures 2 and 3 for: (A) model A and (B) model B. The vertical axis in each plot is the norm of the total prediction error defined by equation (19). The first- and second-order prediction errors are as defined in the text. The actual locations of these synthetic events lie along three mutually perpendicular lines (Figure 1). To allow the results to all be plotted together, each event is referenced by its true distance from the "origin" defined as the point where all three lines intersect. The perpendicular to the strike line is defined as positive to the northeast, the parallel to the strike line is defined as positive to the southeast, and the vertical line is defined as positive downward.

imation can give a reasonable measure of the size of the nonlinear term $\boldsymbol{n}$. The level of the approximation is of the order of the convergence parameter ϵ for model B, where the hypocentral bias is comparable to what may exist in actual data from this area. For model A, the level of approximation is not as good, and we can only say that $\boldsymbol{n}_2$ is a reasonable measure of the size of $\boldsymbol{n}$.

Analyzing Different Sources of Hypocentral Error

Introduction. Having seen the limitations of the approximation given by equation (18), we can approach the practical problem of what to do with this result. The problem is that with real data the vectors $\boldsymbol{e}_{\text{model}}$, $\boldsymbol{n}$, and $\boldsymbol{e}$ are fundamentally unknown. Our only information about them is provided by a projection of the residual vector, $(\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^{+})\hat{\boldsymbol{r}}$. Unfortunately, the above analysis shows that $\hat{\boldsymbol{r}}$ is a sum of all three error terms and is evaluated at the wrong place in space-time. The basic idea here is to use auxiliary information to appraise the relative importance of each term. This provides a valuable error appraisal tool to provide a more complete and realistic appraisal of location uncertainties. I now consider methods for estimating each of these terms. The order in which they are considered is significant.

Measurement error term. Of the three terms in equation (18), $\boldsymbol{e}$ is unique in that it is the only truly statistical quantity. Thus, although $\boldsymbol{e}$ may be unknowable, we can assume we know something about its statistics [see Freedman (1966) or Leaver (1984) for examples of particularly careful studies]. For impulsive arrivals measured from analog records, the e_i are approximately normally distributed with zero mean (Buland, 1976). For emergent arrivals, the e_i tend to have a distribution skewed toward positive numbers due to a tendency to pick weak arrivals too late (Anderson, 1982). Finally, with newer computer picking methods, the distribution of e_i may be somewhat complicated, but at least the gross details of the distribution are known (Allen, 1982; Leaver, 1984). In any case, if we know something about the probability density function of $\boldsymbol{e}$, it is a standard exercise in regression analysis (see Flinn, 1965; Hoel, 1971; Jordan and Sverdrup, 1981; and their associated references) that the errors induced by $\boldsymbol{e}$ can be appraised by examination of confidence ellipsoids. For the hypocenter location problem, this amounts to outlining an ellipsoidal region in space-time characterized by

$$(\boldsymbol{h} - \hat{\boldsymbol{h}})^{T}\boldsymbol{C}_{h}^{-1}(\boldsymbol{h} - \hat{\boldsymbol{h}}) \leqq \kappa_{\alpha}^{2} \tag{20}$$

$\boldsymbol{C}_{\text{h}} = \boldsymbol{A}^{+}\boldsymbol{C}(\boldsymbol{A}^{+})^{T}$ where $\boldsymbol{C} = \text{cov}(\boldsymbol{e}) = \boldsymbol{E}[\boldsymbol{e}\boldsymbol{e}^{T}]$ is the data error covariance with $\boldsymbol{E}[\ ]$ denoting expectation value. κ_{α} is a critical value which depends upon the distribution of $\boldsymbol{e}$ and the confidence level, α, specified. For example, if the elements of $\boldsymbol{e}$ are independent samples from a normal distribution with zero mean and variance σ^2, κ_{α} for $\alpha = 95$ per cent is given by $\kappa_{\alpha}^{2} = 9.488\ \sigma^2$ where 9.488 is the 95 per cent critical level for a χ^2 distribution with 4 degrees of freedom (e.g., Hoel, 1971, p. 392).

Confidence ellipsoids constructed from equation (20) are a classical result, but there is a major difference between the general concept of equation (20) and what is conventional in seismology. Since the original work of Flinn (1965), the standard approach has been to scale the data covariance, $\boldsymbol{C}$, by the rms residuals. That is, one sets $\boldsymbol{C} = s^2\boldsymbol{I}$, where

$$s^2 = \frac{\| \hat{r}_i^{\,2} \|}{m - 4}. \tag{21}$$

Then, under an assumption of normally distributed errors,

$$\kappa_{a}^{2} = 4s^2 F_{\alpha}(4,\ m - 4) \tag{22}$$

where $F_\alpha(4, m-4)$ is an F statistic with 4 and $m-4$ degrees of freedom for the critical level α (Flinn, 1965). This almost always leads to error estimates that are considerably larger than those based on measurement errors alone. The conventional wisdom is that this is justified, as it is one way to account for what I have called $\boldsymbol{e}_{\text{model}}$ and $\boldsymbol{n}$. I hope to convince the reader in the following sections that the conventional wisdom is wrong and a different analysis is necessary.

Model error term. Unlike $\boldsymbol{e}$, it is not really proper to consider $\boldsymbol{e}_{\text{model}}$ as a statistical quantity. The failure of the statistical model to predict systematic biases in hypocenter location estimates is well documented (e.g., Bolt *et al.*, 1968; Brown and Lee, 1971; Morrison *et al.*, 1976; Uhrhammer, 1981). Nonetheless, it is useful to consider results from the synthetic data tests considered above as they demonstrate at least part of the reason for the failure of the statistical model to correctly predict errors due to $\boldsymbol{e}_{\text{model}}$.

Figure 2 shows the principle axes of the 95 per cent confidence ellipsoids defined by the standard F statistic given in (22). None of the ellipsoids enclose the true locations of any of the events examined here. This failure is not due to measurement errors as the error induced by $\boldsymbol{e}$ is less than 30 m for every event. The best explanation is that the application of equation (22) ignores a well-known caveat of the properties of the F distribution summarized in the following quote from an introductory statistics book

> Unfortunately, the preceding test (F test) is not reliable if X and Y do not possess normal distributions. Just as in the case of applying the χ^2 distribution to test a single variance, the F distribution may be highly unreliable if X and Y possess distributions whose fourth moments are considerably larger than for a normal distribution (Hoel, 1971, p. 273).

For the case considered here, that turns out to be precisely the problem. Figure 5 shows the distribution of individual components of $\hat{\boldsymbol{r}}$, $\boldsymbol{e}_{\text{model}}$, $\boldsymbol{e}$, and $\boldsymbol{n}$ for all model B events. (Model A results are similar except the time scale is roughly an order of magnitude larger. These are not shown for the sake of brevity.) If one did not look carefully at the residuals (the only information we have in real life), one might be led to think the residuals were Gaussian and jump to the conclusion that $\boldsymbol{e}_{\text{model}}$ was also Gaussian. In reality, we see the components of $\boldsymbol{e}_{\text{model}}$ have a distribution that is basically bimodal due to the fact that the data were generated from a two quarter-space model. A bimodal distribution is an example of a distribution with a large fourth-order moment. This is obscured in the distribution of $\hat{\boldsymbol{r}}$ by the added influence of $\boldsymbol{e}$ and $\boldsymbol{n}$, causing the distribution of $\hat{\boldsymbol{r}}$ to be nearly Gaussian. Nonetheless, the fourth-order moment of the distribution of $\hat{\boldsymbol{r}}$ was apparently sufficiently large to make the F distribution error ellipsoids unreliable.

A statistician would say that the error ellipsoid problem described above was caused by the violation of our assumption that the residuals were Gaussian, and we should take steps to account for the actual distribution of $\hat{\boldsymbol{r}}$. I would argue that would accomplish little and misses the basic point. That is, there are two kinds of errors in hypocenter estimates: (1) systematic biases and (2) scatter in the relative location of events. The former is dominated by $\boldsymbol{e}_{\text{model}}$. The latter is largely controlled by $\boldsymbol{e}$ and short spatial wavelength structures in $\boldsymbol{e}_{\text{model}}$, but relative location scatter can also be induced by interaction with $\boldsymbol{e}_{\text{model}}$ through variations in effective array geometry caused by arrivals not being recorded at all stations (e.g., Pavlis and Hokanson, 1985). I assert that these two kinds of errors should be appraised differently. Relative location errors can and should be appraised by error ellipsoids based on known properties of the measurement error statistics. Systematic biases

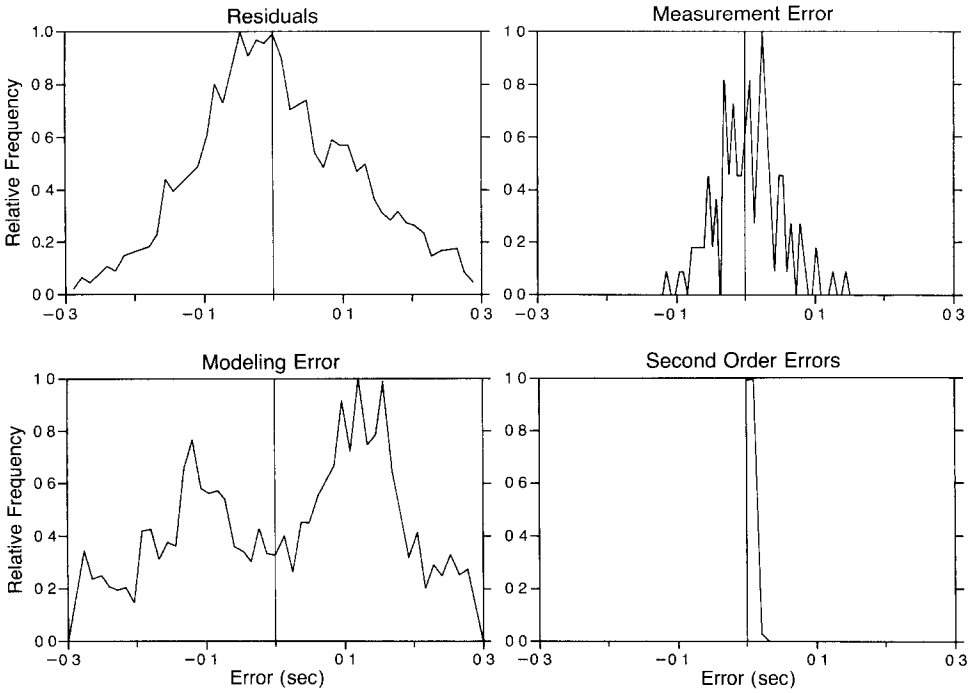

FIG. 5. Frequency distributions of different travel-time error terms from model B synthetic data. Distributions were derived from all 53 events (5413 total travel times). Each plot is normalized to peak frequency to facilitate comparisons. For the second-order, nonlinear errors, only two points fell outside in the first three bins (interval 0.01 sec). The two extremes are beyond the dynamic range of the plot and occur at times ≈0.04 and 0.06, respectively.

should be treated separately. I now advance a new methodology for bounding systematic mislocations.

To bound $\boldsymbol{e}_{\text{model}}$, we need to recognize its relation to the seismic velocity structure of the earth. Assuming ray theory and given the definition of $\boldsymbol{e}_{\text{model}}$ [equation (3)], it follows that

$$(\bar{e}_{\text{model}})_i = \int_{\Gamma_i^{\text{true}}} u_{\text{true}}\, ds - \int_{\Gamma_i^{\text{model}}} u_{\text{model}}\, ds$$

$$\approx \int_{\Gamma_i^{\text{model}}} (u_{\text{true}} - u_{\text{model}})\, ds \tag{23}$$

where u_{true} is the actual slowness as a function of position within the earth and Γ_i^{true} is the actual ray path corresponding to the phase defined as arrival i. Similarly, Γ_i^{model} is the ray path corresponding to the same arrival, but based on a model of the slowness function, u_{model}. In the earth, Γ_i^{true} and Γ_i^{model} can be expected to not differ enormously and the $\approx$ in (23) is reasonably satisfied. Since rock slowness is always finite, the integral in equation (23) can be bounded. This yields the bound

$$|(\bar{e}_{\text{model}})_i| \leqq \bar{s}_i \Delta u \tag{24}$$

where Δu is an upper bound on $|u_{\text{true}} - u_{\text{model}}|$ and $\bar{s}_i$ is the ray path length, $\int ds$, along Γ_i^{model}. Thus, each component of $\bar{\boldsymbol{e}}_{\text{model}}$ can be bounded in terms of a common

scale, Δu, and the variable $\bar{s}_i$. A theorem proven in the Appendix then gives that the error in $\boldsymbol{h}$ due to $\boldsymbol{e}_{\text{model}}$, $\delta\boldsymbol{h}_{\text{model}}$, is bounded by

$$|(\delta h_{\text{model}})_i| \leqq \Delta u \sum_{j=1}^{m} |A_{ij}^{+}| \, s_i \tag{25}$$

where $\boldsymbol{s} = \boldsymbol{W}\bar{\boldsymbol{s}}$ as in equation (10).

The bounding criteria given in (25) is applied to the data from model B in Figure 6. (Similar results were obtained for model A, but I again omit them for brevity's sake.) The bounds shown are calculated from equation (25) using $\Delta u = (5.5^{-1} - 5.6^{-1}) = 0.00325$ sec/km. Note that the estimated bounds contain the true hypocenters of the corresponding events and are not overly pessimistic for any of the spatial coordinates. This is in sharp contrast to bounds based on more conventional matrix bounding methods which I also tried. That is, given the bounds in equation (24), one could also bound $\delta\boldsymbol{h}_{\text{model}}$ as

$$\| \delta\boldsymbol{h} \| \leqq \| \boldsymbol{A}^{+} \| \, \| \boldsymbol{s} \| \, \Delta u$$

for any consistent matrix and vector norm. However, bounds calculated this way using the same Δu as above are ~60 km for L_1 and ~10 km for L_2 and L_∞, which renders them essentially useless. On the other hand, one must recognize that the test of the bounding criteria given in equation (25) is not entirely fair either. The chosen scale for Δu is a nearly perfect estimate, as it is only slightly greater than the actual upper bound (slowness differences between quarter space velocity of 5.5 km/sec and constant velocity model velocity of 5.6 km/sec). Such a precise bound was available only because these were synthetic data. With real data, how to properly choose Δu is an open question. A promising technique for estimating Δu is to use nonparametric statistical techniques such as the bootstrap, jackknife, delta, or cross-validation methods (Efron, 1981; Efron and Gong, 1983). This approach could cast Δu as a statistical quantity, but has the important advantage of not requiring any *a priori* assumptions about the distribution of Δu. The cost, however, is a very high computational overhead, which may prove impractical. This or other alternative schemes for estimating Δu are subjects for future work.

Nonlinear error term. Even in his original work on the subject, Flinn (1965) recognized the potential problem posed by nonlinearity in appraisal of hypocentral location error estimates. Flinn's ideas about how to correct error ellipsoids to account for nonlinear effects, however, have been largely ignored as far as I can judge. Other notable attempts to provide more complete error calculations that account for nonlinearity are given by Tarantola and Valette (1982) and by Rowlett and Forsyth (1984). The former is particularly novel and complete, but its practicality for routine use is questionable (Thurber, 1986). In this paper, I propose a simple, but fundamentally different method for appraising the potential influence of nonlinearity based on the second-order approximation of equation (17) and a bounding criterion similar to that used in equation (25) above.

If we knew $\delta\boldsymbol{h}$ *a priori*, equation (17) provides a means for estimating $\boldsymbol{n}$ as shown by the results in Figure 4. In real life, however, $\delta\boldsymbol{h}$ is unknown. On the other hand, we usually know at least something about its scale,

$$\rho = \| \delta\boldsymbol{x} \| . \tag{26}$$

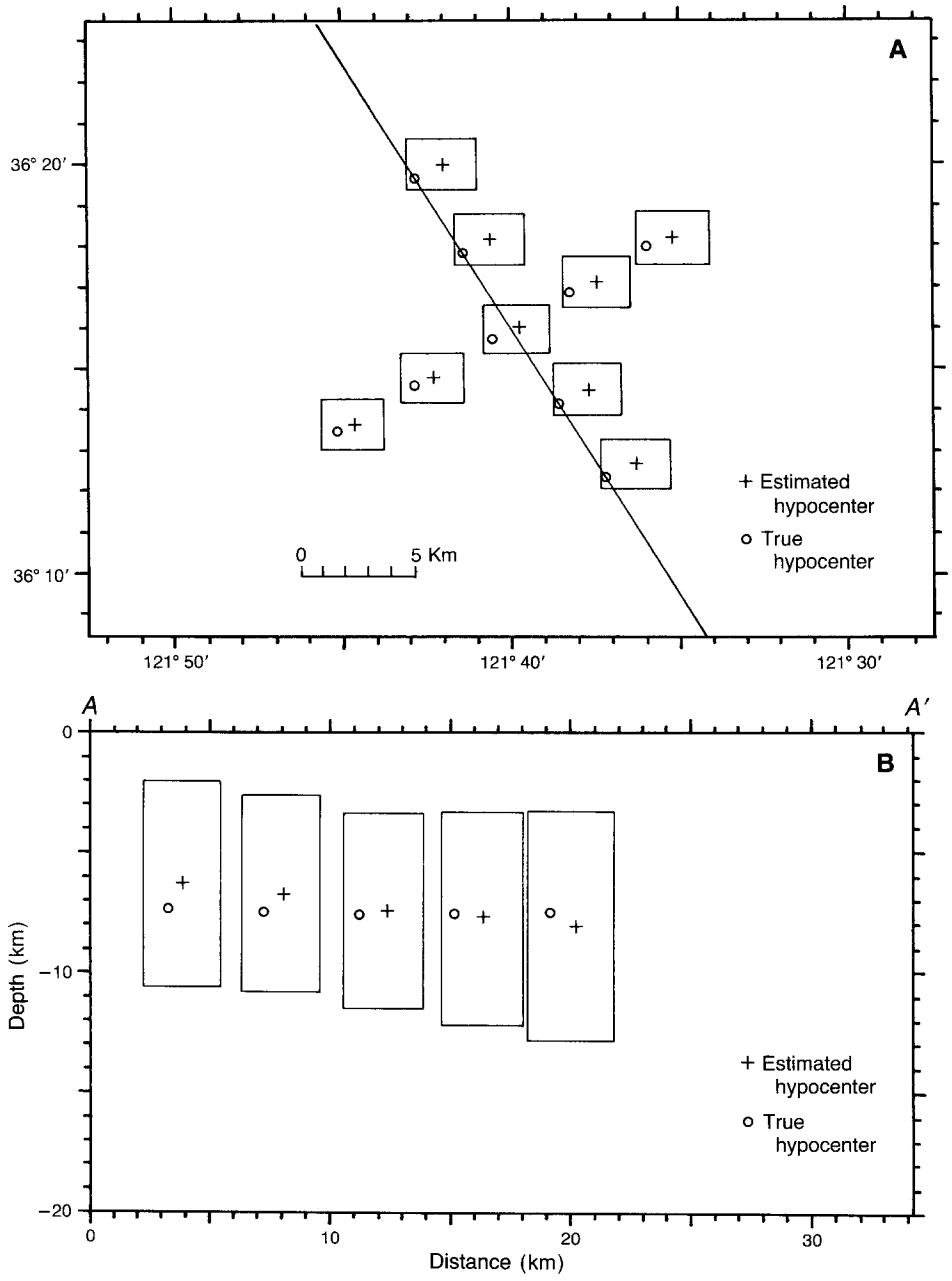

FIG. 6. Extremal bound estimates for model B results. (A) Map view and (B) cross-sections are the same frame as in Figure 1. Boxes are coordinate bounds on systematic mislocation errors caused by modeling errors. For clarity of presentation, results are only shown for every fourth event along the horizontal profiles, and no events along the vertical line are shown. Cross-sections are projections of extreme edges of these parallel-piped-shaped regions. Projected edges are omitted for clarity.

Then, it follows from equation (17) that within the limits of the second-order approximation and our imperfect knowledge of ρ, that each component of $\boldsymbol{n}$ is bounded as

$$\bar{n}_i \leqq \tfrac{1}{2}\rho^2 \| \boldsymbol{H}_i \| \tag{27}$$

where $\| \boldsymbol{H}_i \|$ implies the L_2 or spectral norm of $\boldsymbol{H}_i$ because the second-order term in equation (17) is quadratic. The scale, ρ, in equation (26) involves only the spatial coordinates, as the Hessian does not depend on the origin time (Thurber, 1985). We can then again apply the theorem in the Appendix to bound the error, $\delta \boldsymbol{h}_n$, due to nonlinearity as

$$\| (\delta h_n)_i \| \leqq \tfrac{1}{2}\rho^2 \sum_{j=1}^{m} | A_{ij}^{+} | \| \boldsymbol{H}_j \| . \tag{28}$$

These bounds, like those used for $\boldsymbol{e}_{\text{model}}$, are based on a scale factor, ρ, and a set of variable factors, $\| \boldsymbol{H}_j \|$, which vary widely for different arrivals. The scale factor is especially crucial in this case, however, as the calculated bounds depend on ρ^2. Therefore, before proceeding, it is necessary to discuss how ρ can be estimated.

Normally, it can be expected that the total error is dominated by the effects of $\boldsymbol{e}_{\text{model}}$. In this case, the most obvious estimate of ρ is $\| \delta \boldsymbol{h}_{\text{model}} \|$ where $\delta \boldsymbol{h}_{\text{model}}$ is the vector of bounds in equation (25). This will certainly yield a set of absolute bounds. Unfortunately, the results will tend to be overly pessimistic. A graphic illustration of this is shown in Figure 7. The set of bounds labeled "c" were constructed this way. Figure 7 shows only results for the y coordinate of the series of synthetic events perpendicular to the strike of the "fault" shown in Figure 1, but these results are representative. The bounds are so large that they are not only useless, but also very misleading. This occurs because we are using the square of one bound to construct another, which inevitably gets too pessimistic. I experimented with estimating ρ from rms residuals, but this had the opposite problem. We saw above that error ellipsoids scaled using rms underestimated the systematic bias in these locations. For the same reason, any estimates of ρ derived from rms according to any conventional recipe will underestimate ρ leading to an underestimate of the bounds due to $\boldsymbol{n}$.

An estimate for ρ that proved adequate, although still not entirely satisfying, was a crude guess one might be able to make with the real data that these synthetic data models represent. That is, all events are offset perpendicular to the "fault" by epicentral distances of ~9 and ~1.5 km for models A and B, respectively. If we presume this is representative for one degree of freedom of the solution, then a reasonable guess for the scale ρ is $9\sqrt{3}$ km and $1.5\sqrt{3}$ km for models A and B, respectively, since $\delta \boldsymbol{x}$ has three degrees of freedom. Using these values results in the bounds shown in Figure 7 labeled "b." These are still somewhat overly pessimistic, being approximately an order of magnitude larger than the actual error. This is largely caused by the fact that these estimates of ρ are not perfect either. Ideal values of ρ are values of ~10 and ~1.5 km for models A and B, respectively. Hence, if we knew ρ *a priori*, we would be able to reduce the bounds shown in Figure 7 by a factor of 3. In that case, these bounds would do a remarkably good job of measuring the potential impact of nonlinearity in the solution. It is important to recognize that this is a major advantage of the bounding criterion based on equation (28). That is, provided ρ can be estimated accurately by some auxiliary means, the bounds provided by (28) can be expected to be reasonable. This is again in contrast to simpler schemes I first tried using standard bounding methods based on matrix norms (i.e., $\| \delta \boldsymbol{h}_n \| \leqq \| \boldsymbol{A}^{+} \| \| \boldsymbol{n} \|$ for any pair of consistent matrix and vector norms, provided $\| \boldsymbol{n} \|$ is an upper bound on the true norm of $\boldsymbol{n}$). These always gave terrible results even when ρ was chosen exactly.

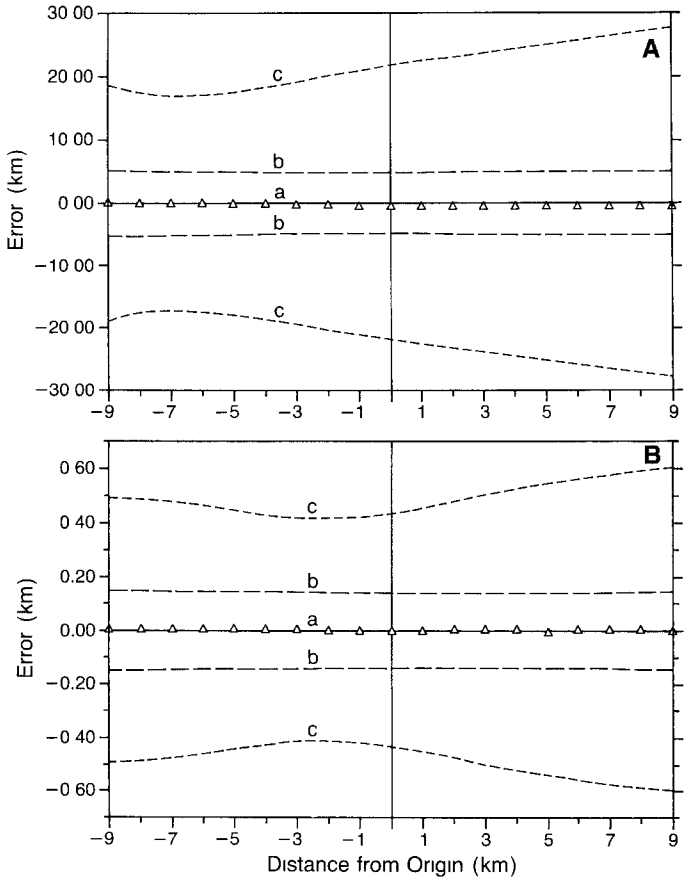

FIG. 7. Example of nonlinear error bounding estimates. (A) Model A and (B) model B. Results shown are for the east-west epicentral coordinate for the line of events oriented perpendicular to the strike (northeast to southwest) in Figure 1, which are representative. The *abscissa* is as described in Figure 4. The actual nonlinear errors are plotted as discrete points as triangles labeled a. The bounds constructed from the extremal bounds shown in Figure 6 are those labeled as c. Bounds constructed from the guess described in the text are those labeled as b. Note the extreme difference in the scale between results for models A and B.

DISCUSSION AND CONCLUSIONS

The first major result of this paper is equation (18). It is significant that this was not the result I originally expected when I began this study. My original intent was to consider the impact of the fact that the matrix of partial derivatives [$\boldsymbol{A}$ defined by equations (7) and (10)] was calculated from a model of the earth's velocity structure in the same way the travel time is. It is well known that errors in calculating coefficients of a matrix cause errors in solutions of linear equations. Such errors are commonly appraised by matrix peturbation analysis methods (for least-squares problems, see, e.g., Lawson and Hanson, 1974, pp. 41–52) to bound the possible influence of computational errors in real computers with a finite precision. The analysis leading to equation (18) shows that even though $\boldsymbol{A}$ is not calculated perfectly, it makes little difference as long as a stable solution can be found. The basic reason is that in a linear problem, $\boldsymbol{A}$ is fixed; here $\boldsymbol{A}$ is variable. We try to minimize $\| \boldsymbol{r} \|$ by a sequence of steps given in equation (6). Each step is a linear one designed to minimize the norm $\| \boldsymbol{r} - \boldsymbol{A}\delta\boldsymbol{h} \|$ based on the current values of $\boldsymbol{r}$ and $\boldsymbol{A}$, which vary from step to step. The net result is that when the solution

converges, the only relevant measure of $\boldsymbol{A}$ is the current one. Hence, the fact that it may be wrong due to inadequacies in calculating ray takeoff angles is irrelevant.

Equation (18) shows that hypocentral errors are a composite of three terms: (1) measurement error; (2) what I have called modeling errors; and (3) a nonlinear term. Of these, I claim only the first should really be viewed as a statistical quantity. This term can be appraised with confidence ellipsoids, but the scale of the process should be determined from independent studies of measurement errors such as that by Freedman (1966) or Leaver (1984). Scaling confidence ellipsoids by rms following Flinn (1965) is a dangerous step that can produce misleading results as shown by the results given in Figure 2. This occurred because the synthetic examples studied here model a feature of most real data. That is, the errors are dominated by modeling errors. In this paper, I introduced an alternative means for appraising the influence of modeling errors using a component-wise bounding criteria based on a theorem proven in the Appendix. These bounds are based on arrival-dependent ray arc lengths and a common scale factor, Δu, which is a bound on the average slowness along a given ray segment. The result is a parallelepiped-shaped bounding region for each earthquake location whose relative dimensions are fixed and whose absolute scale depends linearly on Δu.

The third source of earthquake location errors, nonlinearity, is almost always ignored in conventional location procedures. The analysis presented here indicates nonlinearity can be viewed as an additional component of the systematic location bias superimposed on top of that caused by modeling errors. Evidence from the synthetic examples presented here suggests that such errors can be approximated to adequate precision for any reasonable location estimate using a second-order approximation. Using a second-order approximation and a component-wise bounding procedure similar to that used for modeling errors, the expected size of the nonlinear error can be bounded. These bounds are constructed from the spectral norm of the Hessian for each arrival (Thurber, 1985) and a common scale factor, ρ, which is a guess of an upper bound on the total hypocentral error for that event. Again, these bounds form a parallelepiped-shaped region of space whose relative dimensions are fixed. In this case, however, the scale of the bounding region is determined by the square of the scale factor ρ. Consequently, the bounds estimated this way are reasonable only when ρ is chosen reasonably.

How to best estimate the scale factors Δu and ρ used in the bounding procedures described here is an open question. Δu is probably best estimated by nonparametric statistical methods (Efron and Gong, 1983) or from a random media viewpoint as advocated by Leaver (1985). The best approach for ρ in many cases is probably to make a reasonable guess of its size based on other information. For example, if a group of earthquakes appear to be systematically offset a distance d from a mapped fault trace, ρ could be estimated from d as described above. Lacking such information, ρ may be estimated from some measure based on rms residuals or the extremal bounds on systematic biases described here. In doing so, one must recognize, however, that the former may underestimate ρ, and the latter will always overestimate ρ. In any case, one should recognize the main advantage of a set of bounds based on a common scale factor. If one decides the original scale factors used to calculate the bounds were in error, recalculating them based on the revised scale is trivial.

I would claim that one of the most significant facts about the error appraisal techniques described here is their practicality. The error estimates are easy to calculate, and the results are easy to understand at a glance. Furthermore, imple-

menting them requires only minor modifications to most location programs. Rescaling confidence ellipses to reflect only measurement errors is trivial and is already an option in at least one common location program I am aware of (HYPOINVERSE, Klein, 1978). Calculating unscaled model and nonlinear bounds requires only a minor modification to any program that explicitly calculates $\boldsymbol{A}^{+}$. One then only has to calculate the vector of component bounds defined in equations (24) and (27). The model error bound estimate requires the calculation of ray arc lengths. For local networks, this calculation can probably be approximated adequately as the total source receiver spatial separation. For teleseismic locations, it would probably require a table. The nonlinear error bound calculation requires calculation of the spectral norm of the Hessian matrix for each arrival. Analytic forms are presently known only for constant velocity media and a layer over a half-space model (Thurber, 1985). These second-order derivatives could presumably be calculated numerically for any arbitrary model, but I expect that is probably unnecessary in most cases. Thurber points out that the Hessian gives a measure of local wave front curvature. My experience from this work is that this causes the nonlinear errors to be dominated by the one or two nearest stations to the source where the wave front curvature is largest. At nearby stations, the wave fronts will not differ that dramatically from a constant velocity medium. Therefore, I suspect the use of Hessians appropriate to constant velocity media would normally give reasonable results.

Finally, it is important to stress the two limitations of the techniques discussed here. First, the error analysis presented here is based on the fundamental assumption that the location estimate is well constrained. That is, I assumed no auxiliary constraint like fixing the depth is necessary to obtain a stable solution. In the case of poorly constrained locations, one probably should resort to the techniques of Tarantola and Valette (1982) or Rowlett and Forsyth (1984). Second, this error analysis is what might be called a single-event theory. That is, each event is treated as an independent entity, which is an unstated, implicit assumption in most routine locations. In future work, I hope to extend these ideas to measures of relative location precision of a set of earthquake locations. That problem is complicated by the fact that different earthquakes will, in general, be recorded by different stations with different levels of precision for corresponding arrivals. How this data heterogeneity interacts with the biasing influence of modeling errors is a nontrivial question.

ACKNOWLEDGMENTS

Special thanks are due to Anne Turner for handling the seemingly endless generations of this paper that emerged from the word processor and to Kim Sowder for turning my crude computer graphics into works of art. I thank Robert Uhrhammer and Cliff Thurber for providing very constructive comments. I thank Robert Uhrhammer especially for suggesting the possibilities of nonparametric statistical methods. This work was sponsored by the National Science Foundation under Grant EAR83-19418 and by the U.S. Geological Survey under INTER 14-8-1-G1084.

REFERENCES

Allen, R. (1982). Automatic phase pickers: their present use and future prospects, *Bull. Seism. Soc. Am.* **72,** S225–S242.

Anderson, K. R. (1982). Robust earthquake location using *M*-estimates, *Phys. Earth Planet. Interiors* **30,** 119–130.

Bolt, B. A., C. Lomnitz, and T. V. McEvilly (1968). Seismological evidence on the tectonics of central and northern California and the Mendocino escarpment, *Bull. Seism. Soc. Am.* **58,** 1725–1768.

Brown, R. D., Jr. and W. H. K. Lee (1971). Active faults and preliminary earthquake epicenters (1969–1970) in the southern part of the San Francisco Bay region, *U.S. Geol. Surv. Map MF-307.*

Buland, R. (1976). The mechanics of locating earthquakes, *Bull Seism. Soc. Am.* **66,** 173–187.
Cockerham, R. S. and J. P. Eaton (1984). The April 24, 1984, Morgan Hill earthquake and its aftershocks: April 24 through December 30, 1984, in *The 1984 Morgan Hill, California Earthquake, California Division of Mines and Geology, Spec. Publ. 68,* J. Bennett and R. Sherburne, Editors, Sacramento, California.
Dewey, J. W. (1971). Seismicity studies with the method of joint hypocenter determination, *Ph.D. Dissertation,* University of California, Berkeley, California.
Dewey, J. W. (1972). Seismicity and tectonics of western Venezuela, *Bull. Seism. Soc. Am.* **62,** 1711–1751.
Douglas, A. (1967). Joint epicentre determination, *Nature* **215,** 47–48.
Efron, B. (1981). Nonparametric standard errors and confidence intervals, *Canad. J. Statist.* **9,** 139–172.
Efron, B. and G. Gong (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation, *Am. Statist.* **37,** 36–48.
Flinn, E. A. (1965). Confidence regions and error determinations for seismic event location, *Rev. Geophys.* **3,** 157–185.
Freedman, H. (1966). The "little variable factor": a statistical discussion of the reading of seismograms, *Bull. Seism. Soc. Am.* **56,** 593–604.
Herrmann, R. B. (1979). FASTHYPO—A hypocenter location program, *Earthquake Notes* **50,** 25–37.
Hoel, P. G. (1971). *Introduction to Mathematical Statistics,* Wiley, New York, 409 pp.
Jordan, T. H. and K. A. Sverdrup (1981). Teleseismic location techniques and their application to earthquake clusters in the south-central Pacific, *Bull. Seism. Soc. Am.* **71,** 1105–1130.
Klein, F. W. (1978). Hypocenter location program HYPOINVERSE. Part 1. User's guide to versions 1, 2, 3, and 4, *U S. Geol. Surv., Open-File Rept. 78-694, 113 pp.*
Lawson, C. H. and R. J. Hanson (1974). *Solving Least Squares Problems,* Prentice-Hall, Englewood Cliffs, New Jersey, 340 pp.
Leaver, D. S. (1984). Mixed stochastic and deterministic modeling of the crustal structure in the vicinity of Mount Hood, Oregon, *Ph.D. Dissertation,* University of Washington, Seattle, Washington.
Lee, W. H. K. and J. C. Lahr (1975). HYPO71 (revised): a computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes, *U.S. Geol. Surv., Open-File Rept. 75-311.*
Lee, W. H. K. and S. W. Stewart (1981). Principles and applications of microearthquake networks, in *Advances in Geophysics,* Academic Press, New York, 293 pp.
Morrison, P. W., B. W. Stump, and R. Uhrhammer (1976). The Oroville earthquake sequence of August 1975, *Bull. Seism. Soc. Am.* **66,** 1065–1084.
Pavlis, G. L. and N. B. Hokanson (1985). Separated earthquake location, *J. Geophys. Res.* **90,** 12777–12789.
Rowlett, H. and D. W. Forsyth (1984). Recent faulting and microearthquakes at the intersection of the Vema fracture zone and the Mid-Atlantic Ridge, *J. Geophys. Res.* **89,** 6079–6094.
Stewart, G. W. (1973). *Introduction to Matrix Computations,* Academic Press, New York, 441 pp.
Tarantola, A. and B. Valette (1982). Inverse problems = quest for information, *J. Geophys.* **50,** 159–170.
Thurber, C. H. (1985). Nonlinear earthquake location: theory and examples, *Bull. Seism. Soc. Am.* **75,** 779–790.
Thurber, C. H. (1986). Analysis methods for kinematic data from local earthquakes, *Rev. Geophys.* **24** (in press).
Uhrhammer, R. A. (1981). Northern and central California earthquakes, 1979, in *United States Earthquakes, 1979,* C. W. Stover, Editor, U.S. Geol. Survey, Golden, Colorado, and National Oceanic Atmospheric Administration, Boulder, Colorado, 106–108.

DEPARTMENT OF GEOLOGY
INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47405

Manuscript received 25 March 1986

APPENDIX

The component-wise error bounding criteria developed in this paper depends upon the following theorem. The proof is analogous to the consistency proof for the L_1 and L_∞ matrix norms given in Stewart (1973, p. 179). The interested reader should compare the two to see why component-wise bounding leads to a less pessimistic bond than a simple one based on L_1.

Theorem: Given $\boldsymbol{A} \in R^{m \times n}$ and a vector $\boldsymbol{x} \in R^n$ whose components can be bounded as

$$|x_j| \leqq b_j, \tag{A1}$$

then the components of the vector $y \in R^m$ defined as $\boldsymbol{y} = \boldsymbol{Ax}$ are bounded by

$$|y_i| \leqq \sum_{j=1}^{n} |A_{ij}| b_j. \tag{A2}$$

Proof: The statement $\boldsymbol{y} = \boldsymbol{Ax}$ means

$$y_i = \sum_{j=1}^{n} A_{ij} x_j. \tag{A3}$$

Thus,

$$|y_i| = \left| \sum_{j=1}^{n} A_{ij} x_j \right| \leqq \sum_{j=1}^{n} |A_{ij}| |x_j|.$$

Using (A1), it follows directly that

$$|y_i| \leqq \sum_{j=1}^{n} |A_{ij}| b_j.$$