SEISMIC VELOCITY ESTIMATION IN THE MIDDLE EAST FROM MULTIPLE WAVEFORM FUNCTIONALS: P & S RECEIVER FUNCTIONS, WAVEFORM FITTING, AND SURFACE WAVE DISPERSION

Jay Pulliam¹,², Mrinal K. Sen², Nathan Simmons³, and Michael Pasyanos³

Baylor University¹, The University of Texas at Austin², and Lawrence Livermore National Laboratory³

Sponsored by the National Nuclear Security Administration

Award Nos. DE-AC52-09NA29327¹,² and DE-AC52-07NA27344³
Proposal No. BAA09-58

ABSTRACT

A novel velocity estimation algorithm that will be developed and applied that incorporates the constraints of three seismic functionals: Receiver functions, waveform modeling, and surface wave dispersion measurements. These functionals have distinct but complementary sensitivities to Earth structure and have all been employed in modeling studies previously, yet each fails to resolve ambiguities between realistic models when used alone. By incorporating constraints from all three functionals we can minimize the part of the model space for which data constraints fail to distinguish between models. While several recent studies attempt to model two of these functionals jointly, none explores or assesses adequately the improvements in model reliability that are gained by incorporating additional constraints.

The general search techniques and assessment tools we will employ allow us to describe the size and characteristics of the acceptable model space, proscribe the set of unacceptable models, and identify characteristics that are required by the data, rather than those that are simply consistent with the data. As a result, the method will quantitatively evaluate the strengths of the constraints imposed by each data functional on each model parameter. The Bayesian formulation we propose provides a formal mechanism for incorporating the results of one inversion (as a “prior”) into the inversion of a second type, to produce a multi-step “joint” (as opposed to “simultaneous”) modeling procedure. The final results will be models that are accompanied by detailed, quantitative assessments of model reliability. In addition, these tools will guide us toward the types and characteristics of additional data that should be sought in order to improve model constraints and, therefore, model reliability.

Another benefit of the proposed modeling method is that it will allow seismologists to produce models for relatively aseismic regions of the world, where station density is sparse, and where, for example, one particular data functional is rarely observed. It will therefore result in a method that is more broadly applicable than methods that rely upon one data functional alone.

Lastly, numerous areas have been studied with “campaign”-style temporary deployments, which often produce a relatively small dataset compared to permanent stations. Broadening the range of data types employed in modeling will generally produce more reliable models for sparsely sampled regions. We will ultimately apply our modeling to stations in the Middle East. Besides being an area of monitoring interest, the Middle East is tectonically complex. As such, it is a region where fitting multiple datasets can choose among various non-unique models, as well as reconcile seemingly inconsistent data.

This award begins September 1, 2009, so in this paper we describe our approach and goals for this project.
OBJECTIVES

Knowledge of crust and upper mantle velocity structure is essential to locating seismic events and, specifically, for constraining event focal depths. Seismic locations, discriminants, and yield estimates are all affected negatively by errors in seismic velocity models. Accurately assessing and quantitatively describing model errors is a critical component of estimating velocity and attenuation structure and such assessment is especially difficult in nonlinear modeling. Here we propose a multi-step modeling procedure for the estimation of crust and upper mantle velocity models using distinct types of regional and teleseismic waveform data.

One goal of the BAA to which this proposal responds is to develop optimal procedures for the use of multiple datasets. Due to the inherent variability, inconsistency, and peculiarities of disparate datasets and the well-known nonlinearity and non-uniqueness associated with geophysical modeling, such procedures must include methods for evaluating the performance and contribution of each dataset to the final results.

Our approach makes use of quantitative assessment tools and a well-developed Bayesian approach to explore and evaluate each step of the modeling process, rather than to simply toss all constraints into a simultaneous fitting procedure to find the single solution that satisfies particular criteria. The procedure we propose, best characterized as velocity analysis via optimization, is analogous to velocity analysis in exploration seismology, rather than “inversion”. It will provide quantitative error measures of structural parameter estimates that can then be translated to earthquake location errors, and thus guide seismologists toward the most effective and efficient ways to improve model reliability.

RESEARCH ACCOMPLISHED

We have already made substantial progress on several aspects of the proposed research: waveform modeling, receiver functions, and one proposed variant of simulated annealing. During the last several years we developed, tested, and applied the method for waveform modeling via global optimization described above (Gangopadhyay et al., 2007; Pulliam and Sen, 2005; Pulliam et al., 2006). Earlier we developed, with Lian-She Zhao and Cliff Frohlich, a VFSA-based inversion algorithm for P-wave receiver functions (Zhao and Frohlich, 1996; Zhao et al., 1996). Additional tools developed at UTIG include fast, parallelized reflectivity code that incorporates anisotropic model parameters, additional global optimization methods based upon genetic algorithms, and numerous imaging algorithms.

Specifically with regard to waveform modeling: We developed a code that computes synthetic seismograms with a parallelized reflectivity method and fits the observed waveforms by global optimization. Assuming a one-dimensional, isotropic, layered Earth, our code computes the synthetic seismograms independently for all layers, frequencies, and ray parameters. The code implements a global optimization algorithm using Very Fast Simulated Annealing (VFSA) that allows for searching within a broad spectrum of models in order to find the global minimum, and hence minimizes dependency on the starting model. Our method also computes the Posterior Probability Density (PPD) function and correlation matrix, to evaluate the uniqueness of the resulting models, and the trade-offs between individual model parameters therein. We applied the code to determine the crust and upper mantle structure beneath permanent broadband seismic stations in Africa using teleseismic earthquakes (M 5.5–7.0 and 200-800 km focal depth) recorded at these stations. We modeled the S, SP, SsPmP, and shear-coupled PL waves from these earthquakes and our P- and S-wave velocity models compare well with, and in some cases improve upon the models obtained from other existing methods. We obtained P- and S-wave velocities simultaneously and our use of the shear-coupled PL phase wherever available improved constraints on the models of the lower crust and upper mantle(Gangopadhyay et al., 2007). Preliminary results of applications to data from permanent broadband seismic stations in China and Canada are summarized below.

In China, based on our selection criteria, the number of earthquakes recorded at each station range between 3 and 31. The stations encompass tectonic provinces such as the north and south China blocks, Tibetan plateau, and Tien Shan Mountains. Results from the application of our technique are consistent with regional tectonics and models obtained from earlier studies using receiver functions and seismic tomography. The crustal thicknesses beneath stations in the north China block range between ~35–42 km with the crust in the central part of the block near station BJT being thicker. However, within the north China block towards its western edge, the crust appears anomalously thick (~55 km) beneath station LZH. This region also coincides with the border of the north China block and Tibetan plateau. The south China block consists of widely varying crustal thicknesses: ~37 km beneath station ENH in the northern part of the block and ~52 km beneath station KMI in the southern part, for example. The southern part of the south China block near station KMI also coincides with its border with the Tibetan plateau, where the
Moho significantly deeper than elsewhere in China, implying a deep crustal root. Station LSA, in Lhasa, Tibet, has a crustal thickness of ~53 km. In the Tien Shan Mountains in northwestern China, the crust is thicker than average (~42 km) but not as thick as that beneath the Tibetan plateau.

The seismic stations in Canada each recorded between 3 and 26 earthquakes suitable for analysis. The chosen stations span major tectonic provinces within the Canadian landmass such as the Cordilleran orogen, western plains and Slave province in western Canada, Grenville province and Appalachian orogen in eastern Canada, and the Canadian Arctic. Initial results from the application of our technique are consistent with earlier studies and regional tectonics. Crustal thicknesses beneath stations in the northern Cordilleran orogen, western plains, and Slave province range between 35-37 km. The Moho appears to be slightly shallower (31-35 km) beneath stations in the southern Cordilleran orogen. In eastern Canada, the crust beneath stations of the Grenville province and Appalachian orogen is generally thick (~44 km) with the exception of that beneath station DRLN, in the northeast, where it is ~33 km. Moho depths beneath stations in the Canadian Arctic range between ~35-42 km. We also observe low-velocity zones in the crust and uppermost mantle at some stations, however the constraints on them are not strong. Taken together, the results provide a comprehensive snapshot of the velocity structure beneath Canada.

These results demonstrate that the waveform modeling method is tractable and effective when the data are available. Unfortunately we discovered during our global search that SPL waves, in particular, are relatively uncommon. It makes sense, therefore, to include additional data functionals for joint modeling, in order to broaden the applicability of the enhanced method to more regions of the world and to increase sensitivities to Earth structure, and therefore model reliability.

### Estimation of Uncertainties

Geophysical inverse problems are often non-unique. That is, their error functions either have broad minima or are multi-valued, indicating that models that are slightly different from the best-fitting model satisfy the data nearly as well, in the first case, or that one or more very different models also satisfy the data, in the second case. It is therefore necessary to explore the model space and thus identify the range of models that fit the data, and perhaps to identify characteristics of the models that are required by the data, rather than which simply are allowed by the data.

VFSA conducts such a search efficiently, and the products of multiple such searches enable us to evaluate the uncertainty in a single, best-fitting solution. This evaluation is particularly necessary in seismic waveform modeling because more than one model can often explain the observed data equally well and trade-offs between different model parameters are common (Pulliam and Sen, 2005). The waveform inversion method we use in this study incorporates important statistical tools that allow the user to evaluate the uniqueness, and physical feasibility of the resulting model. The most useful of these tools in evaluating the results’ reliability are the PPD function, and the parameter correlation matrix. To estimate these statistical parameters we cast the inverse problem in a Bayesian framework (Sen and Stoffa, 1995; Tarantola, 1994), and employ “importance sampling” based on a Gibbs’ sampler (GS) (Pulliam and Sen, 2005; Sen and Stoffa, 1996).

The goal of “importance sampling” is to concentrate sample points in the regions that are the most “significant”, in some sense (perhaps, for example, where the error function is rapidly varying, or many acceptable solutions lie). Because this concentration is achieved using a Gibbs’ probability distribution, it has been named the “Gibbs’ sampler” (Sen and Stoffa, 1996). The PPD function \([\sigma(m|d_{obs})]\) is defined as a product of a likelihood function \([e^{-\frac{1}{2}(m)}]\), and prior probability density function, \(p(m)\). The prior probability density function \(p(m)\), describes the available information on the model without the knowledge of the data and defines the probability of the model \(m\) independent of the data. In our application here, we use a uniform prior within a minimum and maximum bound for each model parameter. The likelihood function defines the data misfit and its choice depends on the distribution of error in the data (Sen and Stoffa, 1996). Sen and Stoffa (1996) examined several different approaches to sampling models from the PPD and concluded that a multiple-VFSA based approach, though theoretically approximate, is the most efficient. In a multiple-VFSA approach we make several VFSA runs with different random starting models and use all the models sampled along to characterize uncertainty in the model. We use all these sampled models to compute approximate marginal PPD and posterior correlation matrices to characterize uncertainties in the derived results. The posterior correlation matrix measures the relative trade-off between individual model parameters and is computed by normalizing the covariance between two model parameters. Computationally, the correlation between \(i^{th}\) and \(j^{th}\) model parameters is given by their covariances divided by the square root of the product of the covariances of each parameter with itself. Next we discuss an application of the technique to seismological data recorded in Africa, including descriptions and interpretations of the resulting computations of the PPD and correlation matrix.
Figure 1. (Top panel) P- and S-wave velocity models up to 100 km for station TAM from the inversion results for individual events recorded at TAM. (Lower panels) Model parameter correlation matrices for events 1 (left panel), in which SPL was not observed, and 3 (right panel), for which SPL was observed, recorded at TAM. Each small square represents a model parameter (Vp, Vs, Thickness of layer, and Density) on both the horizontal and vertical axes. The correlations range between -1 and 1. Sparse colored squares off-diagonal in the lower crust - upper mantle in event 3 (right panel) compared to that in event 1 (left panel) indicate better resolution and confidence (less trade-off) in this region.
In a paper published in Geophysical Journal International (Gangopadhyay et al., 2007) we generated $P$- and $S$-wave velocity models up to a depth of 100 km for the broadband station TAM, in North Africa (Figure 1). We observed some variability in the models, and hence computed the uncertainties for each model using the statistical tools described earlier to choose the “best” model. In Figure 1 we show examples of parameter correlation matrices computed from the modeling results of two events (we call them events 1 and 3). Each small square along an axis of the parameter correlation matrix, either horizontally or vertically, represents a model parameter. Since every model layer consists of four independent model parameters ($V_p$, $V_s$, thickness, and density), four small squares combined together represent a model layer on both axes. Correlation values range between -1 and 1 and are symmetric about the diagonal of the matrix, hence, for clarity, we show only values below the diagonal. Values along the diagonal are ones, simply indicating that each parameter is perfectly correlated with itself. Off-diagonal colored squares indicate significant cross-correlation (trade-offs) between corresponding model parameters. In the parameter correlation matrices for both events, layers comprising the upper crust have greater independence, as indicated by the sparse distribution of off-diagonal cross-correlations whose absolute values are greater than ±0.5 (colored squares). Also, for both events, the level of tradeoffs among model parameters in these shallow layers is similar. For event 1, however, the layers comprising the lower crust and upper mantle have larger off-diagonal cross-correlations, indicating significant tradeoffs (Figure 1, lower left). On the contrary, for event 3, even the lower crustal and upper mantle layers appear better constrained (Figure 1, lower right). The main difference between the two events is that an SPL phase is also observed in the seismogram of event 3 but not in that of event 1. This observation attests to the fact that if SPL is present in the seismogram and is well modeled, we are able to better constrain the structure of the lower crust and upper mantle. This result confirms our expectation that, due to the sensitivity of SPL to those parts of the model, adding SPL to the modeling, when present, improves constraints on parameters at and just below the Moho.

**CONCLUSIONS AND RECOMMENDATIONS**

**Joint Modeling of Multiple Datasets**

There are several advantages to jointly modeling multiple datasets. First, each data functional has unique sensitivities to Earth structure. For example, receiver functions are primarily sensitive to shear wave velocity contrasts and vertical traveltimes while surface wave dispersion measurements are sensitive to vertical shear wave velocity averages (Julia et al., 2000). Full waveform modeling of $S$, $Sp$, $SsPmP$ phases, in contrast, are more sensitive compressional wave velocity contrasts and vertical traveltimes; adding SPL to the modeling improves sensitivity to the uppermost mantle shear wave structure and to velocity contrasts across the Moho (Gangopadhyay et al., 2007). Mutually satisfying constraints imposed by the three datasets may constrain a larger subset of model parameters than a single set alone, resulting in a single model that better represents the true structure beneath a station.

Second, the amplitudes and signal-to-noise characteristics of waves that produce these data functionals depend on several factors, including epicentral distance, event focal depth, fault mechanism and radiation patterns, source time function, properties of the intervening Earth structure (including attenuation, low velocity zones, velocities, heterogeneity, and anisotropy), and characteristics of the recording seismometer. Since regions of high seismicity are highly restricted, many stations are not well situated to record a large number of ideal events. Or a station may be deployed only temporarily, resulting in inadequate sampling to produce some functionals. An algorithm that incorporates more than one type of data will most likely be more widely applicable than one that relies solely on a single type.

However, with these advantages come some disadvantages. For example, combining disparate data types requires great care in their treatment and assessment (Roy et al., 2005). Benefits of additional data may be null if the method used to model them preferentially fits one type. Or, worse, minimizing an inappropriate criterion in conjunction with incompatible data may “split the difference” between them to choose a model that is wholly inaccurate and inappropriate for its intended purpose. In geophysical modeling, one does not often have the luxury of choosing a model based solely upon its numerical characteristics. If the model is to be used for regional or local earthquake locations, for example, it would be a mistake to rely on the best fit to surface wave dispersion. The judgment and experience of seismologists who keep a clear eye on their goal is critical, and this experience must be combined with rigor in the computational modeling. This experience and judgment can be incorporated after the fact, as is sometimes the case with, for example, the smoothing applied to 3D tomographic models but it is usually better to
acknowledge the “prior” explicitly at the outset. This is one advantage of the Bayesian approach that we propose here. While such priors are sometimes referred to pejoratively as “bias”, the explicit statement of a prior during the formulation of the modeling algorithm forestalls serious criticism and enables a clear discussion of the “bias” that was imposed on the modeling.

Next, as a general rule of thumb, a greater number of data functionals incorporated into modeling will result in a broader range of model parameter sensitivities, and it will be less likely that a linear inversion approach will be adequate. This is unfortunate because linear approaches are much more tractable and straightforward than nonlinear methods. But only a broad search of the model space will demonstrate whether a linear approach is valid. Non-linear global optimization algorithms require no change in the algorithm to include multi-part objective function with different norms. A variety of methods for nonlinear inversion are now available and the only real cost for conducting a thorough search and finding the single best-fitting model is in computational effort and complexity. The downside is that theoretically exact methods for assessing model uncertainties and model reliability are not generally tractable, and approximate methods result in significantly greater cost and complexity than is required to find the best-fitting model alone. Nevertheless, our previous work has demonstrated that useful methods are indeed tractable, and the method we propose will add only marginal increases in computation time. Fortunately, with two of the world’s largest parallel computers at our disposal at UT, we have the tools we need to undertake these tasks.

While surface wave dispersion and receiver functions have been modeled jointly (Ammon et al., 2005; Ammon et al., 2004; Cakir and Erduran, 2004; Chang et al., 2004; Dugda and Nyblade, 2006; Herrmann et al., 2001; Julia et al., 2000; Julia et al., 2005; Lawrenece and Wiens, 2004; O zalaybey et al., 1997; Tkalcic et al., 2006), no study has, to our knowledge, incorporated waveform constraints such as S, Sp, SsPmP, and Shear-coupled PL. Further, none has conducted a thorough, nonlinear assessment of the constraints provided by each functional.

Application: Modeling the crust and uppermost mantle beneath the Middle East

The Middle East has a complex tectonic history. This history has given rise to a complicated earth structure, with highly non-uniform spatial distributions of fundamental crustal properties, such as depth-to-basement and crustal thickness. Similarly large variations in upper-mantle velocity occur. These factors strongly affect the group velocities of surface waves propagating across the region, as well as discontinuities that convert P and S body wave energy.

Table 1. Sensitivities of the three seismic functionals under discussion. While this table summarizes what is known about sensitivities from previous studies, the degree to which tradeoffs between model parameters will be reduced by the joint modeling we propose is unknown. The degree of tradeoff and resolution will be quantitatively assessed during the course of this study.

<table>
<thead>
<tr>
<th>Region Constrained</th>
<th>Parameter Constrained</th>
<th>Receiver Functions</th>
<th>Surface Wave Dispersion</th>
<th>Waveform Modeling of S, Sp, SsPmP, SPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertical velocity average</td>
<td>Absolute Vs</td>
<td>High</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Absolute Vp</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Internal boundaries in the crust</td>
<td>Vp Contrast</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Vs Contrast</td>
<td>High</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Moho</td>
<td>Vp Contrast</td>
<td>Moderate</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Vs Contrast</td>
<td>Moderate</td>
<td>Low-to-Moderate</td>
<td>High</td>
</tr>
<tr>
<td>Crust</td>
<td>Absolute Vp</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Absolute Vs</td>
<td>Moderate</td>
<td>Low-to-Moderate (depending on frequencies used)</td>
<td>Moderate</td>
</tr>
<tr>
<td>Uppermost Mantle</td>
<td>Absolute Vp</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Absolute Vs</td>
<td>Moderate</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Vertical Resolution</td>
<td>Absolute Vs</td>
<td>High</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Lateral Resolution</td>
<td>Absolute Vs</td>
<td>Low</td>
<td>High</td>
<td>Moderate</td>
</tr>
</tbody>
</table>
Due to its complex tectonics, the Middle East is an appropriate location in which to apply methods that (a) reduce non-uniqueness by fitting multiple datasets and (b) assess the relative constraints placed on model parameters by each data type. These tools can help reconcile seemingly inconsistent datasets, such as the surface wave dispersion data (and models) proposed for the Zagros Mountains region (Pasyanos and Nyblade, 2007) and receiver function results for the same region (Ammon et al., 2005; Paul et al., 2006), although produced with data from different stations. What we need now are better constraints on $P$ velocities and an algorithm that incorporates priors and data functionals in a step-wise manner, with assessment tools to evaluate the results. Lastly, the Middle East has the densest coverage and highest resolution in the surface wave model and is an area of interest for explosion monitoring.

REFERENCES


