SEISMIC SOURCE LOCATIONS AND PARAMETERS FOR SPARSE NETWORKS BY MATCHING OBSERVED SEISMOGRAMS TO SEMI-EMPIRICAL SYNTHETIC SEISMOGRAMS: IMPROVEMENTS TO THE PHASE SPECTRUM PARAMETERIZATION

David. Salzberg and Margaret Marshall

Science Applications International Corporation

Sponsored by Air Force Research Laboratory

Contract No. FA8718-05-C-0019

ABSTRACT

The purpose of this study is to demonstrate the feasibility of full-waveform earthquake location using semi-empirical synthetic waveforms and received data from as few as two regional stations. We previously used a parameterization of the semi-empirical correction (or filter) that assumed range invariant phase spectrum of the correction term. That approach works very well if the two events are close (<10 km). However, with greater event separation, the empirical filter is unable to correct the synthetic to match the data, particularly at higher frequencies (Salzberg et al., 2005).

By modifying the parameterization of the semi-empirical correction to allow for a range dependent phase spectrum, the empirical filter is able to correct the synthetic to match the data even when the event separation is large (> 50 km), and in particular, in cases where the range invariant phase correction fails. The improved parameterization characterizes the spectrum in terms of wave number and range. This approach requires us to isolate individual propagating modes as such. Currently, only the fundamental Rayleigh wave is used. In this method, the ray-theory P-wave arrival time is used for time reference; the synthetic is lagged by the predicted arrival time, and the data uses the measured arrival time. With this technique, we are able to obtain GT5 or better locations using the long-period (>30 second) Rayleigh waves. Also, we are able to simulate the waveforms from a nuclear explosion (JUNCTION) using the Little Skull Mountain Earthquake.

We will present comparisons of the results using both the new approach and the old approach by utilizing ground truth (GT) for a data set of events in California. In particular, we will determine the range at which the accuracy of the new approach (using the long period surface waves) exceeds the old approach (using broadband signals). In addition, we will determine the separation at which the new phase parameterization is no longer able to simulate the observations.

OBJECTIVE(S)

The objective of the research is to provide a method that gives accurate locations (GT5) and source mechanisms using a sparse regional network when two or more seismic stations record the event. Tests using synthetic waveforms indicate that location accuracy on the order of 300–500 km^2 and depth uncertainty of less than 5 km can be obtained with recordings from only two stations low pass-filtered at 0.5 Hz using hold waveforms, or .1 Hz when using just the Rayleigh waves. The objective if this paper is to present the semi-empirical (or empirically filtered) method using a range-dependent empirical filter, and to evaluate the results with the new parameterization when compared to the old parameterization.

RESEARCH ACCOMPLISHED

- Measured the Group Velocity Curves of all reviewed waveforms
- Identified two workable clusters
- Determined the 1-D path specific velocity models
- Processed events to determine frequency dependence of the method
- Identified and corrected problems with the original methodology
- Demonstrated the ability to locate earthquakes and explosions to better than GT5
- Demonstrated the ability to locate small events (M = 4.1)
- Demonstrated the ability to transform waveforms with different mechanisms

Approach

In this section we present our technique for empirical filtering. First, an overview of our approach is presented, followed by a derivation of the approach. The overall goal of our approach is to find a filter that transforms synthetic seismograms into the observed data. When comparing the synthetic waveform with real data, a match is not possible unless the velocity model is well defined. Furthermore, errors in the synthetic waveforms translate into errors in the resolved location and source parameters. Consequently, the primary limitation in matched waveform processing is the ability to produce high-fidelity simulations based on an imperfect environmental model.

One approach that can be used to improve the quality of the synthetic waveforms is to empirically determine a filter that characterizes the propagation portion of the synthetic waveform, yet still use the theoretical excitation (Salzberg, 1996; Velasco et al., 1994). These empirically filtered Green's Functions can then be employed to characterize the seismic wavefield recorded at a seismic station or array using sources from a reasonably homogeneous source region.

In our ongoing project, we defined the filter as a static function of frequency. This approach works well if the two events are co-located. However, any significant separation in range yields a poor empirical filter. This is caused by a phase shift from the difference in propagation distances. To compensate for this, we propose to represent the phase of the empirical filter in wave number space. With this approach, the empirically filtered synthetic seismogram, u_e is

$$u_e\left(\omega,r\right)=\frac{s\left(\omega\right)}{u\left(\omega,r_o\right)}\cdot e^{i\Delta k\left(r-r_o\right)}\cdot u\left(\omega,r\right), \tag{1}$$

where, u is the synthetic seismogram, s is the observed seismogram, r_o is the range to the reference event, r is the range to the new event, and Δk is the wave number corresponding to the phase mismatch reference data and synthetic.

This change significantly improves our ability to match the data to the empirically filtered synthetic seismogram, when there is a substantial distance between the new event and the reference event; or, when r-r_o is large (50 km or more) compared to the static phase approach. The approach will, of course, also work at shorter ranges. However, when the events are close, r-r_o approaches 0, and the formulation reverts to

$$u_e(\omega,r) = \frac{s(\omega)}{u(\omega,r_o)} \cdot u(\omega,r), \tag{2}$$

or

$$u_e(\omega,r) = \Delta u(\omega,r_o) \cdot u(\omega,r), \tag{3}$$

where

$$\Delta u(\omega,r_o) = \frac{s(\omega)}{u(\omega,r_o)}, \tag{4}$$

which is the approach used in our current Air Force Research Laboratory (AFRL) project.

Proposed Wave Number (K-Space Parameterization)

In this section, the *k-space* parameterization will be described. At the end of the derivation, the resulting formulation is Equation (2).

Parameterizing our classic formulation, shown in Equation (4) in terms of amplitude and phase yields:

$$\Delta u(\omega,r_o) = \left|\frac{s(\omega)}{u(\omega,r_o)}\right| \cdot e^{i\left(\phi_s - \phi_{u_o}\right)} = \Delta A(\omega) \cdot e^{i\Delta\phi}. \tag{5}$$

This approach works well if the events are close together, or if there is a dominant phase in the data (e.g., Rayleigh). In the case where the spatial separation of the events approaches or is greater than the wavelength of the highest frequency energy used, the increasing time separation of the body waves and phase mismatch of the surface waves result in increased misfit, as increases occur in the separation of the reference and the second event.

An alternative, range dependent phase parameterization will provide better agreement of the waveforms in these cases. Treating the waveform as a sum of propagating modes, the modes can be represented as a range dependent amplitude and range dependent phase, or

$$u(\omega,t,k,r) = \sum_n A_n(\omega,r) \cdot e^{i\left(k_n r - \omega t\right)}, \tag{6}$$

where u is the displacement, A is the amplitude, k is the wavenumber, and n is the mode.

Using this formulation for both the data and the synthetic waveform (and assuming that the modes of the data can be separated), Equations (5) and (6) can be combined to parameterize the phase misfit to terms of the wave number and distance, or

$$\Delta\phi = \Delta\phi_{reference} - \Delta\phi_{synthetic} = \left(k_{reference} \cdot r_{reference} - \omega \cdot t_{reference}\right) - \left(k_{synthetic} \cdot r_{synthetic} - \omega \cdot t_{synthetic}\right). \tag{7}$$

Assuming the reference time for the data and the synthetic is the same, and the synthetic was computed for the observed distance, then

$$\Delta\phi = \left(k_{reference} \cdot r - \omega \cdot t\right) - \left(k_{synethic} \cdot r - \omega \cdot t\right) = \left(k_{reference} - k_{synethic}\right) \cdot r = \Delta k \cdot r. \tag{8}$$

For each propagating mode, assigning the reference event range to be r_o, allows the empirically filtered waveform for the specific propagating mode at range r to be expressed as

$$u_e(\omega, r) = \frac{s(\omega)}{u(\omega, r_o)} \cdot e^{i\Delta k(r - r_o)} \cdot u(\omega, r), \tag{9}$$

which is Equation (1).

With real data, it is not feasible to separate each propagating mode exactly. However, it is possible to apply time-frequency windowing functions to isolate the body waves from the surface waves.

Examples of the Approach

Two Components of NVAR

This process is demonstrated using two broadband components of the NVAR recording a Parkfield, CA, aftershock, as shown in Figure 1. This test is desirable because, with multiple components recording the same wavefield, source uncertainty is eliminated. Computing synthetic waveforms at both distances, and using the observed Rayleigh wave from NV32 (range = 328km) to determine the filter to transform the 344 km synthetic (which is the distance to NV33), it is possible to accurately simulate the waveform. Comparing the synthetic using the range dependent phase correction to the existing constant phase technique shows that the new approach does a better job simulating the Rayleigh wave than using the existing constant phase method, as shown in Figure 2.

Furthermore, the correct distance separation between the stations can be determined by matching the Rayleigh waveforms (with time relative to the P arrival time), as shown in Figure 3.

San Simeon Event Using Parkfield as a Reference Event

In our presentation at the 2005 SRR (Salzberg et al., 2005), we presented an example where the semi-empirical synthetic seismogram location approach was unable to resolve the event location; those results are shown in Figure 4. With the range-dependent formulation, we are able to correctly resolve the event location, as shown in Figure 5, even though the San Simeon Event was small (M_w=4.1), relatively far from the reference event (60 km), and a difference mechanism (Reverse vs. Strike Slip).

Sensitivity Analysis of GT Events (NTS Nuclear Explosions)

In order to evaluate the solution quality of the approach, we have applied it to Nevada Test Site (NTS) nuclear explosions recorded at PAS and Albuquerque (ANMO). Our primary concern is the sensitivity using different frequency bands. As such, the data were processed using surface waves over a variety of frequency bands. We found that we were able to locate the event HOYA using JUNCTION as a reference event in range and depth to within one grid spacing of GT, as shown in Figure 6. The uncertainty in our estimates, though, were on the order of 2–4 km, with the area of ambiguity being about 5 km^2 (2 σ) or 20 km^2 (3 σ), in when looking at broadband surface waves (filtered to periods between 100 and 7 seconds) as shown in Figure 7.

Locating Little Skull Mountain Earthquake Using NTS Explosions as Reference Events

As a final example, the NTS explosion JUNCTION is used as a reference event for locating the Little Skull Mountain (LSM) earthquake. The geographic context of the events and the station is shown in Figure 8. Our two-dimension (range and depth) examples show that we were able to approximately locate the LSM earthquake, as shown in Figure 9. The absence of GT renders it impossible to assess the accuracy of the results.

CONCLUSIONS AND RECOMMENDATIONS

The range-dependent (wave number based) parameterization of the semi-empirical synthetic approach allows for the waveform-matched location to within a few kilometers using P-wave arrival times for reference and the Rayleigh waveform, though the quality of the depth resolution is uncertain. With this, earthquakes and explosions can be relocated using GT events (either earthquakes or explosions) as reference events. In order to improve the technique to provide for better depth estimates, the P-wave and S-wave arrivals (and depth phases) should be included in the processing.

ACKNOWLEDGEMENTS

All waveform data was obtained from the Incorporated Research Institutions for Seismology (IRIS) Data Management Center. The International Data Center (IDC) event locations were from the SMDC Monitoring Research Web Page.

REFERENCES

Salzberg, D. H. (1996). Simultaneous inversion of moderate earthquakes using body and surface waves: Methodology and applications to the study of the tectonics of Taiwan, Ph.D. thesis, State University of New York, Binghamton, NY.

Salzberg, D. H., K. E. Votaw, and M. E. Marshall (2005). Seismic source locations and parameters for sparse networks by matching observed seismograms to semi-empirical synthetic seismograms, in *Proceedings of the 27th Seismic Research Review: Ground-Based Nuclear Explosion Monitoring Technologies*, LA-UR-05-6407, Vol. 1, 454–462

Velasco, A. A., C. J. Ammon, and T. Lay (1994). Recent large earthquakes near Cape Mendocino and in the Gorda Plate: Broadband source time functions, fault orientations and rupture complexities, *J. Geophys. Res*. 99: 711–728.

FIGURES

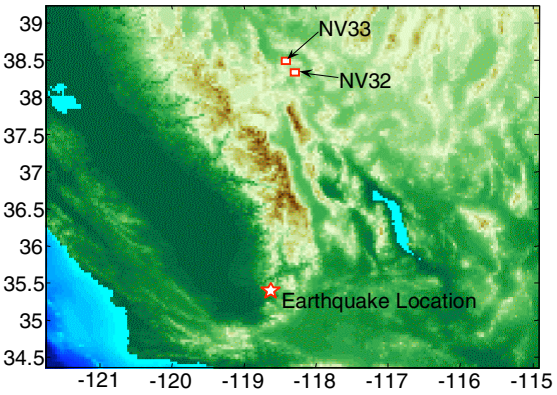

Figure 1. The geographic context of the two NVAR components to the earthquake.

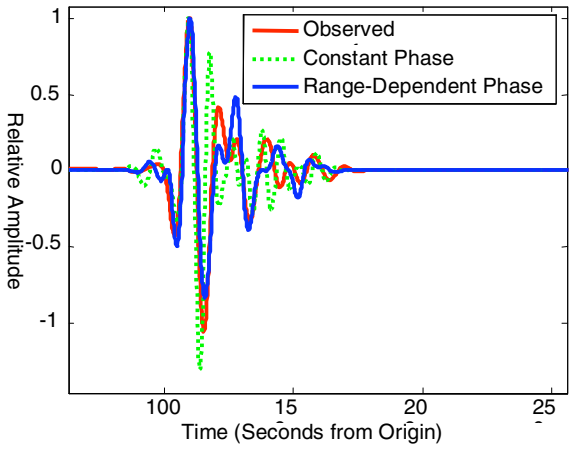

Figure 2. The observed Rayleigh wave recorded at NV33 compared with semi-empirical synthetic based on the constant-phase approach and the range-dependent approach. Note the range-dependent phase (based on wave number) filter does a much better job of fitting the observed data than does the synthetic waveform.

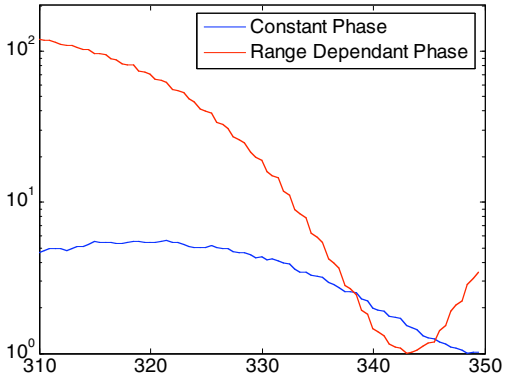

Figure 3. The optimal range determined using the constant phase approach and the range-dependent phase approach. The range-dependent phase is able to resolve the location of NV33 accurately, as the GT range is 344 km.

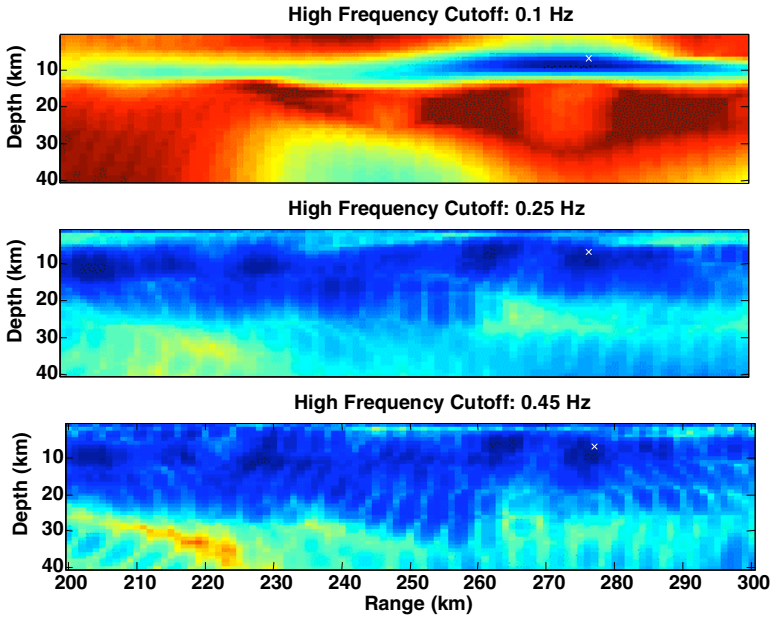

Figure 4. The range-invariant empirical filter only poorly resolves the event location when the new event and reference event are more than a few kilometers. At high frequencies, a minimum was not obtained. At low frequencies (below 0.1 Hz), a broad minimum was obtained. However, the uncertainty is approximately 30 km.

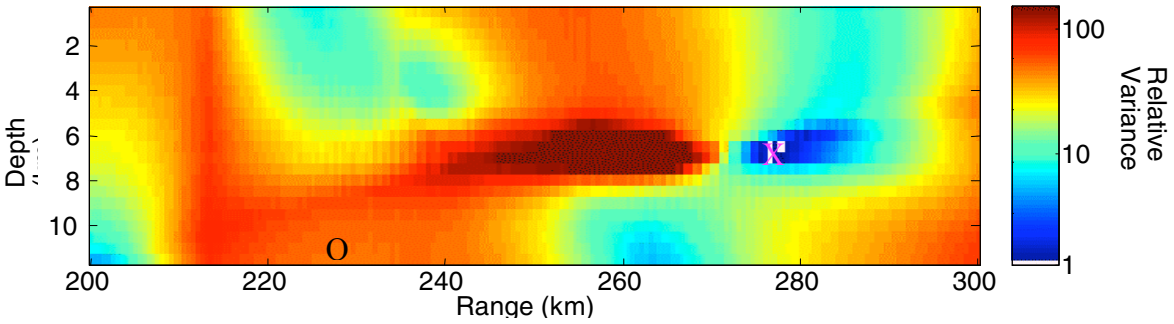

Figure 5. Relocating the same event discussed in Figure 4, but using the range-dependent (wave number based) approach. X marks the GT location, which coincides with the minimum in variance. The reference event, shown by the black O is approximately 50 km. Given the sharpness of the minima, it is clear that the range-dependent approach did obtain a good solution for the event.

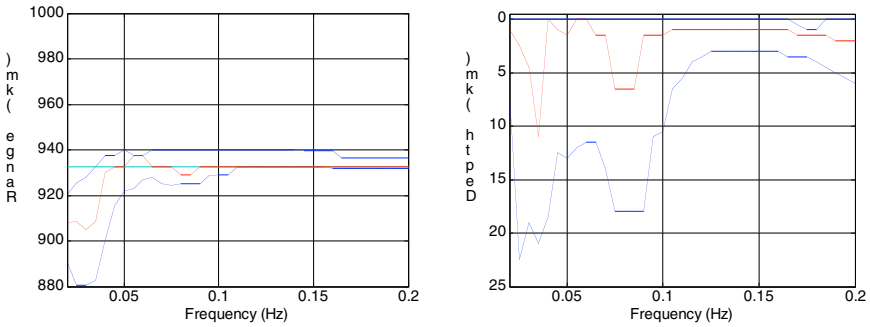

Figure 6. The location sensitivity vs. processing frequency of the HOYA NTS nuclear explosion, using the JUNCTION explosion as a reference event. The blue lines represent the 2 σ errors, based on variance, and the red line is the location of the minimum variance. GT is shown by the cyan line on the range plot and is approximately 300 m on the depth plot. From this, it is clear that the most accurate and precise event locations occur when higher-frequency data are used.

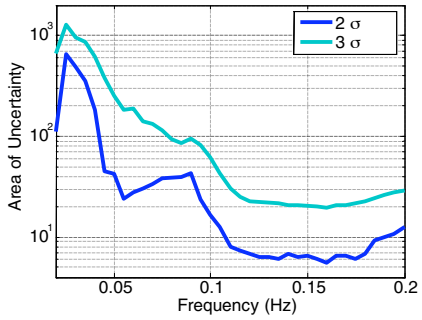

Figure 7. The 2-D area of uncertainty of the HOYA NTS nuclear explosion, using the JUNCTION explosion as a reference event. Note that the minimum uncertainty, which is about 5 km^2, occurs when using data with a low-pass filter of around 7 s (0.15 Hz).

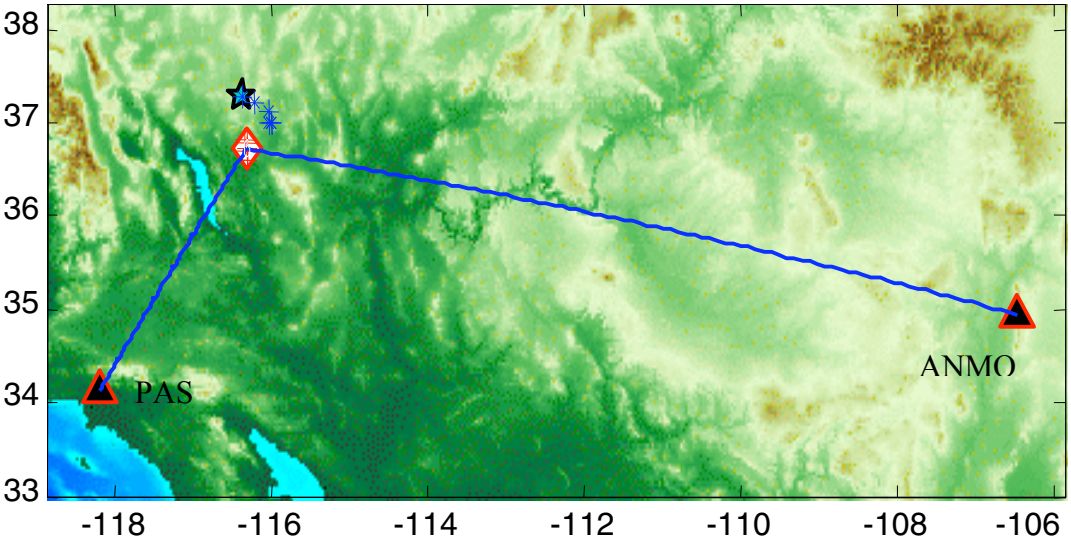

Figure 8. The geographic context of the NTS explosions (blue) and the LSM earthquake (red diamond).

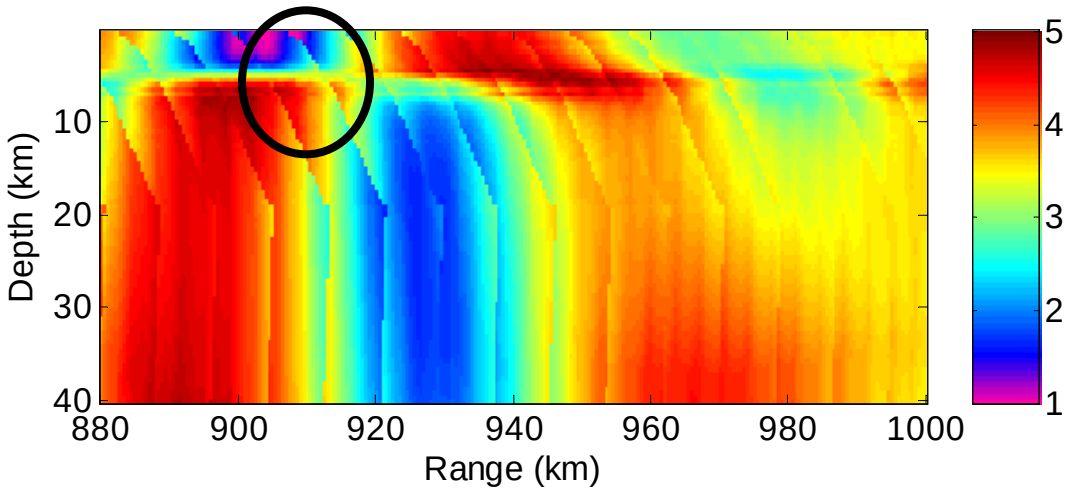

Figure 9. Locating the LSM earthquake using the JUNCTION explosion as a reference event. The GT location is shown by the black ellipse centered near 910 km in range. Note that the location uncertainty overlaps with minimum in residual (magenta region).