LOCATION CALIBRATION IN WESTERN EURASIA AND NORTH AFRICA: GROUND TRUTH, IMPROVED EARTH MODELS, BAYESIAN KRIGING, REGIONAL ANALYSIS, LOCATION ALGORITHMS, ARRAY CALIBRATION, AND VALIDATION

Stephen C. Myers, Megan Flanagan, Michael Pasyanos, William Walter, Paul Vincent, and Craig Schultz

Lawrence Livermore National Laboratory

Sponsored by National Nuclear Security Administration
Office of Nonproliferation Research and Engineering
Office of Defense Nuclear Nonproliferation

Contract No. W-7405-ENG-48

ABSTRACT

Improved seismic location is traditionally accomplished by better predicting arrival times for seismic phases. The National Nuclear Security Administration (NNSA) Ground-Based Nuclear-Explosion-Monitoring Research and Engineering (GNEM R&E) calibration effort at Lawrence Livermore National Laboratory (LLNL) begins with an extensive ground truth integration effort. Accurate locations, arrival picks, and waveforms are continuously added to the LLNL database, where they are cross-referenced with existing holdings and characterized for accuracy. The ground truth (GT) database is used to assess travel-time prediction accuracy for both Earth models that are developed here and models brought in through the integration process. When possible, we use data-driven techniques, such as tomography, to improve Earth models, and in aseismic regions, we build models through geophysical analogy. Regardless of the base model, we use empirical corrections (Bayesian Kriging) where GT events are available to ensure the most accurate travel-time prediction. Component validation is used to assess the importance of each calibration activity, and end-to-end validation of seismic event locations allows us to assess overall progress.

New techniques and approaches are aiding the calibration effort. 1) Earthquake location accuracy is improved by combining seismic and InSAR signals. The addition of InSAR data sets may allow us to reduce the uncertainty of uncalibrated global locations from 15 km to 5 km. 2) Region-specific wave field characterization extracts additional information from each seismic trace, bringing more information to bear on location. 3) We are developing new methods to assess travel-time prediction uncertainty from models, particularly in aseismic areas. Using Bayesian methodologies we assess uncertainty using limited ground-truth data sets. 4) Array calibration improves azimuth and slowness prediction and will help to better identify seismic phases and improve sparse-network location. 5) Lastly, in collaborative efforts, we are working to implement multiple-event grid search locations to improve both event precision and uncertainty propagation.

OBJECTIVE

Improvement in seismic location is accomplished by combining model-based and empirical travel-time and azimuth-slowness corrections. Model-based approaches are aimed at improving travel-time prediction at regional distance (between approximately 1.5° and 13°) and at upper-mantle triplication distances (between approximately 13° and 25°). Beyond upper mantle distance (between 25° and 90°) the IASPEI91 global earth model with a static station correction provides excellent travel-time prediction. Therefore, only the highest quality reference events are useful at teleseismic distances. After a calibration model is applied, empirical corrections are calculated using the Modified Bayesian Kriging algorithm (Schultz *et al*., 1998) with travel-time residuals for suitably well located calibration events. Our approach combines the extrapolative advantages of model-based corrections and the interpolative/geostatistical advantages of Kriging to produce hybrid travel-time predictions and uncertainty models. For ease of use, model-based and empirical corrections are combined to produce one travel-time correction and uncertainty model that is applied to the IASPEI91 global earth model. Validation of improvement and proper error characterization are performed at each step of the calibration process. Therefore, the importance of each calibration component is assessed and representative uncertainty estimates are assured. Final calibrations are tested in an end-to-end relocation of test data sets, providing a final validation of uncertainty estimates and improvement in location accuracy.

Overviews of the calibration process we employ can be found in previous Seismic Research Review proceedings. This year we provide more detailed discussion on selected aspects of the calibration process. The topics covered are 1) collection and characterization of calibration events, 2) assessment of model-based uncertainty, and 3) accounting for calibration events with a range in location accuracy.

RESEARCH ACCOMPLISHED

Calibration-Events

The accuracy of earth model and empirical travel-time predictions is limited to the accuracy of calibration events that are used to develop them. Therefore, collection and validation of a geographically distributed set of calibration events is critical to the success of any travel-time calibration effort. The GNEM R&E program integrates calibration events from a diverse set of sources, including dedicated scientific explosions; nuclear explosions; events with InSAR locations; global, regional, and local seismicity catalogs; mine explosions and collapses; and military activity. The database of calibration events includes not only numerous source types but also a range in location accuracy. Although events with perfectly known locations are most preferable, improvement in travel-time prediction is often achieved using calibration events with non-zero location uncertainty. Further, the goal of covering a wide geographic area requires inclusion of events that are not perfectly located.

Categorization of location accuracy is accomplished using the "GTX" system of Bondár (1998), where X is location accuracy. We modify this nomenclature by defining X as the 95% confidence that an epicenter is within X kilometers of the true location.

The vast majority of calibration event locations is determined seismically. For a minority of events, satellite information (photography or InSAR) can add independent constraints on location, and for a select few events (dedicated explosions) the Global Positioning System (GPS) can provide working level GT0. For events that are determined seismically, however, formal error ellipses are not a reliable metric of location accuracy. Analysis of seismic location accuracy is traditionally based on formal uncertainty. Most location algorithms rely on one of two methods to determine uncertainty. The first is based on the F-statistic, where the *a posteriori* residual distribution is mapped to a location confidence ellipsoid (Flinn, 1965). The second is based on the chi-square statistic, where *a priori* uncertainty for phase picking and travel-time prediction are mapped though the location algorithm to produce a coverage ellipsoid (Evernden, 1969). Proper application of either technique requires compliance with basic statistical assumptions. Both methods require Gaussian, zero mean, uncorrelated error processes. A number of studies suggest that these assumptions are violated in most seismic locations. Picking error tends to have "heavy" tails (Buland, 1986) and may be multimodal (Ryall and Schultz, 2002). The mean path-specific, travel-time prediction error is typically not zero, and travel-time prediction errors are typically correlated for similar ray paths (e.g. Myers and Schultz, 2000a).

Because formal error ellipses cannot be used to assess location accuracy, we are forced to use other information to assess location accuracy for potential calibration events.

Network configuration is a good indicator of seismic location accuracy. Events that are located within the footprint of a seismic network tend to be accurately located because sensitivity to travel-time prediction errors are minimized (see Myers and Schultz, 2000a for discussion). It is, therefore, useful to develop metrics for network coverage that are 1) easily derived from seismic catalogs and 2) empirically correlated to location accuracy. In particular, the maximum azimuthal gap (Sweeney, 1998) in station coverage and the secondary gap (maximum gap when any given station is removed from the network) (Bondár *et al*., 2002) are easy to calculate and are found to be indicative of location accuracy. Here we focus on the results of a collaborative effort among LLNL, the Center for Monitoring Research, and the University of Colorado that are reported in Bondár et al. (2002). Local network locations with at least 10 stations within 250 km, an azimuthal gap less than 110° and a secondary azimuthal gap less than 160° and at least one station within 30 km from the epicenter produce GT5 epicenters (Figure 1) are used. Regional networks between 2.5° and 10° with secondary azimuthal gap less than 120° produce GT20 epicenters, but only at the 90% confidence level. Lastly, teleseismic networks with secondary azimuthal gap less than 120° produce GT25, again at the 90% confidence level. These criteria are conservative and may be improved by using enhanced phase identification or simple travel-time corrections; this is the case for Engdahl *et al*. (1998). Myers and Schultz (2000a) find that when these additional measures are taken, teleseismic locations improve to GT15 (Figure 2). We note, however, that subduction zones are not sampled in the GT15 test data set, and locations in these regions are probably not GT15.

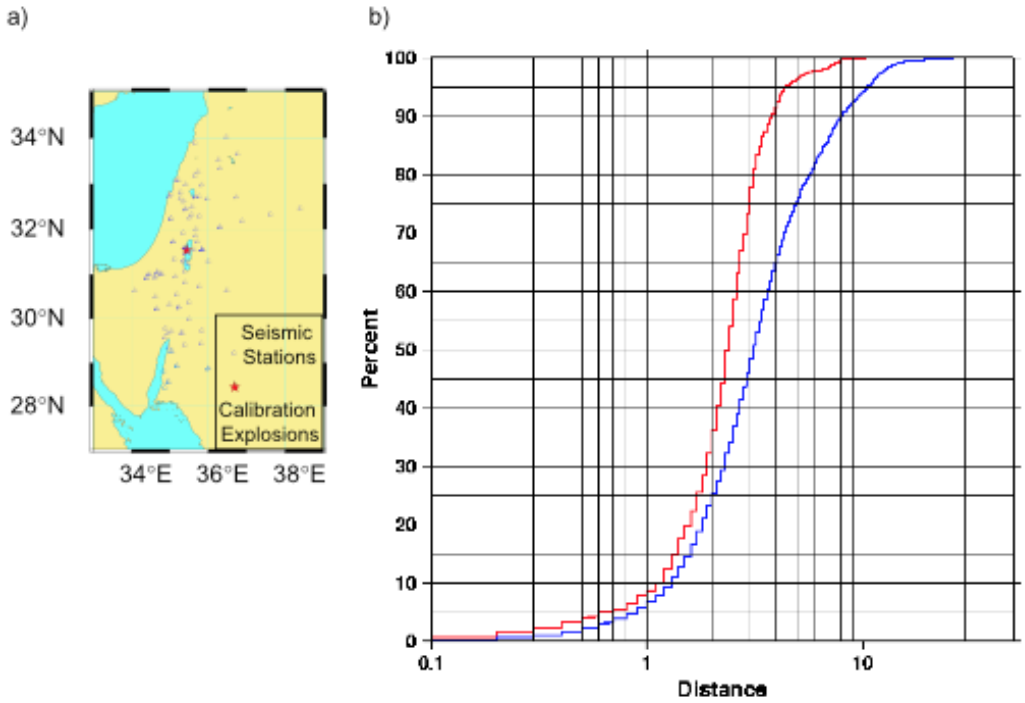

Figure 1. The 1999 Dead Sea calibration explosions and a Swiss munitions explosion are used to validate GT5 network criteria. The explosions are repeatedly located using 10 randomly selected stations from the dense local network. A) Dead Sea network configuration. B) Cumulative distribution of event mislocations: red is the curve for networks meeting the GT5 criteria, and blue is the curve for all network configurations (From of Bondar *et al*., 2002).

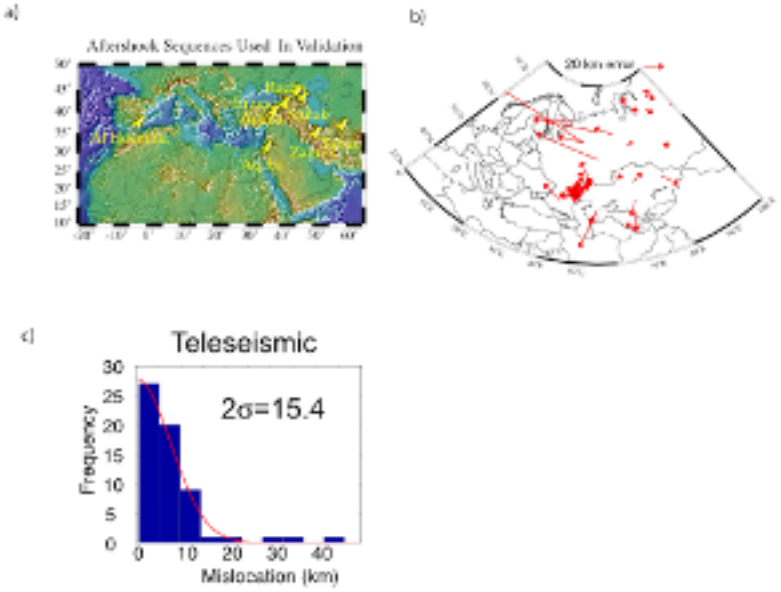

Figure 2. Validation GT15 for teleseismic locations of Engdahl *et al*. (1998) using local-network locations and peaceful nuclear explosions (PNEs). These locations are improved over routine global network locations by reassociating phases and use of regional travel-time corrections. A) Aftershock sequence locations used to validate teleseismic GT15. B) GT1 locations of Sultanov *et al*. (1999) with mislocation vectors of teleseismic locations. C) Distribution of teleseismic mislocation confirms that 95% of teleseismic locations are within 15 km of local-network or GT1 locations.

Location accuracy can be improved by using data complementary to seismic observations. InSAR data are particularly useful for constraining earthquake and mine activity that results in static surface deformation of more then a few centimeters. We have a cross-disciplinary effort to use InSAR for seismic calibration purposes. This effort includes expertise in processing interferometric data, allowing us the flexibility to combine satellite and seismic data into a joint location estimate. Figure 3 shows seismic and InSAR locations for the mb=5.3, April 10, 1998, earthquake in eastern Iran. In this example seismic locations are systematically offset to the north by approximately 20 km, compared to the InSAR location. Each of the seismic locations relies heavily on stations at teleseismic distance and, in this case, the locations meet the GT25 criteria outlined above (locations do not include additional processing). Based on modeling of the InSAR-constrained static displacement field, the source dimension of this earthquake is probably on the order of 5 km. Because we do not have sufficient resolution to attribute teleseismic first arrivals to moment release on a sub-portion of the fault, the InSAR epicenter is estimated to be GT5.

Cross-referencing all sources of calibration-event information is an important check on our database. In many instances several location estimates may be available for a single event. These sources of information may or may not make use of similar data sets, so estimated GT level for each catalog/study may be different. As part of the validation process, we cross-reference epicenters for each event to test the consistency predicted by the assigned GT level. This procedure is used not only to validate the GT level of events, but we have frequently found outlier data that may otherwise go unnoticed (Figure 4). In this example, satellite locations of Novaya Zemlya nuclear explosions (Skorve and Skogan, 1992) are compared to seismic locations determined relative to one satellite location (Marshall, 1994). Location differences for all but two of the events are consistent with GT1 for satellite locations and GT5 for teleseismic relative locations. We find, however, two events with anomalously large location differences. Although we are not completely certain about the origin of the differences, we note that subtraction of one integer degree from the satellite locations makes the location differences comparable to location differences for other events. It is important to note that the purpose of cross-referencing is not to identify preferred calibration-

event contributors, because each source of calibration information may have some outliers. The purpose is to check our GT assignments and flag suspect locations.

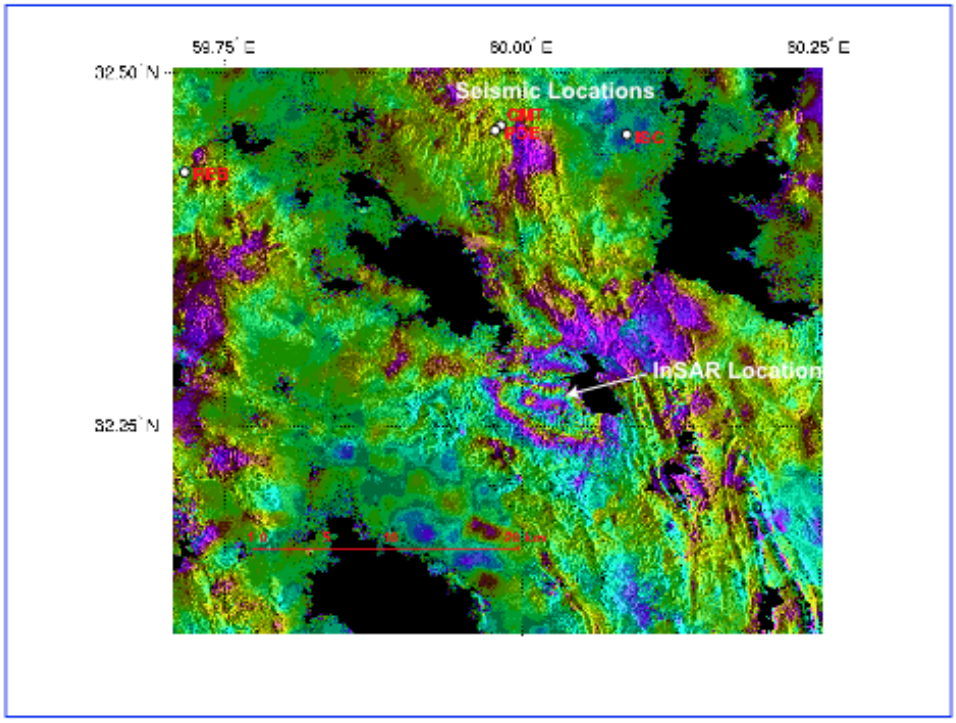

Figure 3. InSAR location of the mb=5.3, April 10, 1998, earthquake in eastern Iran. Global seismic estimates (red text) are biased by approximately 20 km to the north. (InSAR processing by Vincent, 2002)

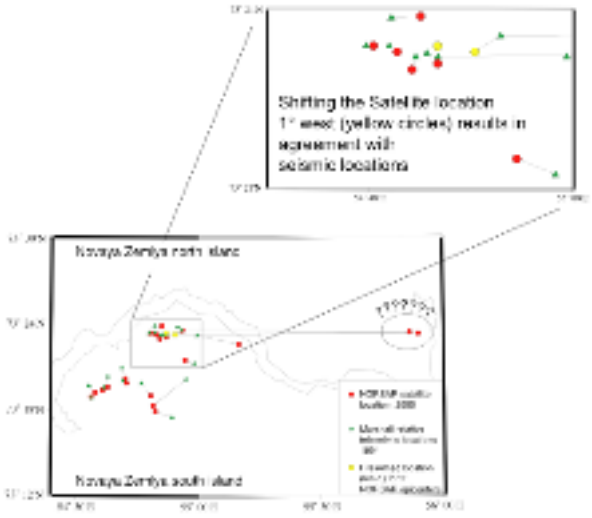

Figure 4. Cross-referencing can flag outlier data that may otherwise be used in calibration. In this example satellite locations of Skorve and Skogan (1992) are generally in good agreement with relative seismic location of Marshal, 1994. However, two satellite locations are outliers. Subtracting one integer degree of longitude brings the locations into agreement. These events are discarded in our calibration procedure.

Model Uncertainty Estimation

Geophysical models improve location performance in areas of limited empirical calibration. Similar to calibration-event collection, efforts begin with a survey of existing models. Each model is evaluated first and foremost using a travel-time prediction metric, which is more directly related to location improvement than metrics involving model velocities or discontinuities. Travel-time prediction accuracy is assessed in a non-stationary framework. Figure 5 shows an example of stationary and non-stationary uncertainty models. Although stationary (scalar) uncertainty is most desirable, we find that uncertainty for most models is dependent on event-station distance. For simple models, even distance-dependant uncertainty models do not capture the complexity of travel-time uncertainty. In these cases uncertainty must be assessed by geographic region, as show in Figure 5c.

Travel-time prediction uncertainty is most directly assessed by comparing model predictions with observations from calibration events. However, proper application of empirically determined error distributions requires representative geographic sampling. In areas with sparse data coverage, especially when non-stationary status requires higher-dimensional error models, representative data sampling is not achieved. In these cases we use a Bayesian approach that takes advantage of prior travel-time distribution information. With the Bayesian approach, we begin with a conservative, global estimate of travel-time uncertainty and reduce the distribution as data coverage improves.

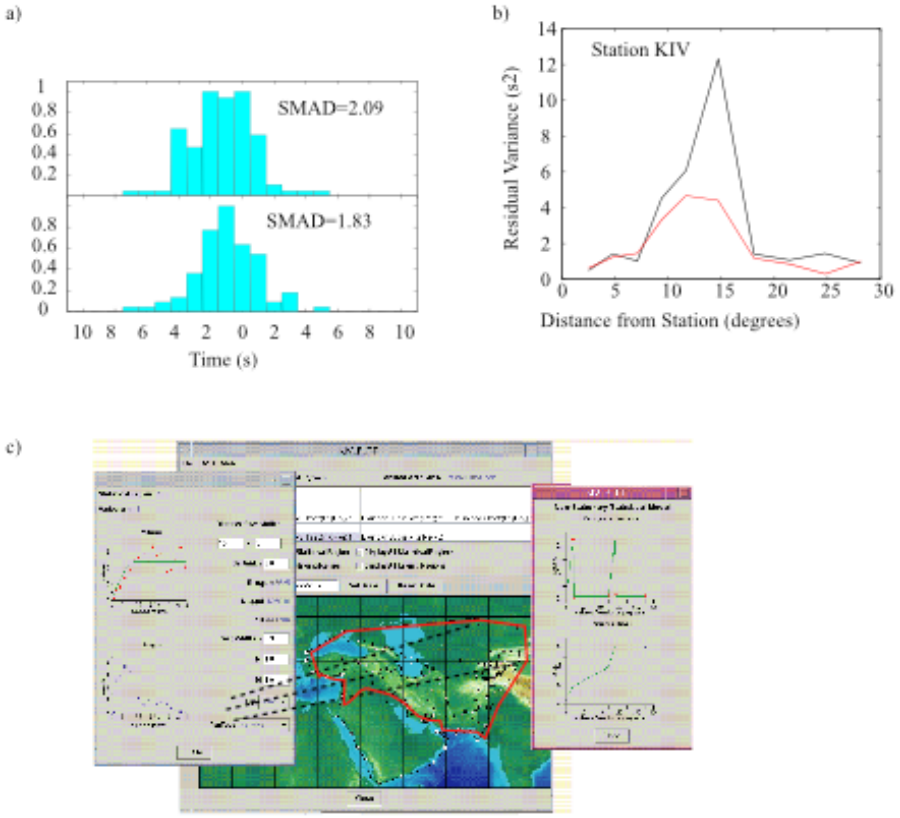

Figure 5. Example of a stationary and non-stationary travel-time residual characterization. a) All travel-time residuals are used to build one distribution. Improved prediction is suggested in the lower panel. b) Distance-dependent analysis shows that, in this instance, the improvement in prediction comes from events between about 7° and 18°. c) For simple models uncertainty must be assessed in distinct geographic regions. Simple distance dependence is not sufficient.

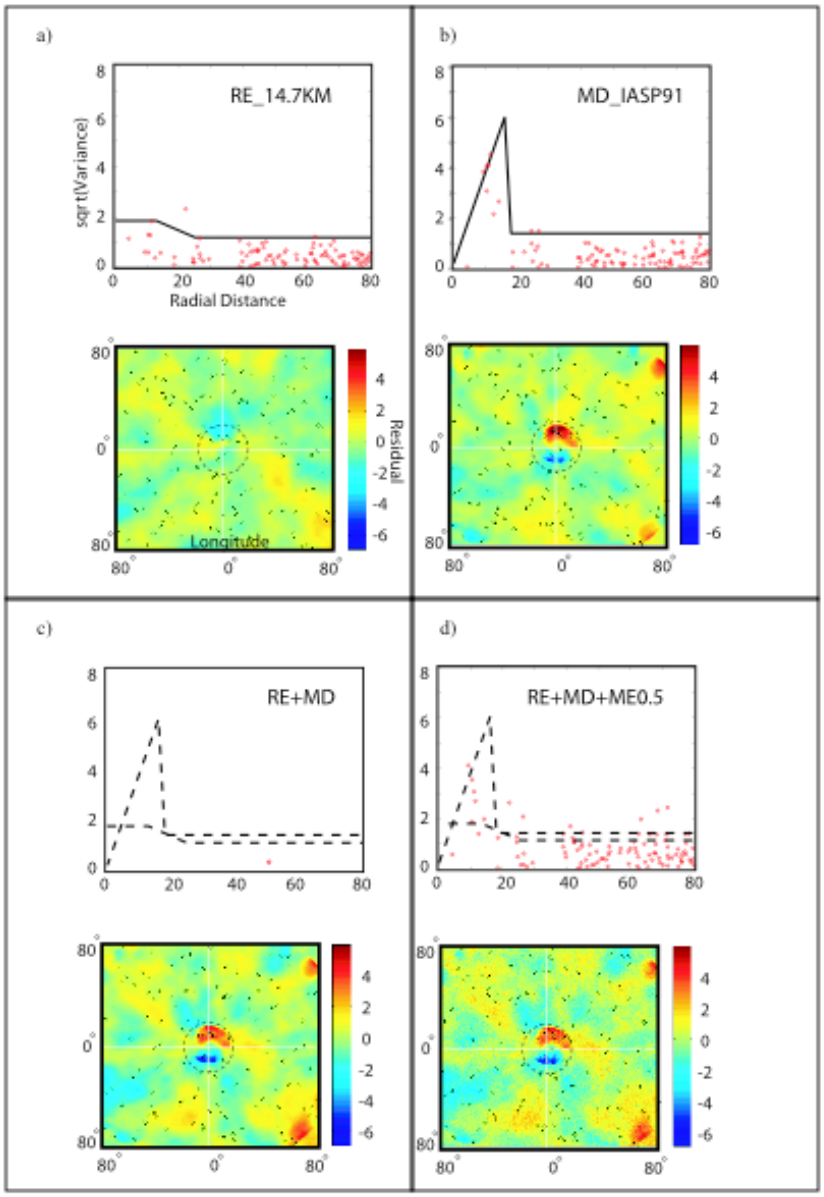

Figure 6. Synthetic generation of non-stationary travel-time residuals are used to test new techniques for assessing model uncertainty. a) Non-stationary travel-time residuals caused by calibration-event error. b) Non-stationary travel-time residuals caused by model error. c) Stationary residuals cause by picking error. d) The result of adding error process from a) b) and c).

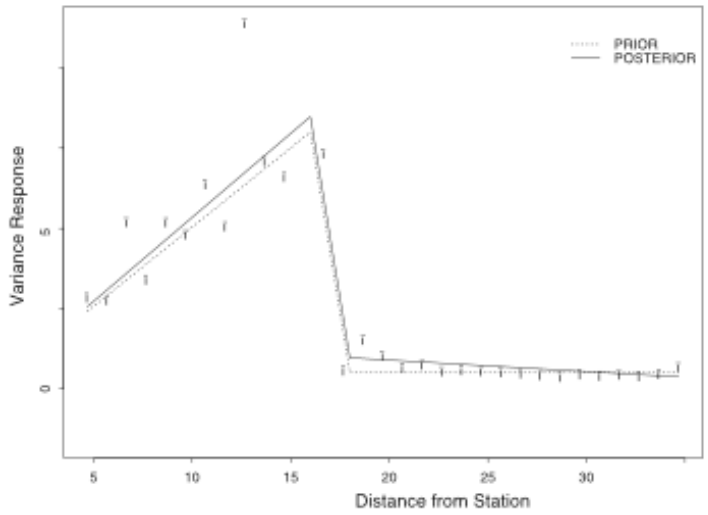

Figure 7. Example of Bayesian regression to determine model error. The *a priori* uncertainty (dotted line) is modified by observations to produce *a posteriori* estimate of uncertainty (solid line).

The Bayesian model estimator is currently tested using synthetic data sets. The synthetic data sets include additional error processes that complicate estimation of the desired model error. Inclusion of additional error processes make the synthetic tests as realistic as possible. Figure 6 is an example of a synthetic travel-time residual surface constructed using model, calibration event, and picking error. Panel a) shows non-stationary travel-time errors attributed to calibration-event location error (discussed below). Panel b) is the non-stationary model error, which is the error process we hope to extract from the data, and c) shows a stationary measurement error. All three of these error processes are combined to produce a synthetic error distribution from which example instances are constructed. An example of the Bayesian regression is shown in Figure 7. The prior estimate is the dotted line and the posterior estimate is the solid line. Although one outlier at a distance of approximately 13° suggests substantially increased uncertainty, the Bayesian prior tempers the posterior estimate.

Calibration-Event Error Propagation

Combining events with variable location accuracy is accomplished in a rigorous statistical framework. Travel-time measurement derived from more accurately located events should receive more weight than corrections from events with poorer location accuracy. Likewise, the uncertainty of travel-time corrections derived from more accurately located events should be reduced. Mapping location uncertainty to travel-time uncertainty is the first step in assessing the appropriate weights. This relationship is established by Myers (2001).

$$\sigma_{tt}^2 = \left(\frac{\partial t}{\partial \Delta} \sigma_{dist} \right)^2 \qquad [1]$$

where σ^2_{tt} is travel-time variance, $dt/d\Delta$ is the derivative of the travel-time curve, and σ^-_{dist} is the variance of the epicenter accuracy. Figure 8 shows the mapping of several GT levels into travel-time error as a function of distance. Therefore, the stationary calibration-event uncertainty maps to a non-stationary travel-time uncertainty, and proper propagation of errors requires application of Equation [1].

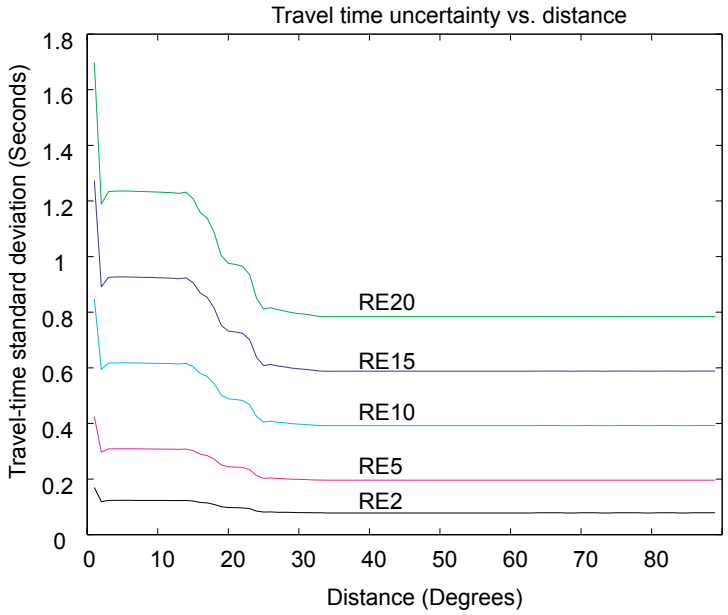

Figure 8. Travel-time uncertainty resulting from epicenter uncertainty as a function of distance for IASPEI91 P-waves is shown. The epicenter uncertainty is a worst-case scenario, where the error occurs along the great-circle path. Each curve is labeled by reference event (RE) level, where the number is the 95% confidence in the location accuracy (km). Note that overshoot at changes in gradient of the curves is the result of errant spline interpolation and should be ignored (From Myers, 2001)

Accounting for calibration-event bias

In some instances reference-event error contains a biased component. The GNEM R&E processing accounts for this through a declustering algorithm (Myers, 2001). Declustering combines events that are close to one another – events that may have similar vector mislocation – into one observation. Within each decluster bin, locations derived from the same source (e.g. teleseismic or local-network locations) are grouped and averaged separately. Then each of the groups are combined using a weighted average. Uncertainty for the group average is given by:

$$\tau_{re_i}^2 = \frac{1}{N^2} \sum_{j=1}^{N} \sum_{k=1}^{N} C_{jk} \qquad [2]$$

where τ^2 is the variance of the mean, N is the number of events, C_{jk} is the covariance, j and k count the events in the group.

After averaging the events belonging to individual groups, we combine the groups into a single decluster value via a weighted sum.

$$t_{decluster} = \sum_{i=1}^{M} w_i \bar{t}_{res_i} \qquad [3]$$

where $t_{decluster}$ is the declustered value, w_i is the weight of the i_{th} group, and t_{res} is the group average. The weights are a normalized inverse of the decluster uncertainty.

$$w_i = \frac{N_o}{\tau_i} \qquad [4]$$

where N_o is a normalizing factor.

The variance of the declustered point is then:

$$\tau_{decluster}^2 = \sum_{i=1}^{M} (w_i \tau_i)^2 \qquad [5]$$

CONCLUSIONS AND RECOMMENDATIONS

The LLNL location program encompasses each aspect of seismic calibration. Only a few of the components are highlighted here. In this paper we focus on collection, database building, validation, and cross-referencing of calibration events, as well as model error assessment and rigorous propagation calibration event location uncertainty to travel-time uncertainty.

In summary:
Collection of a diverse set of calibration events is critical to broad area calibration. LLNL combines seismic and satellite methods to obtain the most accurate source locations for calibration events. Cross-referencing calibration-event locations validates GT assignments and flags outliers. Models are judged on travel-time prediction. Assessment of model error using empirical observations is complicated by calibration-event and phase picking uncertainty. The LLNL Bayesian regression tool addresses GT and picking uncertainties and provides a quantitative means to account for data coverage during model uncertainty estimation. Stationary calibration event epicenter errors map to non-stationary travel-time uncertainties. Differing levels of GT are combined through a weighting system that is related to epicenter accuracy.

ACKNOWLEDGEMENTS

We thank Dave Harris, Bill Rodi, Istvan Bondar, Bob Engdahl, Eric Bergman, and Keith Koper for discussions and collaborations.

REFERENCES

Bondár, I., S.C. Myers, E.R. Engdahl, and E.A. Bergman (2002), Epicenter accuracy based on seismic network criteria, in preparation.

Buland, R. (1986), Uniform reduction error analysis, *Bull. Seismol. Soc. Am*., **76**, 217-230.

Engdahl, E.R., R. van der Hilst and R. Buland (1998), Global Teleseismic Earthquake Relocation with improved Travel Times and Procedures for Depth Determination, *Bull. Seism. Soc. Am.*, **88**, 722-743.

Evernden, J. (1969), Precision of epicenters obtained by small numbers of world-wide stations, *Bull. Seism. Soc. Am*., **59**, 1365-1398.

Flinn, E. (1965), Confidence regions and error determinations for seismic event location, *Rev. Geophys.*, **3**, 157-185.

Myers, S.C. and C.A. Schultz (2000a), Calibration of seismic travel time using events with seismically determined locations and origin times, *EOS Trans. AGU*, **81**, F845.

Myers, S. and C. Schultz (2000), Improving Sparse-Network Seismic Location with Bayesian Kriging and Teleseismically Constrained Calibration Events, *Bull. Seism. Soc. Am*., **90**, 199-211.

Myers, S.C., Methods of travel-time residual declustering for the Knowledge Base Calibration and Integration Tool (KBCIT), UCRL-ID-142521, 2001.

Ryall, F. and C.A. Schultz (2001), Seismic waveform characterization at LLNL: analyst procedures and issues, Lawrence Livermore National Laboratory, UCRL-ID-146229.

Schultz, C., S. Myers, J. Hipp, and C. Young (1998), Nonstationary Bayesian Kriging: Application of Spatial Corrections to Improve Seismic Detection, Location and Identification, *Bull. Seism. Soc. Am*., 1275-1288.

Skorve, J. and J. Skogan, The NUPI satellite study of the northern underground nuclear test area on Novaya Zemlya, NUPI Report No. 164, Oslo, Norway, 1992.

Sultanov, D.D., J.R. Murphy and Kh. D. Rubinstein, A seismic source summary for Soviet peaceful nuclear explosions, *Bull. Seism. Soc. Am*., **89**, 640-647, 1999.

Sweeney, J. (1998), Criteria for Selecting Accurate Event Locations from NEIC and ISC Bulletins, Lawrence Livermore National Laboratory report UCRL-JC-130655.