

**RELATIVE LOCATION ACCURACY AND WAVEFORM SUBSPACE DETECTOR**

Stephen C. Myers,<sup>1</sup> David B. Harris,<sup>1</sup> Megan L. Anderson,<sup>1</sup> William Rodi,<sup>2</sup>  
William R. Walter,<sup>1</sup> Megan P. Flanagan,<sup>1</sup> and Flori Ryall<sup>1</sup>

Lawrence Livermore National Laboratory<sup>1</sup>, Massachusetts Institute of Technology<sup>2</sup>

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

Contract No. W-7405-ENG-48

**ABSTRACT**

We present two Lawrence Livermore National Laboratory (LLNL) research projects in the topical areas of location and detection. The first project assesses epicenter accuracy using a multiple-event location algorithm, and the second project employs waveform subspace correlation to detect and identify events at Fennoscandian mines.

Accurately located seismic events are the bases of location calibration. A well-characterized set of calibration events enables new Earth model development, empirical calibration, and validation of models. In a recent study, Bondar et al. (2003) develop network coverage criteria for assessing the accuracy of event locations that are determined using single-event, linearized inversion methods. These criteria are conservative and are meant for application to large bulletins where emphasis is on catalog completeness and any given event location may be improved through detailed analysis or application of advanced algorithms. Relative event location techniques are touted as advancements that may improve absolute location accuracy by 1) ensuring an internally consistent dataset, 2) constraining a subset of events to known locations, and 3) taking advantage of station and event correlation structure. Here we present the preliminary phase of this work in which we use Nevada Test Site (NTS) nuclear explosions, with known locations, to test the effect of travel-time model accuracy on relative location accuracy. Like previous studies, we find that the reference velocity-model and relative-location accuracy are highly correlated. We also find that metrics based on travel-time residual of relocated events are not a reliable for assessing either velocity-model or relative-location accuracy.

In the topical area of detection, we develop specialized correlation (subspace) detectors for the principal mines surrounding the ARCES station located in the European Arctic. Our objective is to provide efficient screens for explosions occurring in the mines of the Kola Peninsula (Kovdor, Zapolyarny, Olenogorsk, Khibiny) and the major iron mines of northern Sweden (Malmberget, Kiruna). In excess of 90% of the events detected by the ARCES station are mining explosions, and a significant fraction are from these northern mining groups. The primary challenge in developing waveform correlation detectors is the degree of variation in the source time histories of the shots, which can result in poor correlation among events even in close proximity. Our approach to solving this problem is to use lagged subspace correlation detectors, which offer some prospect of compensating for variation and uncertainty in source time functions.

## **OBJECTIVE**

One of the most important and challenging aspects in seismic calibration is assessment of location accuracy for candidate reference events. Location accuracy criteria based on network coverage has been developed for routine bulletins (e.g. Bondar et al., 2003), where single-event algorithms make use of contributed phase picks. These objective criteria allow efficient assessment of large volumes of bulletin data. However, there is an increasing trend towards the application of multiple-event algorithms combined with careful analyst review to develop reference-event bulletins (Armbruster et al., 2002; Engdahl et al., 2002). The currently accepted criteria (Bondar et al., 2003) are not applicable to – and probably over estimate – location errors determined by using multiple-event algorithms. One of the primary benefits of multiple-event locations is the determination of station corrections based on the best-located events; in effect, diminishing the importance of reference velocity model accuracy. Here we present preliminary results for seismic location accuracy studies using the Nevada Test Site (NTS) nuclear explosions with known locations. To test the importance of the velocity model accuracy, 74 events are relocated using 4 distinct travel-time models.

In the second part of this study we examine the use of waveform subspace detectors in Fennoscandia. The National Nuclear Security Administration (NNSA) Ground-Based Nuclear Explosion Monitoring Research and Engineering (GNEM R&E) Program at Lawrence Livermore National Laboratory (LLNL) is conducting research on subspace (i.e. generalized correlation) detectors for use in mine-event screening. Regional mining explosions dominate detections at some seismic monitoring stations. This is particularly true of the ARCES array, where the large mining districts of the Khibiny Massif, Zapolyarny, Olenogorsk, Kovdor and northern Sweden (iron mines Malmberget and Kiruna) constitute up to 90 percent of detections. Here, we present initial efforts to develop screening detectors for application to the mines of the Khibiny Massif.

## **RESEARCH ACCOMPLISHED**

### **Location Accuracy**

#### **Dataset**

We make use of the NTS dataset (Walter et al., this Proceedings), with known event locations, to study the accuracy of relative seismic locations. The 74 explosions and 61 regional stations used in this study are shown in Figure 1. An individual LLNL analyst re-picked these events to produce a high-quality dataset. The 1577 LLNL phase arrivals, with best-fit regressions, are shown in Figure 2. National Earthquake Information Center (NEIC) phase picks augment the LLNL dataset in some of our relocations, but only picks whose residuals are within three standard deviations of the empirical curves are kept. The number of picks outside the three standard deviation bound far exceeds statistical expectations (including the LLNL picks). In large part, transient clock errors are thought to account for the “heavy tails” in the residual distribution in this dataset (Walter et al., this Proceedings).

#### **Travel-time models**

Events are re-located using four travel-time models. The first model consists of linear, empirical fits to LLNL phase picks (Figure 2); two other models are derived from regional surface-wave modeling (Patton and Taylor, 1984 and Priestly and Brune, 1978). Finally, we test IASPEI91 for comparison with a global average. The Pn arrivals reduced by predictions from the four models are shown in Figure 2b. The regression model fits best, followed by Patton & Taylor, Priestly & Brune, and IASPEI91. Ordered ranking of model fit to Pg and Lg arrivals (not shown) is the same as for Pn, and we use this order when assessing the importance of the travel-time model accuracy for multiple-event location accuracy.

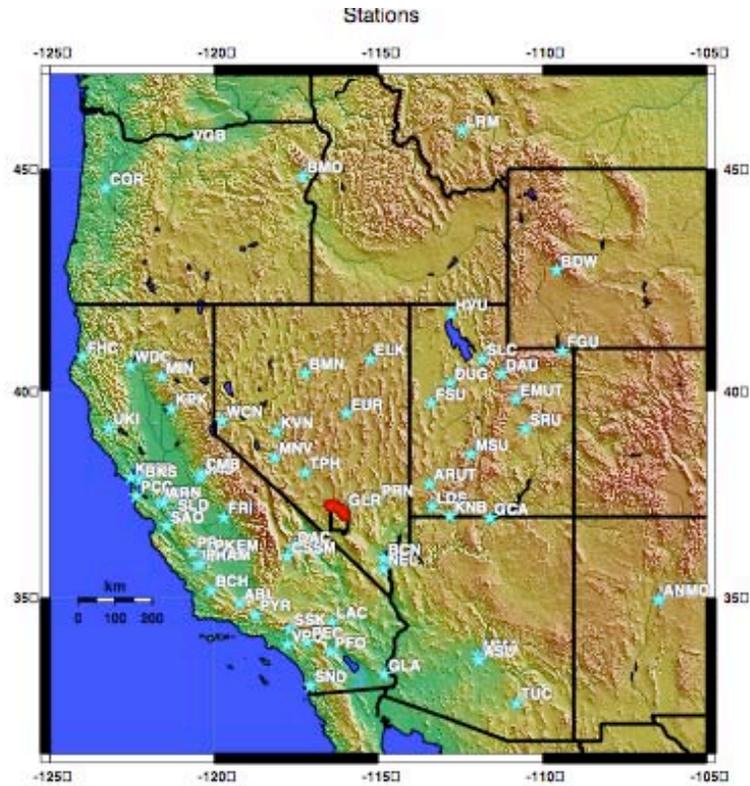


Figure 1. NTS explosions (red cluster of stars near the southern Nevada boarder) and regional stations used for relocation (blue cluster of stars shown near the northern region of Utah).

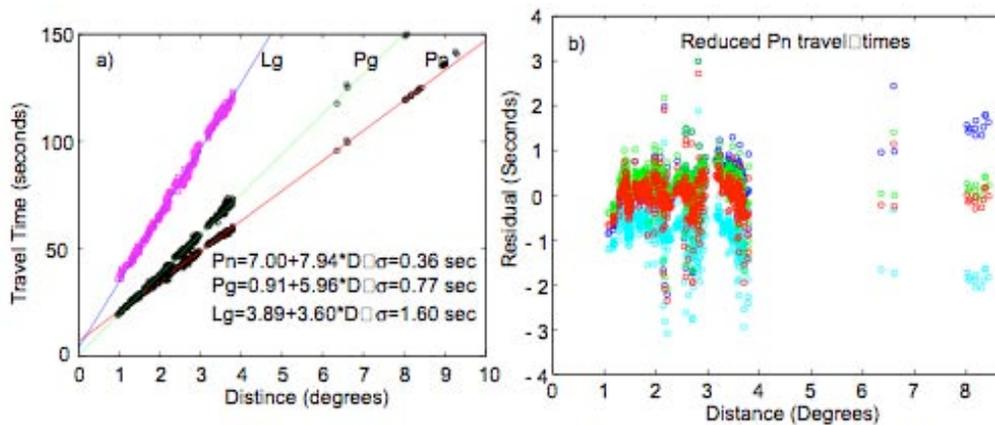


Figure 2. a) Empirical travel time curves for NTS explosions with travel-time models from linear regression. b) Pn arrival times reduced using IASPEI91 (blue), Priestly & Brune (cyan), Patton & Taylor (green) and a new LLNL model (red).

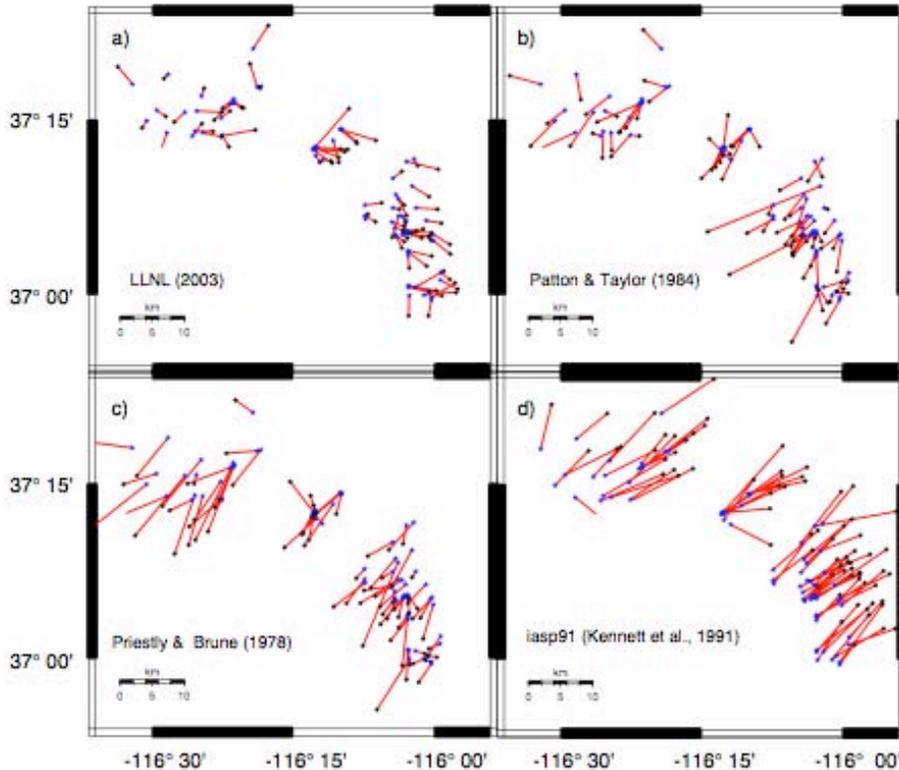
### Location algorithm

We use the Grid-search Multiple Event Location (GMEL) algorithm (Rodi et al., this Proceeding) and briefly note its highlights here. GMEL is formulated within a maximum likelihood estimation framework and implemented numerically with grid search and Monte Carlo techniques; it does not rely on the usual assumption of local linearity

of the forward problem. Station-specific travel-time corrections are determined using residuals from grid-search locations. This process is iterated to convergence. One of the relevant features of GMEL is the ability to constrain any number of events to the known location, therefore, supplementary information (e.g. satellite images) may be used to improve accuracy of an event cluster.

### Location experiments

We simultaneously relocate 74 NTS nuclear explosions with known hypocenters using the GMEL locator. New locations are determined using each of the four travel-time models described above. All hypocenters are unconstrained during inversion. We do, however, remove outlier arrival-time data, as identified using known hypocenter information, in an attempt to isolate model error. Figure 3 shows mislocations in map-view with corresponding mislocation histograms plotted in Figure 4. It is apparent that locations derived from the empirical travel-time curves are most accurate, with mean mislocation of 3.3 km and 95% of the locations within 6 km of the known epicenter. Considering the careful review of each phase arrival and the empirical nature of the travel-time curves, this degree of accuracy ( $\sim$ GT6) is likely to approach the lower bound of regional location accuracy for this dataset. We acknowledge that waveform correlation, which is not used for any of our arrival times, could reduce the uncertainty of arrival times, and thus location error.

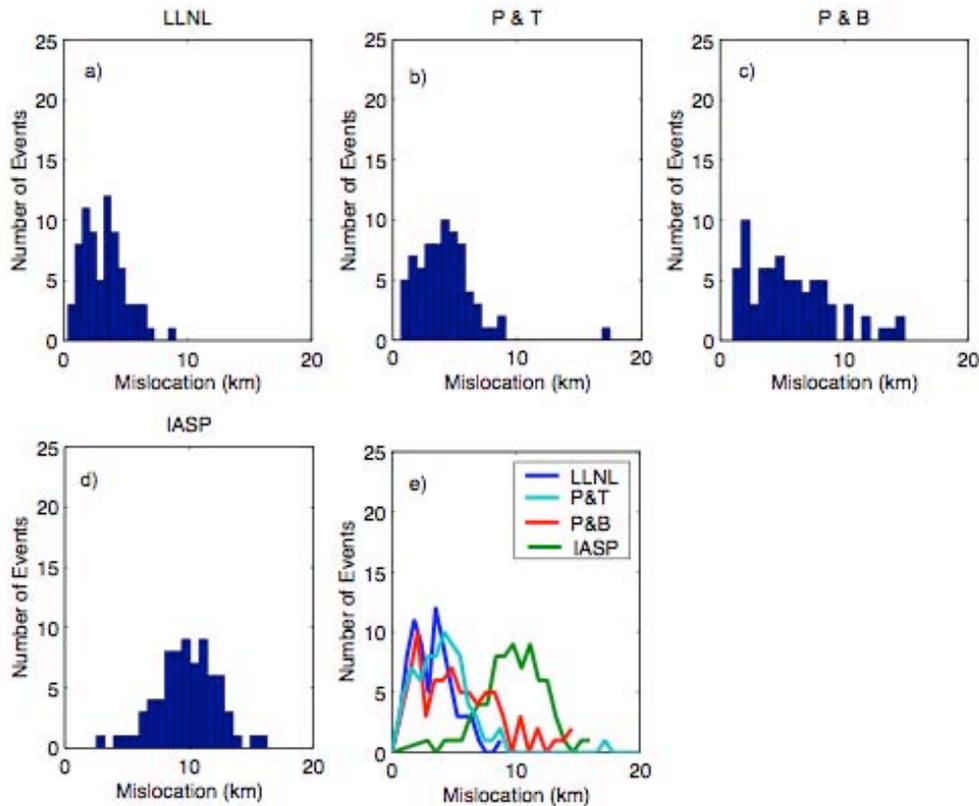


**Figure 3. Multiple-event location epicenter errors for 4 models a) LLNL empirical curves, b) Patton & Taylor, c) Priestly & Brune, and d) IASPEI91. Blue stars and black dots are known and estimated locations, respectively. Note how each model imparts a distinct vector bias.**

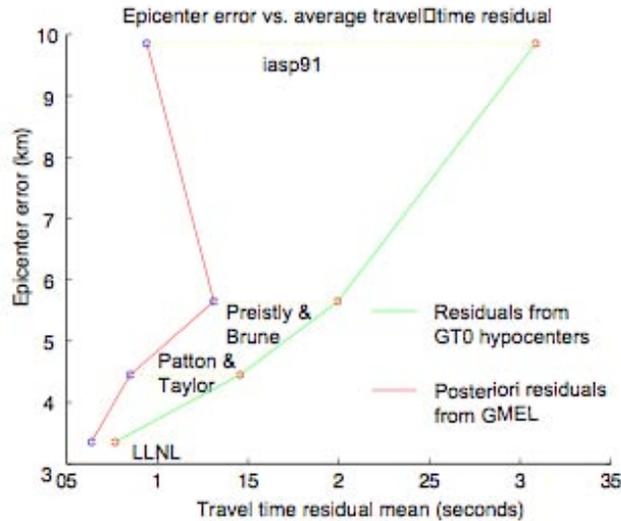
Similar to previous studies, we find that model error can contribute significantly to multiple-event mislocation, and this is apparent in Figures 3 and 4, parts b, c, and d. For the most part, increased error is manifested as a bias, with each model producing a distinct vector shift in the locations. The mean location accuracy of the Patton and Taylor, Priestly and Brune, and IASPEI91 models is 4.5 km, 5.6 km, and 9.9 km, respectively; the 95% bound on location accuracy is 8.7 km, 12.8 km, and 13.3 km, respectively. Using the error from the empirical curves as a baseline, we deduce that the model component of error is between 1.2 (Patton and Taylor) and 6.6 km (IASPEI91) in this study.

We note that in other parts of the world, observed regional IASPEI91 residuals can be more than twice that observed in the Basin and Range of Nevada. Therefore, regional application of a global model is probably unwise, even when using a multiple location algorithm and groomed dataset.

The correlation between travel-time prediction and epicenter error is shown in Figure 5. Epicenter error is plotted against “known” and estimated travel-time prediction errors. “Known” errors are determined using known hypocenter parameters, and estimated travel-time errors are derived from GMEL posteriori residuals. It is apparent that epicenter and known travel-time prediction errors are highly correlated, and that model-based, travel-time prediction error remains significant even when using a multiple-event algorithm. Perhaps more important is the lack of reliability in the correlation between estimated model error and location accuracy. Although the first three models suggest a weak correlation between estimated model error and location accuracy, the IASPEI91 point significantly deviates from this trend. We conclude, therefore, that travel-time residuals should not be used to access location accuracy. This conclusion is consistent with previous studies (e.g. Myers and Schultz, 2000; Bondar et al., 2003), where confidence ellipses (based on posteriori residual distributions) significantly underestimate true location error.



**Figure 4. Histograms of epicenter error for the 4 velocity models. a) LLNL empirical travel-time curves. b) Patton and Taylor (1994). c) Priestly and Brune (1978). d) IASPEI91. and e) is a composite of parts a through d for comparison.**



**Figure 5. Mean epicenter error is plotted against mean absolute deviation (MAD) of travel-time prediction error for each of the 4 models. All phases are used for location (Pn, Pg, and Lg). The green line connects points determined using “known” travel-time errors (i.e., determined using known hypocenters). The red line connects points determined using relocation posteriori residuals. The green line shows a clear correlation between locations and “known” travel-time error, however, the correlation is not as clear when posteriori residuals – which are necessarily used in earthquake studies – are used to assess travel-time prediction error.**

Lastly, our regional results suggest better location accuracy than the broader study of Bondar et al. (2003). Bondar et al., (2003) conclude that regional events with good azimuthal coverage (secondary azimuthal gap <math><120^\circ</math>) are accurate to within 20 km at 95% confidence. In the present study we find that IASPEI91 produces locations with accuracy of ~13km at 95% confidence. The discrepancy is probably due to a combination of factors, including: the use of analyst-reviewed arrival times; the use of the multiple-event location algorithm, which reduces errors attributable to 3-dimensional Earth structure; and IASPEI91 is better suited to the Basin and Range than many other regions of the world.

### **Subspace Detector**

The master-event waveforms from the 40-event cluster were decomposed into time-series basis functions. A subset of the basis is used as the signal representation in the subspace detector. The size (dimension) of the subset is chosen to maximize the probability of detection and minimize false alarms. Figure 8 shows one of the diagnostics used to determine the size of the basis subset. It displays the energy capture, i.e. the fraction of waveform energy contained in the basis (least-square) representation of the waveforms as a function of the selected dimension of the representation. The figure displays 40 separate curves in blue to represent the energy capture for each of the 40 master event waveforms. The single red curve represents the average energy capture for the suite of events. A significant number of the signals are not well represented except at very large subset dimension (approaching 40); this fact suggests a high degree of diversity among the signals generated by this mine and potentially poor correlation detector performance.

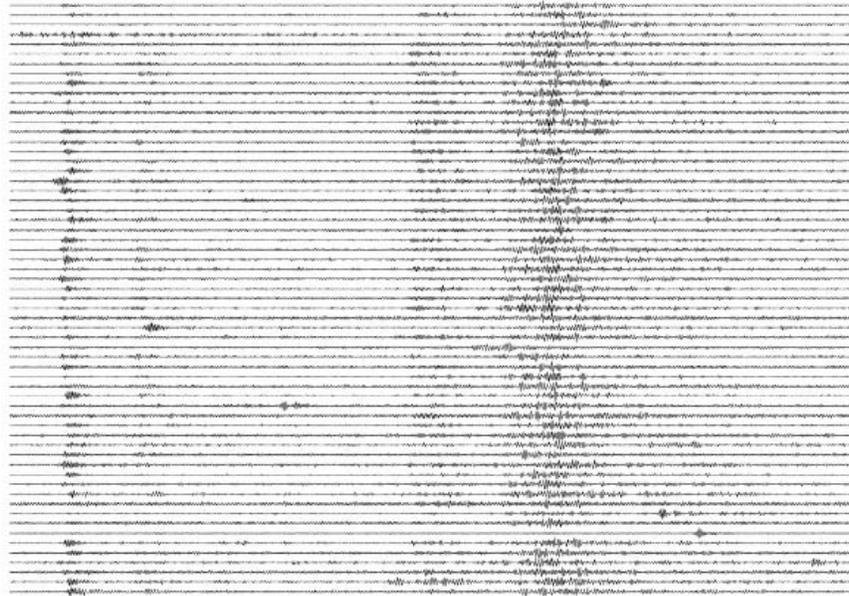


Figure 6. Master waveforms for the 61 event of the Kirovsk mine recorded at ARCES.

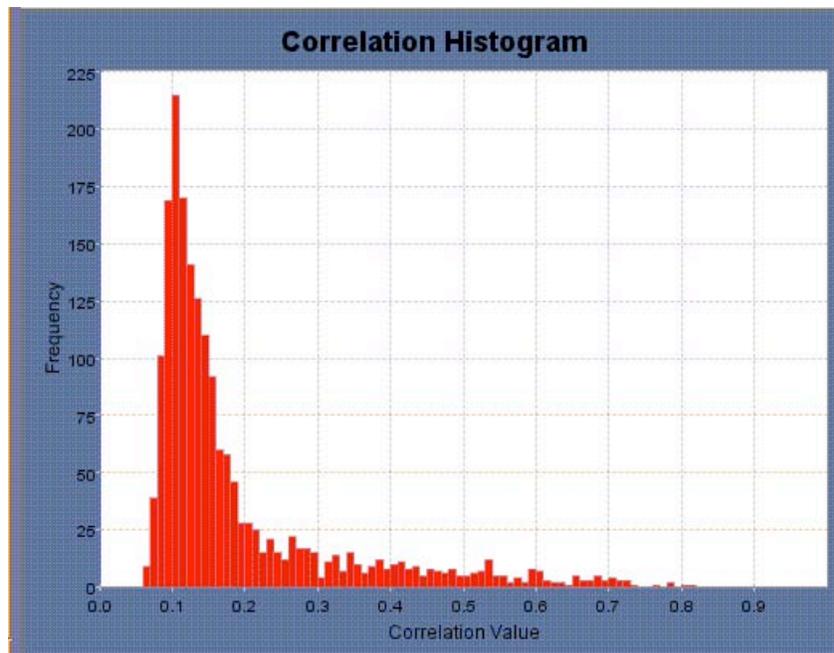
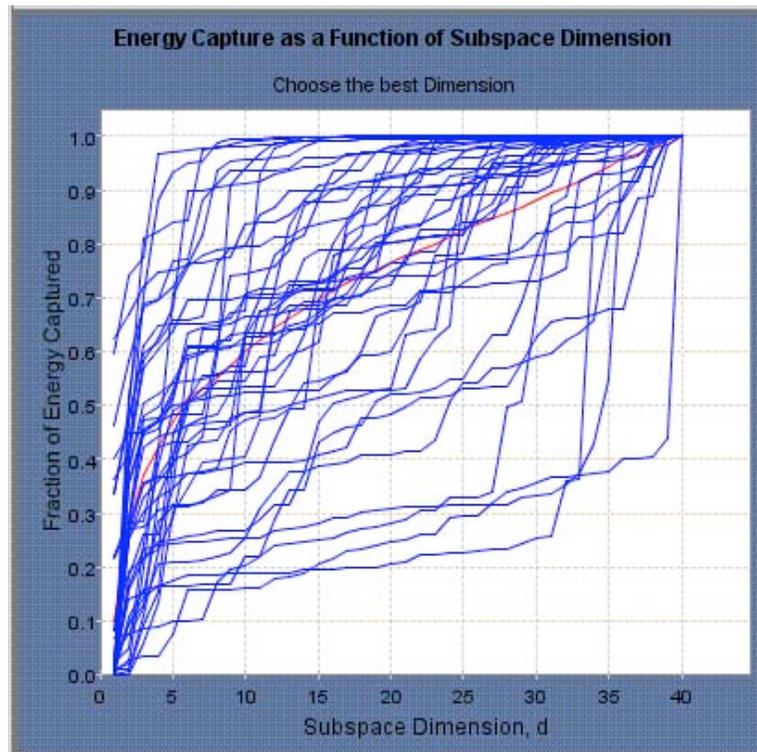


Figure 7. Histogram of correlation values for the 61 master events. Despite our efforts to select events with similar firing practice, waveform correlation tends to be low. The subspace detector makes use of the disparate waveforms to span the possible waveform characteristics for a particular mine.

The master-event waveforms from the 40-event cluster were decomposed into time-series basis functions. A subset of the basis is used as the signal representation in the subspace detector. The size (dimension) of the subset is chosen to maximize the probability of detection and minimize false alarms. Figure 8 shows one of the diagnostics used to determine the size of the basis subset. It displays the energy capture, i.e. the fraction of waveform energy contained in the basis (least-square) representation of the waveforms as a function of the selected dimension of the representation. The figure displays 40 separate curves in blue to represent the energy capture for each of the 40 master event waveforms. The single red curve represents the average energy capture for the suite of events. A

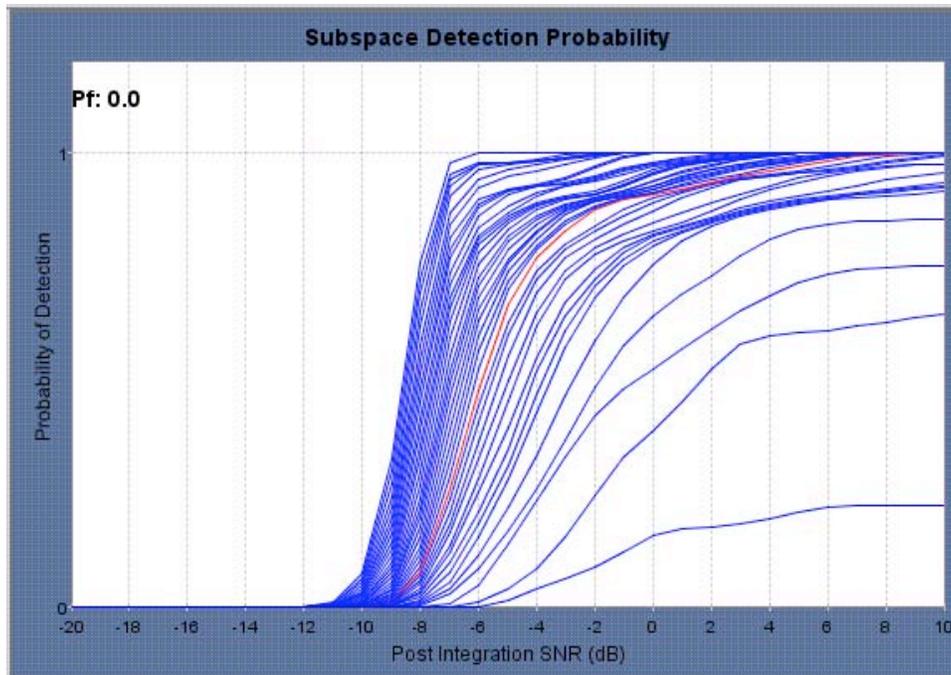
significant number of the signals are not well represented except at very large subset dimension (approaching 40); this fact suggests a high degree of diversity among the signals generated by this mine and potentially poor correlation detector performance.



**Figure 8. Fraction of energy captured for each of the master-event waveforms is plotted against a trial subspace detector with varying dimension. This analysis is used to select detector subspace dimension (see text).**

A theoretical probability of detection can be calculated for a subspace detector constructed with a basis of given dimension. The operating assumption is that the detector conducts a binary hypothesis test on the presence of white Gaussian noise (the null hypothesis) in a detection window, or on noise plus a signal represented as a linear combination of the basis functions (the alternate hypothesis). Figure 9 shows a suite of theoretical probability of detection curves, one for each possible choice of subspace dimension. The single curve highlighted in red shows the probability of detection for a subspace detector of dimension 15 generated from the 40 master event waveforms, for a fixed threshold (0.4). The probability of detection curve actually is a hybrid, consisting of the average probability of detection for the collection of 40 events assuming all were equally likely. At the selected threshold, the probability of a false alarm is vanishingly small.

This detector was run on a data window consisting of the first 40 days of 2003 for a limited number of ARCES elements. The detector made nine detections for this time interval, missing 19 events that are known to have occurred at the Kirovsk mine during this interval. The false alarm rate was zero, as expected. Our result suggests that for the Kirovsk mine, the variability of firing practice and geographic distribution of sources produces an exceedingly diverse set of waveforms. It is possible that a master-event set drawn from a longer time window is needed to capture the full waveform variability.



**Figure 9.** The probability of detection for subspace detectors is a function of the order of the detector. A subspace detector designed from 50 master event waveforms can have an order ranging anywhere from 1 to 40. This suite of curves displays the theoretical probability of detection for each possible detector order. The red curve corresponds to the order 15 detector actually used in our attempt to detect Kirovsk compact underground mining events.

## **CONCLUSIONS AND RECOMMENDATIONS**

### **NTS Regional-Network Location Study:**

- 1) For our groomed dataset, empirical travel-time curves (derived using the known NTS hypocenters) produce GT6 locations. (Note: our dataset does not include correlation picks, which— where applicable – may further improve location accuracy).
- 2) Two other Basin and Range models (based on surface-wave modeling) produce GT9 and GT13 locations. The IASPEI91 model produces GT14 locations.
- 3) Epicenter and “known” travel-time prediction accuracy (derived from known NTS-explosion hypocenters) are well correlated. However, the correlation between epicenter and estimated travel-time prediction (multiple-location posteriori residuals) is not reliable. Therefore, assessment of location accuracy based on travel-time fit is also not reliable.

### **Subspace Detector**

- 1) Subspace detection is a valuable tool for detecting and screening on-going, mine-related seismicity.
- 2) For our case example (Kirovsk mine) the false alarm rate is zero but detection rate is currently about one third.
- 3) For the Kirovsk mine a master-event set spanning a long time window (current set is from a 3-month period) may be needed to improve the detections rate.

## **ACKNOWLEDGEMENTS**

We continue to enjoy a dialog on seismic location research with William (Bill) Rodi, who also provided the GMEL location codes. Much of the NTS dataset was developed in cooperation with Kenneth D. Smith at University of Nevada, Reno.

## **REFERENCES**

- Armbruster, J., V. Burlacu, M. Fisk, V. I. Khalturin, W.Y. Kim, I. Morozov, E. Morozova, P. G. Richards, D. Schaff, and F. Waldhauser, Seismic Location Calibration For Thirty International Monitoring System Stations In Eastern Asia, "*Proceedings of the 24th Seismic Research Review: Nuclear Explosion Monitoring: Innovation and Integration*", 2002, Los Alamos National Laboratory LA-UR-02-5048, pp 231-241. Ponte Verde, FL.
- Bondar, I., S.C. Myers, E.R. Engdahl, E. Bergman, Epicenter accuracy based on seismic network criteria, *Geophys. Jour. Int.*, In Press, 2003.
- Engdahl, E. R., E. A. Bergman, M. H. Ritzwoller, N. M. Shapiro and A. L. Levshin, A Reference Data Set For Validating 3-D Models, "*Proceedings of the 24th Seismic Research Review: Nuclear Explosion Monitoring: Innovation and Integration*", 2002, Los Alamos National Laboratory LA-UR-02-5048, pp 261-270, Ponte Verde, FL.
- Myers, S. and C. Schultz, Improving Sparse-Network Seismic Location with Bayesian Kriging and Teleseismically Constrained Calibration Events, *Bull. Seism. Soc. Am.*, 90, 199-211, 2000. Patton, H.R., S.R. Taylor, Q-Structure of the Basin and Range from Surface-Waves, *Jour. Geophys. Res.*, 89 (NB8): 6929-6940 1984.
- Priestley K.F., J. N. Brune, Surface-Wave Dispersion In Great Basin Of Nevada And Western Utah, *Eos T Am Geophys Un* 57 (12): 951-951, 1976
- Rodi, W, C. A. Schultz, W. G. Hanley, S. Sarkar, and H. S. Kuleli, Grid-Search Location Methods For Ground-Truth Collection From Local And Regional Seismic Networks, "*Proceedings of the 24th Seismic Research Review: Nuclear Explosion Monitoring: Innovation and Integration*", 2002, Los Alamos National Laboratory LA-UR-02-5048, pp 394-402, Ponte Verde, FL.
- 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.
- Walter, W.R., K. D. Smith, J. L. O'Boyle, T. F. Hauk, F. Ryall, S.D. Ruppert, S.C. Myers, M. Anderson, and D.A. Dodge, Improving The Fundamental Understanding Of Regional Seismic Signal Processing With A Unique Western United States Dataset, *Proceedings of the 25th Seismic Research Review: Nuclear Explosion Monitoring: Building the Knowledge Base*, Tucson, AZ, 2003.