A NEW GRID-SEARCH MULTIPLE-EVENT LOCATION ALGORITHM AND A COMPARISON OF METHODS

William Rodi,[1] E. Robert Engdahl,[2] Eric A. Bergman,[2] Felix Waldhauser,[3]
Gary L. Pavlis,[4] Hans Israelsson,[5] James W. Dewey,[6] and M. Nafi Toksöz[1]

Massachusetts Institute of Technology,[1] University of Colorado at Boulder;[2]
Lamont-Doherty Earth Observatory;[3] Indiana University;[4]
Center for Monitoring Research;[5] U.S. Geological Survey[6]

Sponsored by Defense Threat Reduction Agency

Contract No. DTRA01-00-C-0102

ABSTRACT

Multiple-event location methods solve jointly for the location parameters (hypocenters and origin times) of seismic events in a cluster and travel-time corrections at the stations recording the events. This paper describes a new such method based on a grid-search approach. The algorithm (called GMEL) is an extension of an earlier single-event location algorithm (GSEL) and is based on a maximum-likelihood formulation of the location problem. By employing grid-search and root-finding techniques in lieu of a linearized inversion (Geiger's) method, GMEL accommodates non-Gaussian as well as Gaussian models of picking error and handles prior constraints such as the ground truth (GT) level of previously well-located events. We describe the algorithm and report on our efforts to extend our Monte Carlo technique for computing confidence regions on event locations, currently implemented in GSEL, to the multiple-event location problem. This extension will yield confidence regions that quantify the effects on location error of both picking errors and the uncertainty in estimated station corrections, i.e. calibration (or modeling) errors.

We report preliminary results of a test that compares GMEL and four other multiple-event location techniques: hypocentral decomposition (HDC), double differencing (DD), progressive multiple-event location (PMEL), and joint hypocenter determination (JHD). The test is being performed on two adjacent event clusters from the 1999 Izmit/Duzce earthquake sequence, previously analyzed by Engdahl and Bergman (2001) with the HDC method. We are independently applying the other methods to the data they used in an attempt to improve our understanding of the similarities and differences among these techniques and the effectiveness of the multiple-event location approach in general.

OBJECTIVE

Under the contract referenced above, we are developing a methodology for seismic event location that provides improved estimates of location uncertainty that can be used more reliably in the monitoring of small events using sparse seismic arrays. Our approach is based on a general, maximum-likelihood formulation of the location problem and implemented with computational tools such as grid-search and Monte Carlo simulation. With this approach we have been able to address such complexities as nonlinearity of the forward problem and non-Gaussian data errors.

A major objective of our project is to incorporate a more rigorous treatment of the errors in travel-time models (modeling errors) into our uncertainty analysis. Towards this goal we are attempting to link our uncertainty analysis to the calibration process, as performed with data from multiple events. Our first step has been to extend our single-event grid-search algorithm (GSEL) to a multiple-event one (GMEL). The basic task of GMEL is performed, in the Gaussian error case, by several other multiple-event location methods stemming back to the early work of Douglas (1967), Lilwall and Douglas (1970) and Dewey (1971). One part of this paper reports some collaborative work we are doing to compare these other methods and GMEL theoretically and in practice. The remainder of the paper describes GMEL itself and the approach we are developing to use it as a means of including the effects of modeling errors in location confidence regions.

RESEARCH ACCOMPLISHED

Multiple Event Location Problem

GMEL processes arbitrary data sets of seismic arrival times, azimuths and slownesses observed from multiple events for multiple stations and phases. To simplify the discussion here, we consider data sets comprising only arrival times.

We consider arrival times from m seismic events and n seismic stations. We denote the origin time and hypocenter of the jth event as t_j and $\mathbf{x}_j$, respectively. As a further simplification for this write-up, we assume there is at most one arrival time observed for each event-station pair, and denote this time as d_{ij} for the ith station and jth event. It is understood that d_{ij} exists only for a subset of the mn possibilities. The multiple-event location problem can then be written as

$$d_{ij} = t_j + T_i(\mathbf{x}_j) + c_{ij} + e_{ij} \,. \qquad (1)$$

Here, T_i is a known travel-time function for the ith station, obtained from an assumed Earth model; c_{ij} is an unknown correction to this function that accounts for differences between the real Earth and the assumed model; and e_{ij} is an observational ("picking") error.

The multiple-event location problem is to solve equation (1) for the location parameters ($\mathbf{x}_j$, t_j) of the m events jointly with the path corrections (c_{ij}). However, with no constraints on the path corrections, this problem is hopelessly ill posed. Therefore, GMEL and many previous methods assume that the path corrections at a given station are the same for all events, which is tantamount to assuming that the events are in a small cluster. In this case, we can replace c_{ij} with c_i and re-write (1) as

$$d_{ij} = t_j + T_i(\mathbf{x}_j) + c_i + e_{ij} \,. \qquad (2)$$

The problem unknowns are now the $\mathbf{x}_j$, t_j and c_i. The c_i are sometimes referred to as station corrections instead of path corrections, but it is important to realize they are dependent on the location of the event cluster.

GMEL

GMEL addresses the problem stated in equation (2) (with event-independent travel-time corrections) and solves jointly for the corrections and event location parameters. The error assumptions used are identical to those in GSEL (see Rodi and Toksöz, 2001), whereby each e_{ij} has a generalized Gaussian distribution of order p (p = 2 for Gaussian). In GMEL, the data standard errors, s_{ij}, are assumed known in a relative sense, such that

$$s_{ij} = s_i n_{ij} \tag{3}$$

where the n_{ij} are known, but the station-dependent scale parameter, s_i are not. The error model implies a likelihood function, L, given by

$$-\log L = \text{const} + \sum_i n_i \log s_i + \Psi(\mathbf{x}_1, t_1, \mathbf{x}_2, t_2, \ldots, c_1, c_2, \ldots) \tag{4}$$

where n_i is the number of events recorded at the ith station, and Ψ is a data misfit function. We can write Ψ as either a sum of station-specific or event-specific misfit functions, i.e.

$$\Psi = \sum_j \Psi_j^{\text{ev}}(\mathbf{x}_j, t_j, c_1, c_2, \ldots) = \sum_i \frac{1}{(s_i)^p} \Psi_i^{\text{st}}(\mathbf{x}_1, t_1, \mathbf{x}_2 t_2, \ldots, c_i) \tag{5}$$

with

$$\Psi_j^{\text{ev}}(\mathbf{x}_j, t_j, c_1, c_2, \ldots) = \sum_i \frac{1}{(s_i)^p} | d_{ij} - t_j - T_i(\mathbf{x}_j) - c_i |^p / (n_{ij})^p \tag{6}$$

$$\Psi_i^{\text{st}}(\mathbf{x}_1, t_1, \mathbf{x}_1, t_1, \ldots, c_i) = \sum_j | d_{ij} - t_j - T_i(\mathbf{x}_j) - c_i |^p / (n_{ij})^p . \tag{7}$$

GMEL maximizes the likelihood with respect to the event parameters ($\mathbf{x}_j$, t_j) and station parameters (c_i and s_i), subject to a prior upper and lower bound on each parameter. The constraints on $\mathbf{x}_j$ comprise bounds on depth and a maximum epicentral distance from some specified geographic point.

We point out that with the generalized Gaussian error model, maximizing likelihood with respect to the event locations and station corrections, with the s_i fixed, is equivalent to minimizing the data misfit, Ψ, and in the Gaussian case (p = 2) this is a nonlinear least-squares problem.

The scheme GMEL currently uses for maximizing likelihood is an iteration of a two-step process. First, fixing the station corrections and standard errors to their initial values (the average of their upper and lower bounds), each event is located in turn. This is performed by applying the single-event grid-search algorithm (GSEL) separately to each event. The second step is a loop over stations to estimate the station parameters (c_i and s_i) with the event locations fixed. As in the event loop, the stations can be treated individually. In the Gaussian case with unrestrictive prior bounds, each station correction is simply the weighted mean residual at the station, but in general GMEL uses a root-finding procedure to find the correction that minimizes Ψ_i^{st}, the data misfit function for the station. The event loops and station loops are repeated until the likelihood function stops achieving significant gains. We can summarize the GMEL algorithm as an iteration of the following:

1. For each j: minimize Ψ_j^{ev} with respect to $\mathbf{x}_j$ and t_j (with the c_i and s_i fixed)
2. For each i: minimize Ψ_i^{st} with respect to c_i (with the $\mathbf{x}_j$ and t_j fixed)
3. For each i: maximize L with respect to s_I:

$$s_i = (\Psi_i^{\text{st}} / n_i)^{1/p}.$$

Theoretical Comparison of Methods

The multiple-event location methods we compare here are

- JHD: joint hypocenter determination, version JHD89 (Dewey, 1971).
- HDC: hypocentral decomposition (Engdahl and Bergman, 2000; Jordan and Sverdrup, 1981).
- PMEL: progressive multiple-event location (Pavlis and Booker, 1983).
- DD: double-differencing (Waldhauser and Ellsworth, 2000).
- GMEL: grid-search multiple-event location.

We briefly describe the other methods in relation to GMEL. First, we note that PMEL, HDC and JHD all perform a joint least-squares inversion for event location parameters and event-independent station corrections; i.e. they minimize Ψ, defined in equations (5) through (7), with $p = 2$. DD does the same in a special case, which we will explain below. Second, we note that all of the methods except GMEL employ some sort of iterated linearized inversion algorithm for minimizing Ψ with respect to location parameters, as opposed to grid-search in the case of GMEL. The essential differences between the algorithms have to do with the order in which the unknown parameters are solved for. This potentially makes a significant difference between algorithms because of the trade-offs among the unknown parameters in the multiple-event location problem. Jordan and Sverdrup (1981) showed that, under some simplifying approximations, the inverse problem (2) does not have a unique solution in that there is a perfect trade-off between a four-dimensional projection of the station corrections and the mean hypocenter and origin time of the events. However, the relative locations amongst events and the complementary projection of the corrections can be determined uniquely, disallowing the effects of the data errors. However, relaxing their assumptions and allowing for data errors as well as possible deficiencies in the completeness of the data, the trade-offs between station corrections and event locations might be more complex.

We explained that GMEL minimizes the data misfit function with respect to the location of each event and the correction at each station in alternating loops over events and stations. JHD solves for all the parameters of all events and stations simultaneously in a linear inverse problem in which the travel-time functions, $T_i(\mathbf{x}_j)$, have been linearized about the current event locations. (The dependence on the t_j and c_i is exactly linear.) The solution of the linearized problem yields updates to event and station parameters that are added to their current values. The process is iterated, thus solving the nonlinear least-squares problem.

PMEL resembles GMEL in iterating over a two-step process, with one step updating the event locations and the other updating the station corrections. The events are relocated individually as in GMEL, but PMEL uses a conventional single-event location algorithm instead of grid search. The updating of station corrections is quite different from GMEL. PMEL solves for the c_i simultaneously in a linear system which has been pre-conditioned with an “annulling” or “projection” operator (Pavlis and Booker, 1980) that prevents the station corrections from fitting components of the travel-time residuals that can be fit with hypocenter changes on subsequent iterations. Thus, PMEL resolves trade-offs between station and event parameters in favor of the events. Since GMEL updates event locations prior to station corrections on each pass of its iteration, it also gives some preference to fitting the data with event parameters, but not as faithfully as PMEL.

HDC solves separately for the centroid location of the event cluster and the relative locations of the individual events with respect to this centroid (“cluster vectors”). It too repeats a two-step process, but again the steps are different from GMEL. The first step solves simultaneously for all the cluster vectors (relative locations) in the linear inverse problem resulting from linearizing the travel-time functions about the current absolute event locations. HDC pre-conditions this linear inverse problem in the opposite sense of PMEL; i.e., it projects out the sensitivity to station corrections, which couples the problem across events (instead of stations as in PMEL) and which gives preference to the station corrections in fitting the travel-time residuals. (Under the Jordan-Sverdrup assumptions, this preference amounts to the constraint that the cluster vectors sum to zero.) In the second step of the HDC process, the updated cluster vectors are added to the current cluster centroid to get updated absolute locations, and then updated data residuals are computed. The station averages of these residuals are used to update the cluster centroid, and any remaining average residual at each station is its station correction. These two steps are repeated in an iteration.

Of the algorithms discussed thus far, DD follows the HDC paradigm most closely, but the step of finding relative locations is quite different. It solves for the relative locations in a linearized inverse problem obtained by forming arrival time differences, at each station, between pairs of close-by events recorded by the station. Wolfe (in press) has shown that in a simple case (uniform data weighting and all possible event pairs used in the inversion) this differencing procedure is equivalent to the HDC projection technique but with an implicit station-dependent re-weighting of the data (i.e., $s_i^2 = 1/n_i$). In addition to omitting time differences for certain event pairs, another distinguishing feature of DD is that it can weight each difference in accordance with the distance between the two events. As a result of these features (omitting and distance-weighting of time differences), DD relaxes the assumption that path corrections (c_{ij}) are event-independent and, accordingly, does not precisely minimize the data misfit function, Ψ, defined above in equations (5) through (7).

The various algorithms compared here also differ with regard to two other issues: how they re-weight data during the inversion process, and how they incorporate ground truth information. It is much more difficult to compare these aspects of the algorithms, and we defer a discussion of them until a later paper. In the tests shown below, we have attempted to equalize their re-weighting schemes, and we compare only the *relative* event locations each method obtained.

Comparison of Methods Applied to the Izmit/Duzce Clusters

Engdahl and Bergman (2001) applied HDC to two adjacent event clusters comprising 34 events from the 17 August 1999 Izmit earthquake sequence and 41 events from the 12 November 1999 Duzce sequence. They used Pn and P arrival times from the US Geological Survey National Earthquake Information Center (USGS/NEIC), and ground-truth (local network) locations for a few of the events to constrain the centroid of each cluster. The HDC analysis resulted in approximately 3500 defining phases at 650 stations for the Izmit cluster, and 3200 phases at 600 stations for Duzce. The defining phases covered an epicentral distance range of 3 to 100 degrees in each case.

We have independently applied the other methods to the same arrival time data and GT information. The rules under which these tests were conducted are as follows:

- Each method must be applied to the same set of defining phases used by HDC, using the same phase associations (P vs. Pn).
- Outlier rejection cannot be used.
- Event depths must be fixed to the same depths fixed in the HDC analysis.

Here we show a subset of the results we have obtained thus far and offer a limited interpretation of these results. We show only the Duzce clusters and, with one exception, results which did not incorporate local-network (GT) constraints on the locations. The exception is JHD, which was not run without the GT constraints, but we note that the relative locations are unlikely affected since only two of the Duzce events had GT information. We will focus on relative locations and, to facilitate this, we display each solution registered to the HDC result by shifting it to have the same cluster centroid.

Figure 1 compares two GMEL solutions to the HDC solution. In one case (top panel) the data were weighted equally and the weights were not changed in the inversion. In the second case, GMEL adjusted the data variances in a station-dependent manner by setting the standard error for each station (s_I) to its r.m.s. residual, as described in the GMEL algorithm above. (However, each s_i was bounded between 0.5 and 2.0.) It is immediately noticeable from Figure 1 that the GMEL relative locations obtained with this station-dependent weighting are much closer to the HDC locations. In the bottom panel (station-dependent weighting) the GMEL and HDC relative locations differ more than 2 km for only a few events and never more than 5 km, unlike the top panel. This is because HDC used a similar method of station-dependent weighting of the data and because, apparently, how the data are weighted is an important aspect of multiple-event location in this instance.

Figure 2 compares the DD, PMEL and JHD results to the HDC result for the Duzce cluster. In these other methods, the data weighting is not the same as in HDC. Fixed, uniform weighting was used in PMEL and DD. In DD, however, some possible time differences between events were omitted (when the events were separated by more than 30 km), although the retained pairs were not distance-weighted. The JHD algorithm has a complex weighting scheme, but in this example the weights depend mainly on event-station distance and the effect of this might be roughly comparable to station-dependent weighting.

We see from Figure 2 that the other methods also produce relative locations for the Duzce events much like HDC. However, since the data weighting differs from HDC, the agreement is not as close as the second GMEL case. The JHD results (Figure 2, top) match the HDC ones quite well. The DD epicenters are not quite as close and the PMEL result shows the greatest differences from HDC. A possible reason for this is that PMEL used uniform weighting, which we saw caused significant differences from HDC in the first GMEL case. A second possible reason is that PMEL attempts to fit as much of the data as possible with event location differences. Repeating from above, this would not affect the relative locations in an ideal, noise-free case but the PMEL result suggests that it might affect them in practice. Of particular relevance to this is the fact that this test uses a very large number of stations. For the

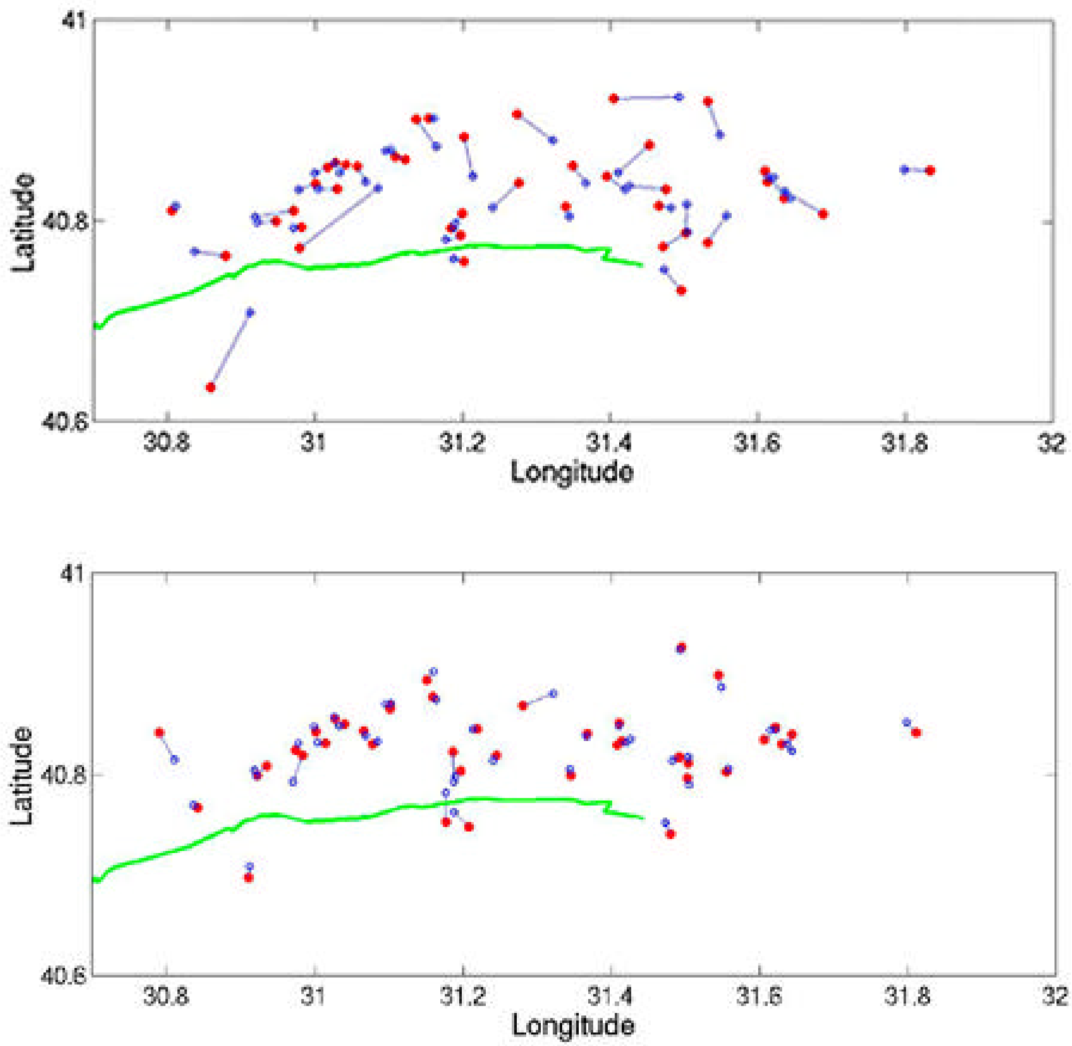

Figure 1: GMEL relative locations (filled circles) vs. HDC relative locations (unfilled) for the Duzce cluster. Two GMEL results are shown. *Top*: using fixed, uniform weighting of the data. *Bottom*: allowing a station-dependent weighting derived from the r.m.s. residual at each station. A line connects the HDC location of each event to the GMEL location of the same event. The surface trace of the North Anatolian Fault is shown for reference.

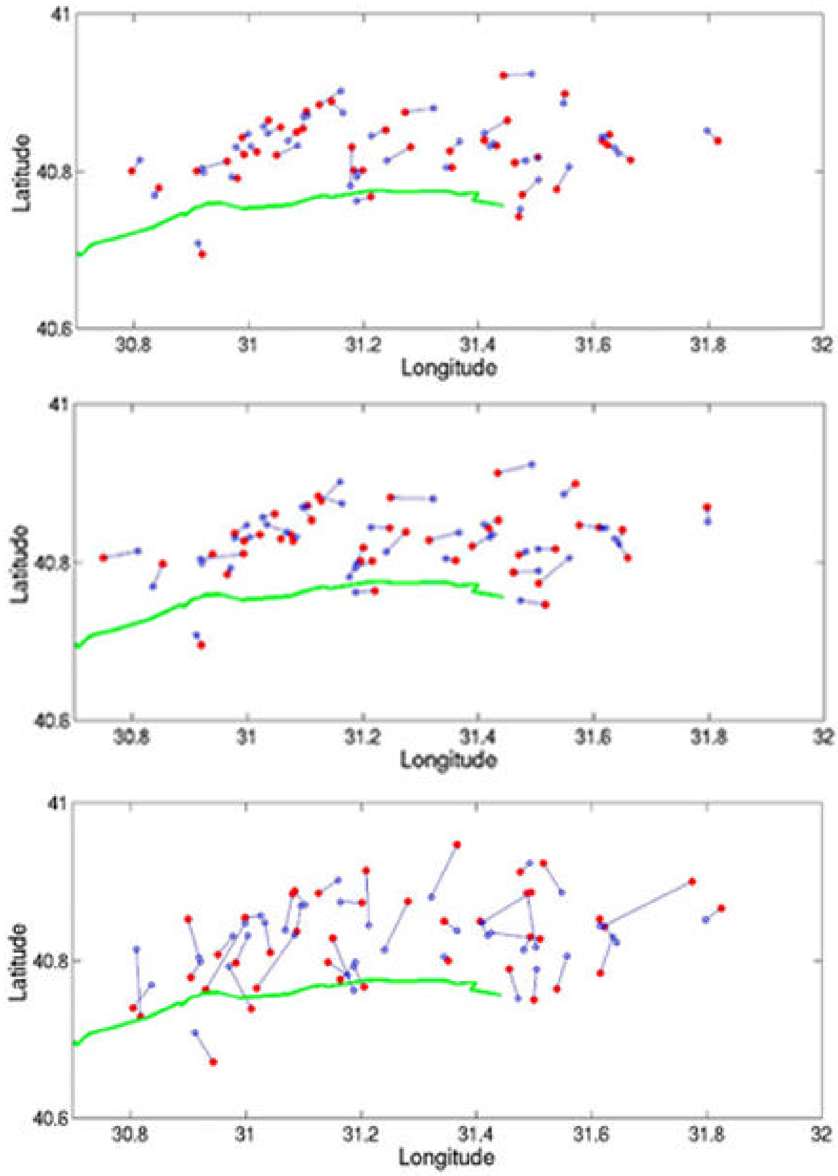

Figure 2: JHD, DD, and PMEL relative locations (filled circles) vs. HDC relative locations (unfilled) for the Duzce cluster. *Top*: JHD results, run with variable data weights. *Middle*: DD results, run with fixed, uniform data weights but with some differences omitted (see text). *Bottom*: PMEL results, run with fixed, uniform data weights.

Duzce cluster, the average number of arrivals per event is high (95) but the number of arrivals per station averages only 5.4. (The situation for Izmit is similar.) This provides the potential for trade-offs between the relative locations and station corrections.

This analysis and its conclusions must be considered preliminary since the comparisons made to date have not considered the errors (confidence regions) in the locations, which might explain some of the larger differences between methods. Since we have limited the comparisons to relative locations, we also have not considered the agreement of each method with available ground-truth locations, which is ultimately a more important test than the agreement between methods.

Confidence Regions and Modeling Errors

In Rodi and Toksöz (2000, 2001) we described a new approach to computing event location confidence regions for the single-event location problem. In the approach, the confidence region on an event hypocenter is defined as the locus of hypocenters which cannot be rejected based on a hypothesis test that compares the likelihood (L) achieved by any given hypocenter to the maximum likelihood achieved by the best hypocenter. The probability distribution of the test statistic, needed to associate a confidence level with the test, is inferred by Monte-Carlo simulation of the test statistic.

In the multiple-event problem, we seek to obtain a confidence region on an event hypocenter, say $\mathbf{x}_1$, allowing for the fact that the data depend also on the origin time of the event, the location parameters of the other $m-1$ events, and the travel-time corrections at the n stations. All these parameters are not uniquely determined by the data, and our uncertainty approach applied to the full multiple-event system will, in principle, automatically account for trade-offs between $\mathbf{x}_1$ and these other parameters. In doing so, the confidence region on $\mathbf{x}_1$ will implicitly account for modeling errors (in this case, uncertainty in the station corrections).

To do this, we must consider two values of the likelihood function, L, given in equation (4). The first, L_{max}, is the value achieved by the maximum-likelihood estimates of the parameters; i.e., L is maximized over all the station and event parameters. The second, which we define as a function of the hypocenter $\mathbf{x}_1$ and denote as $L_{\mathbf{x}_1}$, is the value achieved by maximizing L with respect to all the parameters except the hypocenter of event 1, which is fixed at $\mathbf{x}_1$. The test statistic for defining the confidence region on $\mathbf{x}_1$ is then

$$t(\mathbf{x}_1) = \log L_{\max} - \log L_{\mathbf{x}_1} \quad . \tag{8}$$

This statistic essentially compares how well the data can be fit with $\mathbf{x}_1$ fixed vs. free, with all the other parameters free to fit the data. The confidence region on $\mathbf{x}_1$ at confidence level b (e.g. $b = 90\%$) is then given by the locus of hypocenters $\mathbf{x}_1$ satisfying

$$F[t(\mathbf{x}_1)] = b \ . \tag{9}$$

$F[\ ]$ denotes the cumulative distribution function (c.d.f.) of a random variable.

The c.d.f. of t is determined by the assumed distribution of the data errors, e_{ij}. We are currently developing a Monte Carlo scheme for estimating the distribution of t. The scheme computes $\tau(\mathbf{x}_1)$ from many realizations of synthetic data generated using some assumed true parameters (event locations and station corrections) and varying samples of pseudo-random errors. Applying this approach in the multiple-event case is very computationally intensive since it involves the computation of many multiple-event inversions, one for each point $\mathbf{x}_1$ on a hypocenter grid, using the real data, and two for each realization of synthetic data. We have implemented a basic algorithm and are using it to perform a proof-of-concept of the approach while investigating computational shortcuts and approximations to make it more practical.

CONCLUSIONS AND RECOMMENDATIONS

We have developed a general theoretical and computational framework for characterizing the uncertainty in seismic event locations. At present, we are extending this framework to the multiple-event location problem in an attempt to determine event location confidence regions that rigorously account for the effects of errors in the travel-time forward model (modeling errors) in addition to observational (picking) errors. We extended our single-event grid-search location algorithm (GSEL) to a basic multiple-event algorithm (GMEL) for this purpose, and are now developing a Monte-Carlo technique for computing non-elliptical location confidence regions in the multiple-event problem.

Our comparison of GMEL to other multiple-event location methods (HDC, DD, JHD and PMEL) has helped to understand the similarities and differences among these methods, in theory and in practice. This is an important step in the development of GMEL and its new approach to location uncertainty. The initial comparison tests indicate that the different methods produce quite similar results when applied with the same data and assumptions. However, different manners of weighting the data and resolving parameter trade-offs appear to have a significant effect on the event locations determined by multiple-event location. We recommend more experiments of the type performed here, on other event clusters and data sets, to better understand these effects.

REFERENCES

Dewey, J.W. (1971), Seismicity studies with the method of joint hypocenter determination, Ph.D. Thesis, University of California, Berkeley.

Douglas, A. (1967), Joint epicentre determination, *Nature* **215**, 47-48.

Engdahl, E.R. and E.A. Bergman (2000), Identification and validation of reference events within the area being regionally monitored by IMS stations in Asia and North Africa, *Proceedings*, 22nd Annual DoD/DOE Seismic Research Symposium, New Orleans.

Engdahl, E.R. and E.A. Bergman (2001), Validation and generation of reference events by cluster analysis, *Proceedings*, 23rd Annual DoD/DOE Seismic Research Review, Jackson Hole, Wyoming.

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.

Lilwall, R.C. and A. Douglas (1970), Estimation of P-wave travel-times using the joint epicentre method, *Geophys. J. R. Astr. Soc.* **19**, 165-181.

Pavlis, G.L and J.R. Booker (1980), The mixed discrete continuous inverse problem: application to the simultaneous determination of earthquake hypocenters and velocity structure, *J. Geophys. Res.* **85**, 4801-4810.

Pavlis, G.L, and J.R. Booker (1983), Progressive multiple event location (PMEL), *Bull. Seism. Soc. Am.* **73**, 1753-1777.

Rodi, W. and M.N. Toksöz (2000), Grid-search techniques for seismic event location, *Proceedings*, 22nd Annual DoD/DOE Seismic Research Symposium, New Orleans.

Rodi, W. and M.N. Toksöz (2001), Uncertainty analysis in seismic event location, *Proceedings*, 23rd Annual DoD/DOE Seismic Research Review, Jackson Hole, Wyoming.

Waldhauser, F. and W.L. Ellsworth (2000), A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California, *Bull. Seism. Soc. Am.* **90**, 1353-1368.

Wolfe, C.J. (2002), On the mathematics of using difference operators to relocate earthquakes, *Bull. Seism. Soc. Am.*, in press.