STATION CORRECTION UNCERTAINTY IN MULTIPLE EVENT LOCATION ALGORITHMS AND THE EFFECT ON ERROR ELLIPSES

Jason P. Erickson,[1] Deborah K. Carlson,[2] Anne M. Ortiz,[1] Kevin D. Hutchenson.[1] Linda J. Owiesny,[1] Gordon D. Kraft[1], Dale N. Anderson,[2] Mark A. Tinker[1]

Quantum Technology Services, Inc.,[1] Pacific Northwest National Laboratory[2]

Sponsored by Air Force Technical Applications Center

Contract No. F08650-01-C-0063

ABSTRACT

Accurate location of seismic events is crucial for nuclear explosion monitoring. Of equal importance is an accurate representation of the location uncertainty, as this determines the area, to a certain confidence level, within which the event may be located.

There are many sources of error present in the location process, including measurement error and model bias. In a multiple event process error sources expand to include uncertainty in the station corrections and uncertainty in the location of calibration events. In order to generate meaningful error ellipses, all of these sources of error must be taken into account.

Most location algorithms calculate ellipses using some combination of *a priori* knowledge of the data variance, signal-to-noise weighting, and *a posteriori* measurement of the data residuals. Multiple event algorithms can calculate, and kriged station correction surfaces include, an estimate of the station correction uncertainty. We seek a unified general method for combining the station corrections and uncertainties from different sources and calculating an accurate error ellipse. Furthermore, in some situations it may be desirable to utilize calibration events that are known with a varying degree of accuracy, and in this case the influence of the calibration events should be weighted relative to the confidence in their locations.

We have derived the full covariance matrix for a multiple event location using Fisher's information theory (Lehman, 1983). This formulation of the covariance matrix is generally not used in seismic location, but is common theory for maximum likelihood estimation. The information theory covariance matrix gives complete control over the variance of all of the model parameters. This allows for station corrections from many approaches to be included in the location model and for calibration events with different confidence levels to be given different weighting. We examine the behavior of error ellipses calculated with this covariance with known sources of error. Progressive Multiple Event Location (PMEL, Pavlis and Booker, 1983) is used as a maximum likelihood estimate solver, and is also used to demonstrate the construction of a covariance matrix with the PMEL estimates.

OBJECTIVE

Error ellipses describe the area, to a given confidence level, within which an event is likely to have occurred. There are numerous methods for determining the size of an ellipse, but the shape and orientation are always calculated from the covariance matrix of the model parameters. The purpose of this research is to develop an alternate method of calculating the location covariance matrix. This method is intended to generalize the error ellipse construction by using a formulation of the covariance matrix that allows for any source of error or uncertainty to be incorporated into the location model.

RESEARCH ACCOMPLISHED

Introduction - PMEL

Traditional location methods seek to find a hypocenter or, in the case of multiple event algorithms, a set of hypocenters that minimize the difference between the observed and predicted arrival times (the travel-time residual). A common model for the predicted arrival time is (unweighted):

$$t_{ij} = \tau_i + T(x_i, y_i, z_i, R_j) + S_j + \varepsilon_{ij}, \qquad \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2) \tag{1}$$

where

t_{ij} is the arrival time of the ith event recorded at the jth station,

τ_i is the origin time of the ith event,

x_i, y_i and z_i are the longitude, latitude, and depth, respectively, of event i,

R_j is the location of station j,

$T(\mathbf{x}_i, y_i, z_i, R_j)$ is the travel time from event i to station j through an arbitrary earth model,

S_j is the correction for the jth station, and

ε_{ij} is the measurement error of the ith event at the jth station.

As Equation (1) is non-linear with respect to the hypocenter, the problem is usually linearized by taking a Taylor series expansion and dropping a second order and higher terms. The residual for a given observation then becomes approximately equal to:

$$r_{ij} = \Delta\tau_i + \left.\frac{\partial T}{\partial x}\right|_{x_i} \Delta x + \left.\frac{\partial T}{\partial y}\right|_{y_i} \Delta y + \left.\frac{\partial T}{\partial z}\right|_{z_i} \Delta z + \Delta S_j, \tag{2}$$

and the residuals for all of the observations can be written in matrix form as:

$$\mathbf{r} = \mathbf{A}\Delta\mathbf{x} + \mathbf{S}\Delta\mathbf{s}. \tag{3}$$

PMEL (Pavlis and Booker, 1983) solves for the hypocenter parameters and the station corrections separately through two least squares inversions. In the first inversion, hypocenters are calculated while the station corrections are held constant. The hypocenter term in Equation (3) is then removed using an annulling transformation, resulting in a set of equations relating station corrections to a set of annulled data. The station corrections are then calculated via the second least squares inversion. These station corrections are then substituted into the hypocenter inversion, and the process repeats until the convergence criteria have been met.

The covariance matrix for PMEL is given by:

$$\mathbf{Cov} = \begin{bmatrix} \mathbf{HH}^T + \mathbf{HSC}_S\mathbf{S}^T\mathbf{H}^T & -\mathbf{HSC}_S \\ -\mathbf{C}_S\mathbf{S}^T\mathbf{H}^T & \mathbf{C}_S \end{bmatrix} \qquad (4)$$

where

H is the generalized inverse of the **A** matrix in Equation (3),

S is the matrix of station correction partial derivatives, and

$\mathbf{C}_s$ is the covariance matrix of the station corrections.

<u>Information Theory Covariance Matrix</u>

If we rewrite Equation (1) in terms of the errors:

$$\varepsilon_{ij} = t_{ij} - \tau_i - T(x_i, y_i, z_i, r_j) - S_j\,, \qquad (5)$$

the resulting density function of the model is:

$$f_\varepsilon = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}\varepsilon^2_{ij}\right). \qquad (6)$$

The joint density is the product of the individual densities:

$$j = \prod f_\varepsilon(\varepsilon_{ij}) \qquad (7)$$

and the log-likelihood function is:

$$l = -N\ln(\sigma) - \frac{N}{2}\ln(2\pi) - \frac{1}{2\sigma^2}\sum \varepsilon^2_{ij}\,, \qquad (8)$$

where N is the total number of observations. Maximum likelihood estimates are values of σ, τ, x_i, y_i, z_i, and S_i that maximize equation (8).

The information matrix is defined as the expectation matrix of negative second partial derivatives of the log-likelihood function with respect to each model parameter pair (Lehmann, 1983):

$$\mathbf{I}(\theta)_{pq} = -E\left[\frac{\partial^2 l}{\partial\theta_p \partial\theta_q}\right] \quad p = 1,2,\ldots k; q = 1,2,\ldots k \qquad (9)$$

where θ is the vector of all model parameters (of length k), including the hypocenter parameters for all events and the station corrections.

And finally, the covariance matrix is the inverse of the information matrix:

$$Var(\hat{\theta}) = \left[I(\theta)\right]^{-1}. \qquad (10)$$

In practice, parameter values in $I(\theta)$ are replaced with the maximum likelihood estimates. This matrix contains the variance and covariance of all of the model parameters and its construction provides the framework for the development of covariance matrices when enhancements are made to the fundamental travel time model. Using this covariance matrix it is then possible to calculate confidence ellipses (Flinn, 1965), coverage ellipses (Evernden, 1969), or K-weighted ellipses (Bratt and Bache, 1988). In this report, we demonstrate only coverage ellipses.

Data

Results of each method are demonstrated using two different data sets. The first is a synthetic data set that has had normally distributed random measurement error and model bias added to all of the arrivals. The stations associated with these data are azimuthally distributed at increments of 45 degrees. Four of the stations are located at a distance of 20 degrees from the events, while the other four are at a distance of 35 degrees (Figure 1). Each of the nine events is located at a distance of 0.2 degrees from the adjacent events, and all events were “recorded” by every station (Figure 2). These data will be used to demonstrate the behavior of the ellipses under idealized conditions with known errors.

Ellipses are also shown for a set of mining explosions (Table 1) from the Powder River region in Wyoming. While ground-truth for these regionally recorded events was unavailable, arrivals at a local station provided high confidence correlation of each explosion to a unique mine site.

In order to simulate a sparse network configuration that could be available in a low-yield nuclear monitoring situation, a subset of three stations was used for the locations and ellipse calculations (Figure 3). These stations were located at distances of 3 degrees (PDAR), 4 degrees (ISCO), and 11 degrees (FLAR). The phases used in the locations were Pn and Sn for PDAR, and Pn for ISCO and FLAR. To reduce the variability of the measurement error, the picks were made after the waveforms had been correlated.

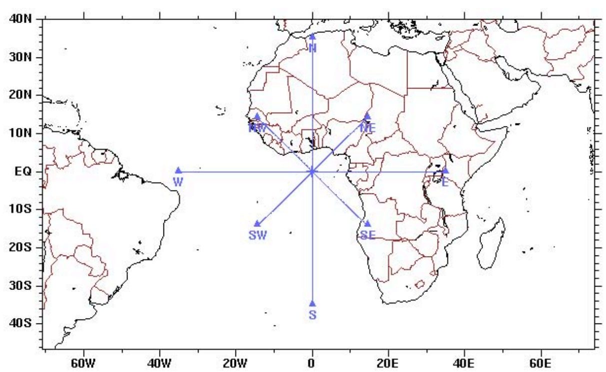

Figure 1. Station configuration of synthetic data.

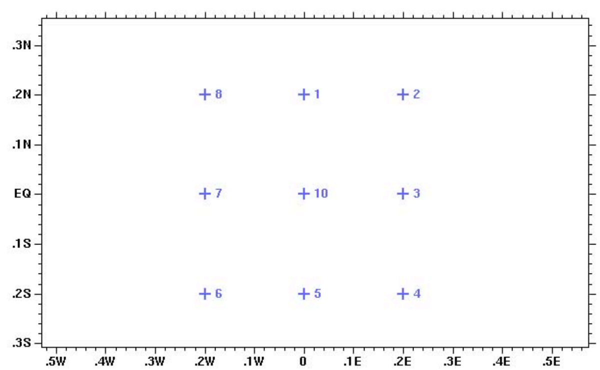

Figure 2. Event configuration of synthetic data.

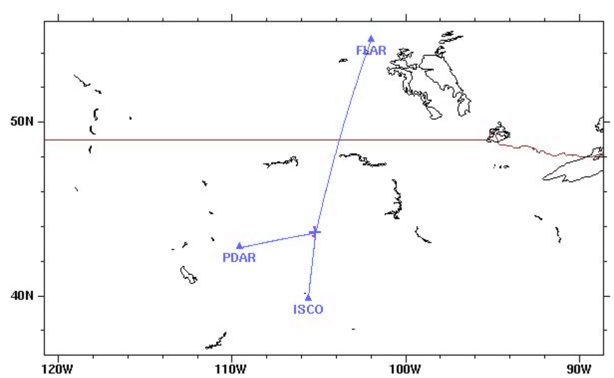

Figure 3. Stations used to calculate locations and ellipses for Powder River mining events.

Table 1. Powder River mining event origin information.

Date	Origin Time	Latitude (N)	Longitude (W)	Depth	M_b
11/30/02	20:00:01.775	43.9796	105.3221	0	3.46
12/06/02	19:07:50.251	43.6259	105.1642	0	3.78
12/09/02	21:19:45.498	43.9870	105.3221	0	3.16
12/13/02	22:09:44.746	43.6386	105.1989	0	3.60
12/14/02	21:06:33.249	43.6246	105.1828	0	3.72
12/20/02	19:17:18.226	43.6224	105.2024	0	3.78

Results

Locations and their corresponding ellipses were calculated with all of the synthetic data having a measurement error variance of 0.3 seconds and with full station coverage (Fig. 4). In this case, both methods give nearly identical results. The only difference is that the information theory derived ellipses are marginally smaller.

A similar set of locations and ellipses were calculated, with the exception that the measurement error for arrivals recorded at stations N, NE, S, and SW was increased to a variance of 0.9 seconds, while those for E, SE, W and NW remained at 0.3. The two methods again produce nearly identical ellipses, all of which are nearly circular (Figure 5), despite the fact that the station corrections to the north, northeast, south, and southwest are generated with data that have a much greater variability than the other stations.

Ellipses have also been calculated using data from mining events located in the Powder River area of Wyoming (Figure 6). With a small number of observations, ellipse calculations can become unstable, sometimes resulting in unreasonably large error regions (as is the case for one of the ellipses in Figure 6). The information theory derived covariance matrix will likely increase the stability of ellipse calculations in the sparse data case providing a more accurate representation of the location uncertainty.

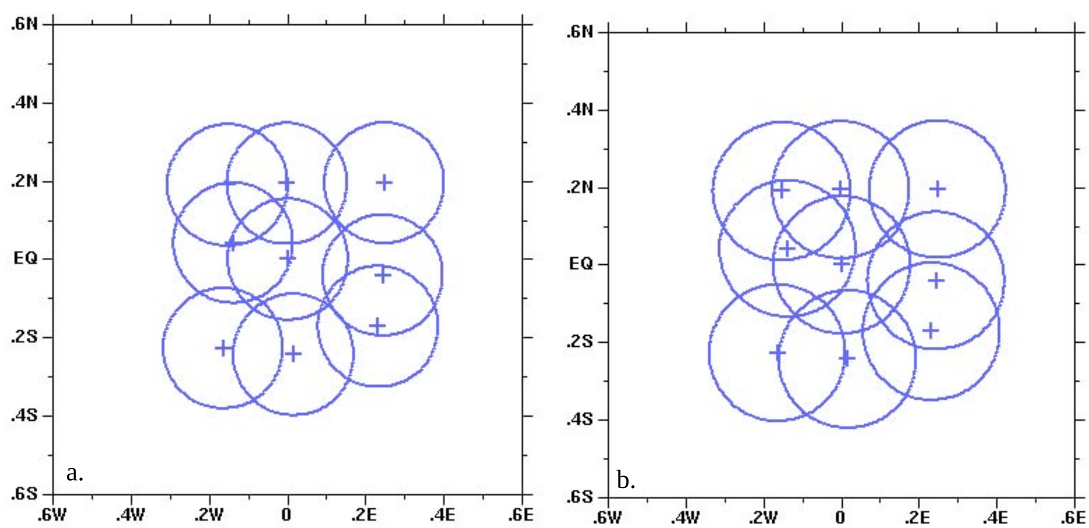

Figure 4. Coverage ellipses with full station coverage and equal random error at all stations using PMEL (a) and information theory (b).

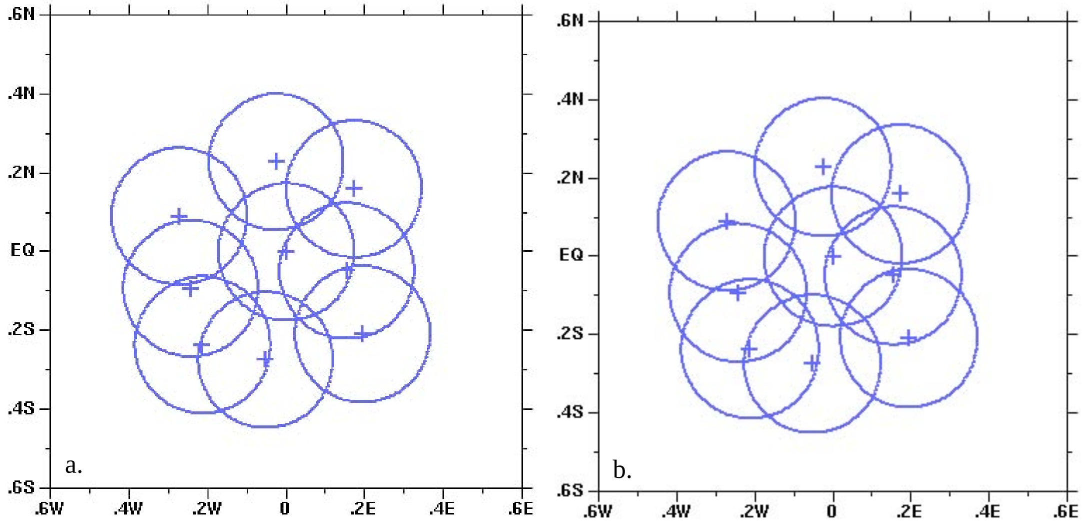

Figure 5. Locations and error ellipses from PMEL (a) and information theory (b). The stations to the north, northeast, south, and southwest all have random error with a variance of 0.9 seconds, while the others have random error with a variance of 0.3 seconds. Despite the difference in this error, both methods result in circular ellipses.

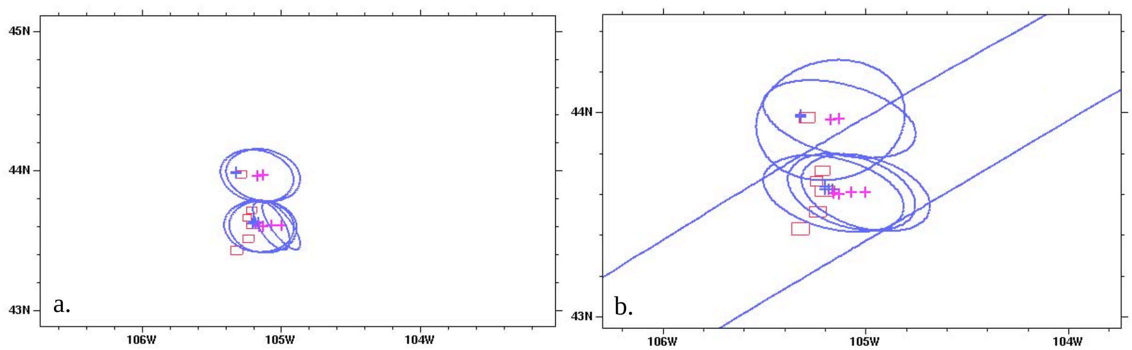

Figure 6. Locations and ellipses for six Powder River mining explosions (seed location – stars, PMEL locations – crosses). The events were located using a calibration event and four arrivals for each event. Due to the small amount of data and limited azimuthal coverage, the resultant error ellipses from PMEL (a) are unstable, while those computed with the information theory (b) covariance matrix are all of similar size.

Discussion

The information theory formulation of a covariance matrix allows for enhanced flexibility in the way that error ellipses are calculated. This work is progressing to more sophisticated formulations of the travel time model.

Information theory covariance construction provides the ability to incorporate a number of enhancements, including station corrections that are not constant, but rather are a function of other parameters. Presently, station corrections are held constant for all events in an inversion. This is based on the assumption that clustered events share the same path to the receiver, and thus the same velocity model deficiencies. As the distance between events increases, this assumption becomes less valid. Allowing station corrections to vary spatially according to some function would allow for a more realistic representation of station corrections and would allow for more distant events to be included in the inversion.

It is also possible to have the station correction variability directly affect the size and shape of the ellipse. Currently, the station correction covariance matrix (C_s from Equation 4) is solely a function of azimuthal coverage and the number of events recorded by a given station. This matrix is unaffected by how well the data are fit by the station corrections, and hence the error ellipses do not adequately represent the fact that some station corrections may be known to a higher degree of confidence than others. This explains why the ellipses in Figure 4 are circular, despite the fact that there is considerably more uncertainty in the stations corrections to the northwest and southeast of the event cluster. By adjusting the hypocenter variances and covariances to account for this uncertainty, a more accurate error region can be obtained.

The information theory covariance matrix provides the ability to use station corrections from different sources (*i.e.* PMEL and kriged correction surfaces) and include their respective uncertainties. There may be instances where a station or stations cannot be included in multiple event processing due to an insufficient number of observations. It may still be desirable to include these arrivals, particularly if they provide additional azimuthal or distance coverage to the network. As currently coded, PMEL allows for the use of observations recorded at stations that do not have corrections. The information theory covariance matrix will allow for station corrections from other sources to be applied to these observations, and will allow for the uncertainties associated with these corrections to be incorporated into the error ellipses, and

Finally, it will be possible to use calibration events with different degrees of confidence (*i.e.* GT5, GT10), and have the error ellipses reflect that uncertainty. This ability would be of particular value in locations where no previous nuclear testing had occurred, and only earthquakes were available for calibration.

CONCLUSIONS AND RECOMMENDATIONS

The information theory covariance matrix has been derived for the basic location model with station corrections and used to calculate error ellipses around seismic locations. This will provide the framework for determining the covariance matrix for a more complicated travel time model and will allow additional information to be incorporated into the uncertainty ellipses, thereby improving their accuracy.

ACKNOWLEDGEMENTS

The authors thank Drs. Gary Pavlis, David Russell, and Terry Wallace for their helpful discussion.

REFERENCES

Bratt, S. R. and T.C. Bache (1988), Locating Events with a Sparse Network of Regional Arrays, *Bull. Seism. Soc. Am.*, 78, 780-798.

Evernden, J. F. (1969), Precision of Epicenters Obtained by Small Numbers of World-Wide Stations, *Bull. Seism. Soc. Am.*, 59, 1365-1398.

Flinn, E. A. (1965), Confidence Regions and Error Determination for Seismic Event Location, *Reviews of Geophysics*, 3, 157-185.

Lehmann, E. L. (1983), Theory of Point Estimation, John Wiley & Sons, Inc., New York.

Pavlis, G. and J. Booker (1983), Progressive Multiple Event Location, *Bull. Seism. Soc. Am.*, 73, 1753-1777.