REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
30 Sep 2009	REPRINT	

4. TITLE AND SUBTITLE
ESTIMATING THE UNCERTAINTY AND PREDICTIVE CAPABILITIES OF THREE-DIMENSIONALEARTH MODELS

5a. CONTRACT NUMBER
FA8718-09-C-0013

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62601F

6. AUTHOR(S)
Delaine T. Reiter[1], William L. Rodi[2], and Stephen C. Myers[3]

5d. PROJECT NUMBER
1010

5e. TASK NUMBER
SM

5f. WORK UNIT NUMBER
A1

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Weston Geophysical Corporation
181 Bedford St., Suite 1
Lexington, MA 02420

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
29 Randolph Road
Hanscom AFB, MA 01731-3010

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RVBYE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-RV-HA-TR-2009-1087

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited.

Weston Geophysical Corporation[1], Massachusetts Institute of Technology[2], andLawrence Livermore National Laboratory[3]

13. SUPPLEMENTARY NOTES
Reprinted from: Proceedings of the 2009 Monitoring Research Review – Ground-Based Nuclear Explosion Monitoring Technologies, 21 – 23 September 2009, Tucson, AZ, Volume I pp 184 - 193.

14. ABSTRACT

Many three-dimensional models of seismic velocity structure in Eurasia have been developed in recent years by the seismic nuclear monitoring community. Most of these models are not accompanied hy quantitative estimates of uncertainty, either in the model velocities themselves or in geophysical observables predicted hy the models (e.g., body-wave travel times). Moreover, the various 3D models produced by these studies have not been compared to one another for their predictive capabilities in any meaningful way. We have recently hegun a new effort to address these issues, which will culminate in a comprehensive evaluation of the current generation of 3D seismic velocity models.

In this paper we show the results of applying two familiar validation techniques, or model evaluation metrics, to three seismic velocity models. The evaluation metrics are regional travel-time prediction and event relocation, each using a ground-truth (GT) dataset that includes events with epicenters known to 7 km or hetter and regional *P* and *S* arrivals within the footprint of the model region. The models include the Joint Weston/MIT (**JWM**) crust and upper-mantle velocity model for south-central Asia, which was derived hy jointly inverting a large set of body-wave travel times and surface-wave group velocities in a coupled nonlinear procedure. We also derived models from the body-wave and surface-wave datasets separately, using the same initial model, inversion grids, constraints and regularization employed in the joint inversion. To make comparisons with the **JWM** model possible, we applied the Poisson's ratio of the initial model to convert the *P* velocity model constructed with travel times to an *S* velocity model (and vice versa). The results of these exercises reveal many factors that complicate the straightforward evaluation of the models.

15. SUBJECT TERMS
Velocity model uncertainty, Travel time prediction, Relocation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Robert J. Raistrick
UNCLAS	UNCLAS	UNCLAS	SAR	10	19b. TELEPHONE NUMBER *(include area code)*

ESTIMATING THE UNCERTAINTY AND PREDICTIVE CAPABILITIES OF THREE-DIMENSIONAL EARTH MODELS

Delaine T. Reiter[1], William L. Rodi[2], and Stephen C. Myers[3]

Weston Geophysical Corporation[1], Massachusetts Institute of Technology[2], and
Lawrence Livermore National Laboratory[3]

Sponsored by the Air Force Research Laboratory and the National Nuclear Security Administration

Award Nos. FA8718-09-C-0013[1,2] and DE-AC52-07NA27344[3]
Proposal No. BAA09-73

ABSTRACT

Many three-dimensional models of seismic velocity structure in Eurasia have been developed in recent years by the seismic nuclear monitoring community. Most of these models are not accompanied by quantitative estimates of uncertainty, either in the model velocities themselves or in geophysical observables predicted by the models (e.g., body-wave travel times). Moreover, the various 3D models produced by these studies have not been compared to one another for their predictive capabilities in any meaningful way. We have recently begun a new effort to address these issues, which will culminate in a comprehensive evaluation of the current generation of 3D seismic velocity models.

In this paper we show the results of applying two familiar validation techniques, or model evaluation metrics, to three seismic velocity models. The evaluation metrics are regional travel-time prediction and event relocation, each using a ground-truth (GT) dataset that includes events with epicenters known to 7 km or better and regional *P* and *S* arrivals within the footprint of the model region. The models include the Joint Weston/MIT (**JWM**) crust and upper-mantle velocity model for south-central Asia, which was derived by jointly inverting a large set of body-wave travel times and surface-wave group velocities in a coupled nonlinear procedure. We also derived models from the body-wave and surface-wave datasets separately, using the same initial model, inversion grids, constraints and regularization employed in the joint inversion. To make comparisons with the **JWM** model possible, we applied the Poisson's ratio of the initial model to convert the *P* velocity model constructed with travel times to an *S* velocity model (and vice versa). The results of these exercises reveal many factors that complicate the straightforward evaluation of the models.

OBJECTIVES

The main objective of our research project is to develop and apply meaningful measures of the predictive capabilities of 3D Earth models. We will specifically focus on seismic velocity models, although our general approach should be applicable to models of other geophysical parameters such as attenuation and density.

Our project will consist of two major elements. First, we will perform a comprehensive and methodical evaluation of a set of regional 3D velocity models for Asia based on data misfit and event mislocation metrics which we and other investigators have used in previous work. These metrics are already generally accepted as informative, if not complete, measures of the prediction capability of Earth models. The models will include our joint inversion model **JWM**, which will serve as the main test bed for the project. Metrics will be evaluated for each model based on a common ground-truth dataset comprising GT5 events and their arrival picks. A large part of this effort will be to establish rigorous validation tests employing standardized methods for model parameterization, forward modeling (e.g., ray tracing), and event relocation. The tools we need for these tests are available to us in-house from our previous tomography and location projects.

The second major element of our project will be the investigation of a new approach for estimating the predictive capability of 3D velocity models. Our approach will convert the uncertainty in the parameters of a 3D velocity model, as a function of position in the Earth, into the uncertainty in travel-time or other geophysical predictions. To do this requires covariance modeling capabilities we have developed for travel times in an earlier project (Rodi and Myers, 2007) and which can be extended to other observables given appropriate model sensitivities.

In this paper we show the results of applying two familiar validation techniques, or model evaluation metrics, to three seismic velocity models. The evaluation metrics are regional travel-time prediction and event relocation, each using a GT dataset that includes events with epicenters known to 7 km or better and regional *P* and *S* arrivals within the footprint of the model region. The results show that there are many factors that complicate the straightforward evaluation of the models.

RESEARCH ACCOMPLISHED

In previously funded work we applied a nonlinear joint inversion of body-wave travel times and surface-wave group velocities to data from a broad region in south-central Asia (Reiter and Rodi, 2009). The forward modeling incorporated in our inversion utilizes fully 3D ray tracing for the body-wave travel times, and a two-step procedure for the surface-wave dispersion data that includes 1D dispersion modeling at a geographic point followed by 2D ray tracing in the phase velocity maps. We numerically solved the inverse problem using a set of iterated inversion steps. Consistency between the *P* and *S* velocities was achieved by imposing bounds on Poisson's ratio and by invoking a regularization constraint that correlates variations in *P* and *S* velocity from an initial, or prior, model. We constructed the prior model as a composite of the CRUST2.0 model (Bassin et al., 2000) in the crust and the 1D AK135 reference model (Kennett et al., 1995) in the mantle. The resulting inversion model, dubbed **JWM** (for **J**oint **W**eston/**M**IT), shows good agreement with the persistent features seen in previous seismic velocity models produced from separate body- or surface-wave datasets, as well as some intriguing differences between the compressional and shear-wave structure.

Our primary goal in developing the new inversion approach and the **JWM** model was to improve regional seismic event location capability in a strongly heterogeneous crust and upper mantle. In our new work we will develop and apply a robust set of location metrics to compare the behavior of seismic velocity models. For this paper we have applied two standard metrics to our **JWM** model and two additional 'in-house' seismic velocity models: the direct comparison of predicted *Pn* and *Sn* travel times to the ground-truth observations; and relocation of the ground-truth events using regional arrivals. In the next sections we discuss the velocity models and ground-truth data used in our evaluation, as well as the results of applying the two metrics.

3D Seismic Velocity Models

JWM was derived by applying our joint inversion procedure to the combined *P*-wave travel-time and Rayleigh-wave group-velocity datasets. In contrast, the other two 3D models considered in this paper were derived from inverting each dataset separately, using the same initial model, grid parameters, constraints and regularization

employed in the joint inversion. To make comparisons with the **JWM** model, we used a simple scaling relationship to convert the final *P*-velocity model constructed with travel times to an *S*-velocity model. Conversely, the *S*-velocity model constructed from group velocities was converted to a *P*-velocity model. The scaled velocities were computed using the prior model's velocity ratio. For the separate surface-wave inversion the scaling relationship was employed at each iteration of our nonlinear inversion procedure, since group velocities depend on both *P* and *S* velocity. We also note that the crustal thickness was held to the values in the CRUST2.0 model in the body-wave only inversion, but was allowed to vary in the surface-wave only inversion.

Each inversion for the joint (**JWM**), body-wave only (**BWo**), and surface-wave only (**SWo**) models comprised ten iterations. The first five of these iterations were completely nonlinear, including the calculation of updated raypaths in the current *P* velocity model and group-velocity dispersion maps. Raypaths were held constant for the last five iterations, although the nonlinear dependence of the group-velocity maps on *S* velocity and crustal thickness was still accounted for. We chose the fifth iteration to stop calculating new ray-path sensitivities based on the insignificant changes that occurred in the ray paths between the 4th and 5th iterations.

We summarize the fit to the inversion datasets for the three models in Table 1, which shows the root-mean-square (RMS) residual from the observed the travel times and group delays achieved by each model. For comparison purposes we also included the fits to the initial and AK135 models. Table 1 shows that the **JWM** inversion model lowers the variance for both datasets with respect to the initial and AK135 models, which both illustrates the success of the inversion method and reveals how well we constructed our initial model.

With respect to the other 3D models, **JWM** fits the travel times nearly identically to the **BWo** model and significantly better than the SWo model. The fit to the group delays presents a more complicated story. The **JWM** model fits the group delays better than the **BWo** model, but not as well as the SWo model. In the joint inversion the influence of the *P* model (derived with travel times) has an effect on the predicted group-velocity curves that are inverted for shear velocities, even though the sensitivity of the group velocities is predominantly to the shear velocity. The SWo inversion used simple scaling to produce a *P* model at each inversion step, which resulted in the best fit to the group-velocity data, but a very poor fit to the *P* travel times (worse than the AK135 and initial models).

Table 1. RMS error of the datasets with respect to the AK135, initial and final inversion models; various subsets of the data are also shown.

Data Set		*RMS Error (s)*				
		AK135	*Initial Model*	*JWM*	*BW Only Model*	*SW Only Model*
P Travel Times	all distances	2.50	2.87	2.02	2.03	2.91
	Δ = 0 - 12°	2.27	2.43	1.86	1.87	2.75
	Δ = 12 - ~18.6°	2.81	3.42	2.24	2.26	3.13
Group Delays	all periods (T)	60.0	35.2	30.2	33.8	24.9
	T = 10 - 20 s	46.9	37.5	28.2	33.2	27.3
	T = 25 - 45 s	82.2	37.6	36.2	37.0	24.9
	T = 50 - 80 s	45.7	28.8	24.1	31.1	20.4
	T = 90 - 150 s	28.3	27.8	24.7	28.9	24.3

We applied our evaluation metrics to these three 3D inversion models to determine the merits of each model in relation to both the AK135 reference model as well as to one another. This represents an ideal comparison case in many ways, since the models were derived using identical datasets and methodologies.

Validation Data: Ground-Truth Explosions and Earthquakes

To show that 3D models improve regional travel-time prediction and event locations, we require a high-quality set of GT events whose locations, including depth and origin time, are known precisely. We used a subset of the reference events developed by the Group-2 Location Calibration Consortium (Bondár et al., 2004a). This valuable database includes nearly 2000 explosions and shallow earthquakes across the Middle East, North Africa, Europe, and Western Eurasia whose epicentral location accuracy is 10 km or less. The Group-2 Reference Event List (REL)

was used effectively by Yang et al. (2004) to demonstrate the validity of a set of regional and teleseismic travel-time models. We filtered the Group-2 REL to include events with epicentral ground-truth accuracies between 0 (GT0) and 7 (GT7) km in our study region (10-50° N, 40-110° E). We supplemented this list with some additional GT events from the EHB event bulletin (Engdahl et al., 1998) and a small list developed for an IASPEI location workshop (Engdahl, 2006). Within the boundaries of our study region there were 404 GT0–GT7 earthquakes and explosions, 319 of which fall into named event clusters. We note that we segregated the GT event database from the data used to develop the inversion models, so that we could use the events and arrivals in independent validation testing. This did not significantly impact the inversion dataset because the EHB bulletin provided a large amount of travel-time data from other well-located events. Figure 1 shows the distribution of the GT events, and Table 2 provides some details about the numbered events or clusters shown in the figure, including the numbers of events and phase arrivals that were used in our model evaluation tests.

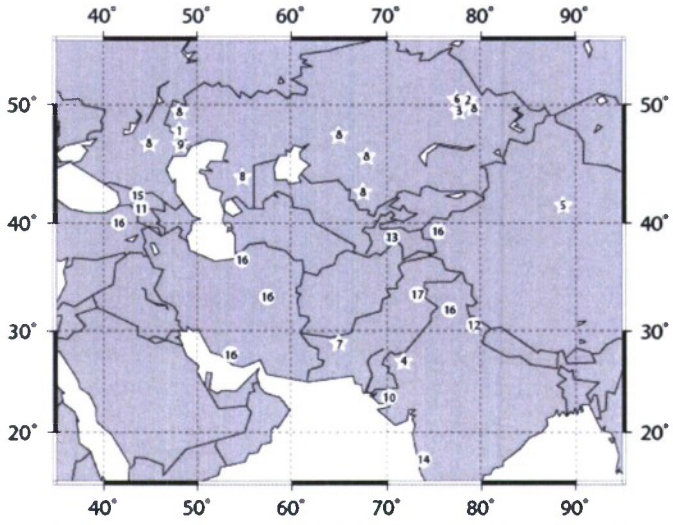

Figure 1. Locations of ground-truth epicenters (GT0-GT7) in the study region. Stars refer to explosions and circles to earthquakes; numbers are referenced to specific events or event clusters listed in Table 1.

Table 2. Details of the ground-truth database used in the model evaluation exercises.

Map Number	Event Type	N_{total} †	$N_{relocate}$ †	$N_{relocate\ S>10\%}$ †	GT Level (km)	P_{ndef}	S_{ndef}
		Nuclear Explosions					
1	PNE: Azgir	8	5	0	1	75	0
2	Balapan (Semipalatinsk)	56	7	7	1	342	81
3	Degelen (Semipalatinsk)	68	3	2	0-1	235	9
4	India	2	2	1	0-1	43	6
5	Lop Nor, China	19	15	7	1-5	395	52
6	Murzik (Semipalatinsk)	7	0	0	0-5	20	0
7	Pakistan	1	1	1	5	27	3
8	PNE: Miscellaneous	7	2	0	0-1	45	2
9	PNE: Vega	13	0	0	1	56	1
	Sum =	181	35	18		1238	154
		Earthquakes					
10	Bhuj, India	76	72	70	5-7	2370	812
11	Caucasus	6	4	4	5	160	83
12	Chamoli, India	58	51	42	5-7	1775	480
13	Garm, Tajikistan	27	27	27	5-7	843	288
14	Koyna, India	14	13	13	5	271	100
15	Racha, Georgia	35	33	28	5-7	891	243
16	Engdahl, 2006	6	6	5	5	239	60
17	Valentine's Day 1977, Pakistan	1	1	1	5	32	10
	Sum =	223	207	190		6581	2076

† N_{total} = total number of events in each geographic area; $N_{relocate}$ = events used in relocation exercise; $N_{relocate\ S>10\%}$ = events with S arrivals making up > 10% of their arrival set

The Group-2 REL does not contain a large number of regional readings, so we cross-referenced our event list to the EHB bulletin to retrieve a larger set of defining regional *P* and *S* observations. We include the *S* arrivals in the validation database because we would like to know whether the new 3D models can predict these observations because of the potentially high value *S* arrivals have for locating small events with sparse networks. However, it is important to bear in mind that *S* travel times are more likely to suffer from picking errors, and we excluded *S* travel times from the inversions.

We groomed the resulting set of arrivals by eliminating observations with residuals relative to AK135 greater than ±8.0 seconds or with clearly misidentified phase arrivals in crossover distances between various travel-time branches. With these criteria we derived a validation dataset of 7,819 regional *P*-wave arrivals and 2,230 regional *S*-wave arrivals observed at stations within a latitude-longitude box defined by (0-60° N, 30-120° E). The great-circle raypath coverage for these *P* and *S* GT observations is shown in Figure 2. The region of the inversion model that is sampled by the GT data raypaths is limited, which illustrates the difficulty of validating a model with travel times alone. However, the Group-2 REL database is currently the best available resource we have to definitively demonstrate the validity of an inversion model.

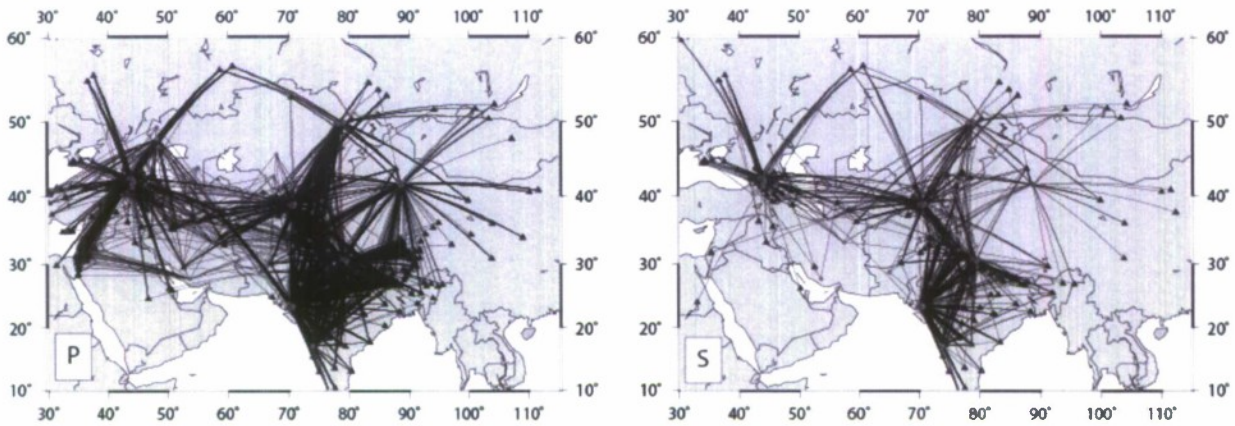

Figure 2. Great-circle raypath coverage for the *P* (left) and *S* (right) regional observations in the ground-truth database used to evaluate the inversion models. Stations are represented by black triangles and events by gray circles.

Model Evaluation Metric 1: Fits to Regional Travel Times

We follow other researchers in assessing our model for improvement in regional travel-time prediction and event location accuracy (e.f. Ritzwoller et al., 2003, Flanagan et al., 2007). The first assessment of the 3D models involves measuring their ability to accurately model the regional travel-time data. As noted by other authors, improved travel-time prediction is the most direct way to estimate the quality of a 3D model, because it eliminates the effects of network geometry and variable pick quality.

One of the more difficult questions we will address in our new project is how to present the model fits in a coherent and meaningful fashion. There are many ways to examine residuals; for example, on a station-by-station or event cluster basis (e.g. Flanagan et al., 2007), or through spatial patterns as a function of distance and station (e.g., Murphy et al. 2005), or by the fits to empirical phase path anomalies (e.g., Ritzwoller et al., 2003). Figure 3 represents a rudimentary attempt at presenting travel-time fits to the regional *P* data in the GT database relative to the AK135 and 3D inversion models. For this figure we binned the data as a function of event-to-station distance and the absolute value of the residual. By defining a visually distinct color scale and the bin sizes (here set to 0.25° in distance and 0.5 seconds in residual size), we can examine a distance-based measure of the fit; however, we remove any sense of the spatial/geographic variation or dependence of the model fits. Nevertheless, we can make some general observations about the models from these plots. The *P* residuals with respect to the AK135 model (Figure 3a) show a distinct pattern that is due to the tectonics of the study region. As noted in Ritzwoller et al. (2003), the increase in spread at the longer regional distances is likely due to some ray paths bottoming well below the bottom of the crust and encountering low velocity zones. The **JWM** and **BWo** models produce nearly identical

plots (Figure 3b,c), although detailed analysis of the differences between the two plots shows that **JWM** performs slightly better than **BWo** at most distances. The **SWo** model does as well as the other 3D inversion models at the shorter distanees, but fits less well at distances greater than approximately 8°.

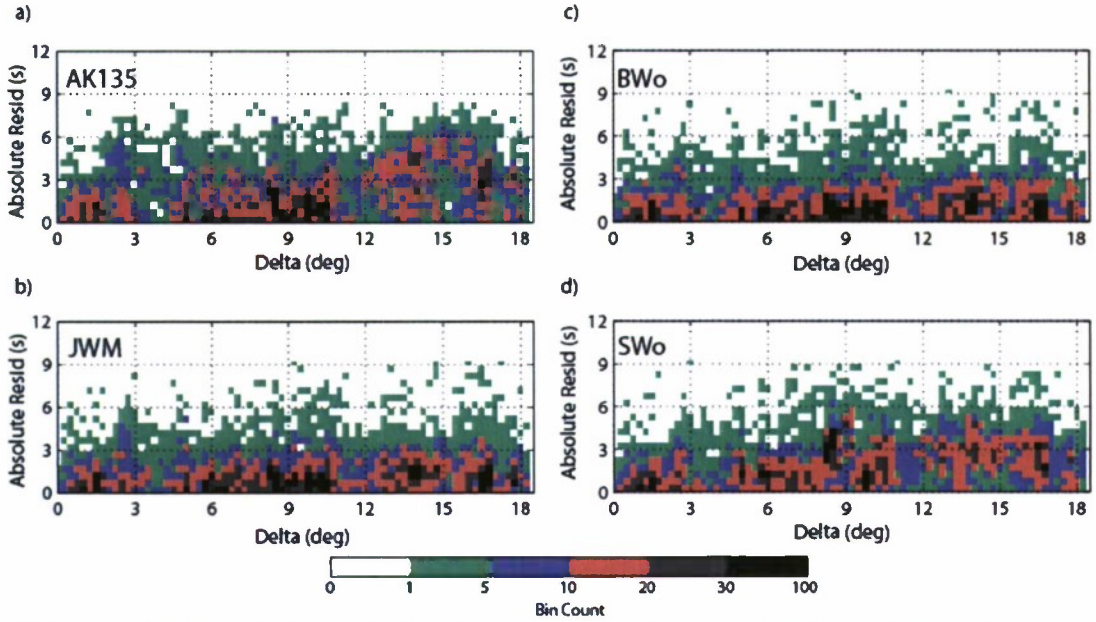

Figure 3. Ground-truth residuals for the *P* arrivals in our GT database, binned according to station-to-event distance and residual (intervals of 0.25° in distance and 0.5 seconds in residual size). The bin hit counts are plotted with the nonlinear color scale at the bottom of the plot. The residuals are calculated as the absolute value of the observations minus a) the AK135 modcl; b) thc JWM model; c) the BWo model; and d) the SWo model.

As mentioned earlier, we are also interested in the performance of the 3D models with respect to regional *S* arrivals. *S* arrival data represent a crucial source of information in the regional recordings of a moderate-sized event, and we believe that inereasing the attention paid to them for location purposes is worthwhile. Figure 4 shows the same information as Figure 3, but summarizes the travel-time predietion results from the significantly smaller and noisier GT *S* database. In this case we used nearly the same color scale, and the bin sizes were enlarged to 0.5° in distance and 1.0 seconds in residual size to eliminate poorly populated bins. It is clear from Figure 4 that the *S* GT observations exhibit much higher scatter and have no distinctive pattern in their variation compared to the *P* observations. There are significantly fewer *S* observations in the GT bulletin, particularly for the explosion events, which makes statistical conclusions more difficult. However, there are several features worth noting.

All of the models fit the data fairly well at distances less than about 5°, mainly because the paths are too short to generate large travel-time differenees, no matter which velocity model is involved. It is also clear that picking error still dominates the *S* travel-time residuals, especially at the crossover distances. The residuals with respect to the AK135 model have an artificial eutoff of 8 seconds imposed on them (Figure 4a) that is not respected with the other models. In our study region, where crossover distances vary considerably, misidentified *S* phases that belong on later travel-time branehes appear regularly, depending on the locations of low-velocity zones in the upper mantle of each respective *S* model. However, in spite of all these difficulties, the **JWM** model produces a slightly better fit to the *S* travel-time residuals, judging by the high bin counts at lower residuals across the entire distance range. Detailed analysis (not shown here) shows that **JWM** has a slight edge over **BWo** and does even better in comparison to **SWo**, in terms of having more hits at lower residual values.

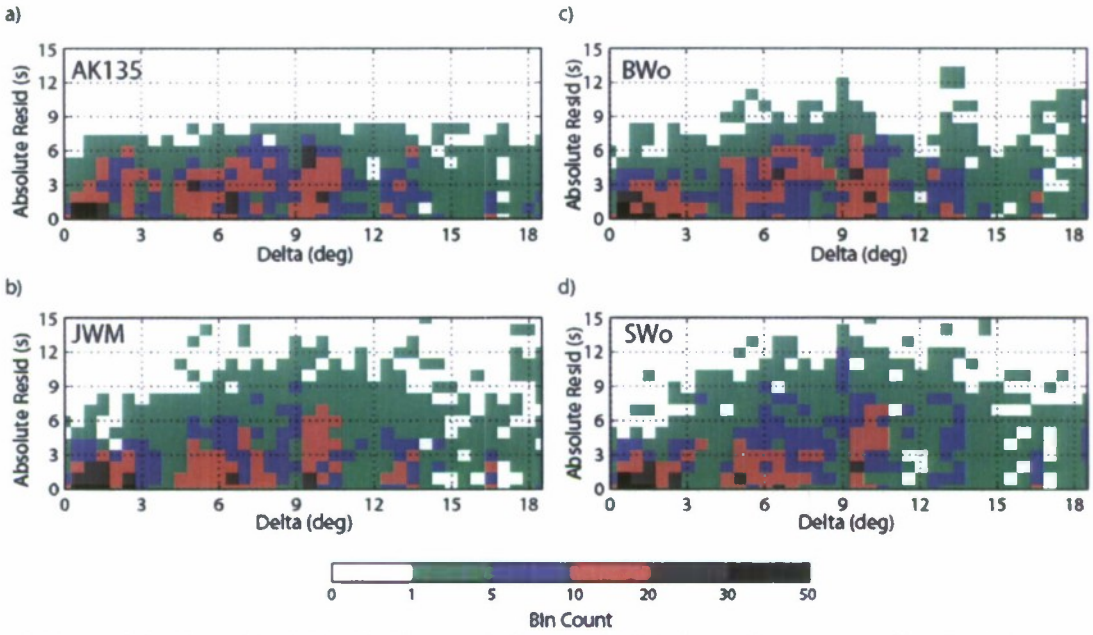

Figure 4. Ground-truth residuals for the *S* arrivals in our GT database, binned according to station-to-event distance and residual (intervals of 0.5° in distance and 1.0 seconds in residual size). The bin hit counts are plotted using the nonlinear color scale at the bottom of the plot. The residuals are calculated as the absolute value of the observations minus a) the AK135 model; b) the JWM model; c) the BWo model; and d) the SWo model.

Model Evaluation Metric 2: Locations Using Regional Travel Times

In the second assessment of the 3D inversion models we used a grid-search location method to estimate the epicenters of a subset of the events in the GT database using the arrivals at stations within the footprint of our models. In this exercise we fix the depth to the value in the GT database, because it trades off with origin time and has little effect on the epicenter. We filtered the GT data to include only those events whose regional station distribution within our model has a secondary azimuthal gap less than 270°; we included events with high secondary azimuth gaps to look at the abilities of the models when the network geometry is very poor. We also removed stations that were within 2.5° of the event and retained events with greater than 5 regional *P* arrivals. Table 2 shows the numbers of events that were relocated in each geographic location in the column labeled $N_{relocate}$.

This filtering reduced the testing dataset to a list of 35 explosions and 207 earthquakes within our region, with 7,041 regional *P*-wave arrivals and 1,826 regional *S*-wave arrivals. We used the Grid-search Multiple-Event Location (GMEL) algorithm (Rodi, 2006) to relocate the events in single-event mode, setting the arrival-time errors to 1.0 s for *P* and 2.0 s for *S* observations.

Figure 5 compares the mislocations for the AK135, **BWo** and **SWo** models versus **JWM** for the case when we locate using regional *P* arrivals. Events to the left of the unity line in these subplots are located better by the **JWM** model; we included overall **JWM** win percentages for explosions (red stars) and earthquakes (blue squares) in the bottom right of the plots. The results in Figure 5 indicate that **JWM** produces smaller mislocations than AK135 and the **SWo** model, which is no surprise; however, **JWM** also produces slightly better locations than the **BWo** model, even though the **BWo** model was generated from *P* travel times. One reason for this is that the **JWM** model locates the Bhuj cluster (No. 10 in Table 2) and the Lop Nor explosions (No. 5 in Table 2) better than the **BWo** model; these clusters account for a large component of the earthquake and explosion events in the exercise.

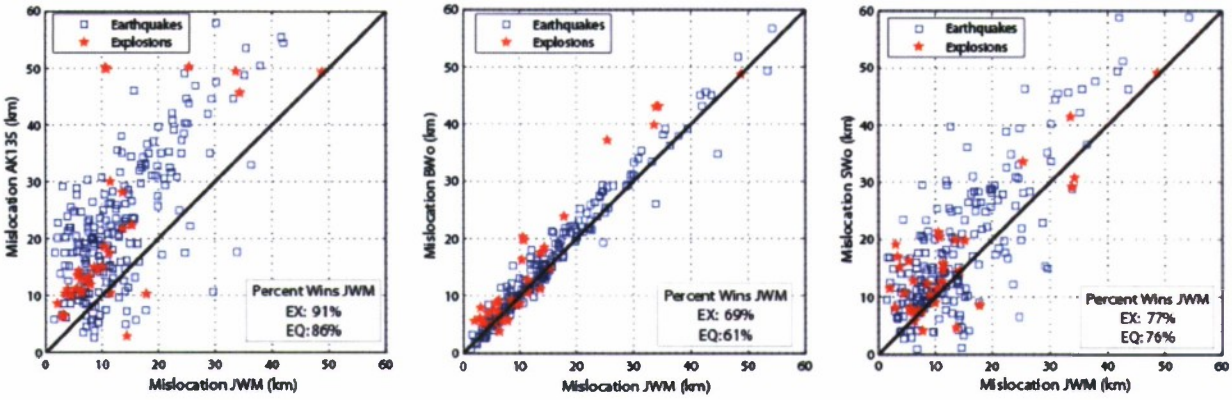

Figure 5. Epicenter mislocation comparisons for the JWM model versus the AK135 model (left), the BWo model (middle) and the SWo model (right). The locations were calculated using only *P* arrivals for the GT events in column $N_{relocate}$ in Table 2. Events to the left of the unity line indicate that JWM located the event better than the other model.

Figure 6 shows the results from locating with the *P* predictions from the 3D inversion models in a different manner, by displaying the difference in epicenter mislocation between **JWM** and the other models as a function of the secondary azimuth gap. All events in the gray-shaded areas below the zero line indicate smaller mislocations when **JWM** predictions were used. These plots demonstrate that as the network geometry degrades, **JWM** performs better than the other models. We note that this does not mean that the absolute mislocation is small for **JWM**, just that it is smaller than the mislocation produced by the **BWo** and **SWo** models.

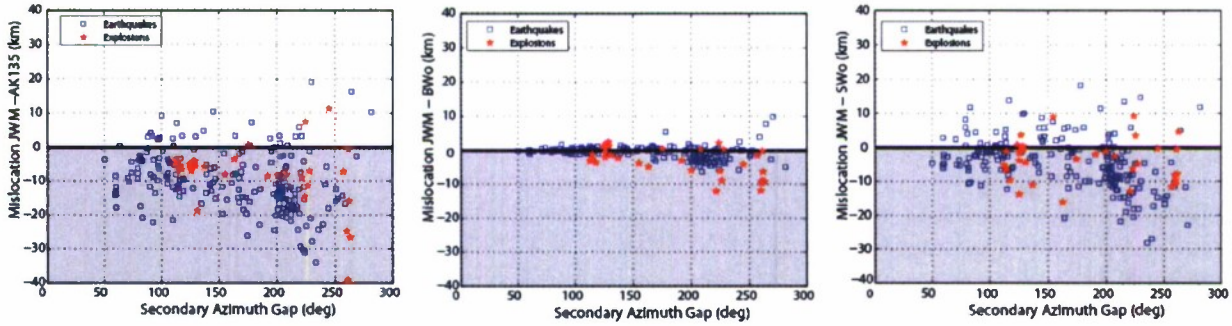

Figure 6. The differences between epicenter mislocations found with the JWM model and AK135 (left), BWo (middle) and SWo (right), as a function of secondary azimuth gap. Events in the gray-shaded regions below the zero line indicate that JWM located the event better than the other model. As the azimuth gap increases, JWM decreases event mislocations with respect to the other models.

We performed another relocation exercise to examine the influence of the *S* data and predictions from the 3D inversion models. In this case we further filtered those events designated $N_{relocate}$ to extract those for which the total regional bulletin had at least 10% *S* arrivals. This extra criterion significantly decreased the number of explosions, since nuclear tests typically do not have many *S* arrivals in the bulletin. This relocation exercise comprised 18 explosions and 190 earthquakes, listed as the column labeled $N_{relocate\ S>10\%}$ in Table 2, with 5,225 regional *P*-wave arrivals and 1,573 regional *S*-wave arrivals.

Figure 7 shows the results of relocating the $N_{relocate\ S>10\%}$ events, displaying them as the percentage improvement found when both *P* and *S* arrivals are included in the relocation compared to using the *P* arrivals alone, as a function of secondary azimuth gap. Events that are in the gray-shaded region below the zero line are the events for which the combined *P* and *S* arrivals resulted in an improved location. The total wins for the combined *P* and *S* relocation cases are shown plot insets. The results for **JWM** are shown on the left, **BWo** in the middle, and **SWo** on the right. The results show that the **JWM** model predicts the *S* data well enough to benefit the *P* and *S* locations over the

P-only locations. Adding the **BWo** *S* predictions to the locations produces better locations for the explosions but not for the earthquakes. Conversely, the **SWo** model predicts the *S* data well for the earthquakes, but not as well for the explosions. There is no strong correlation in percent improvement with secondary azimuth gap. However, for the cases in which the *P* and *S* combined bulletins produce smaller mislocations, **JWM** does a significantly better job of improving the results over **BWo** and **SWo**. In other words, when using both *P* and *S*, **JWM** has a stronger beneficial effect compared to the other two inversion models.

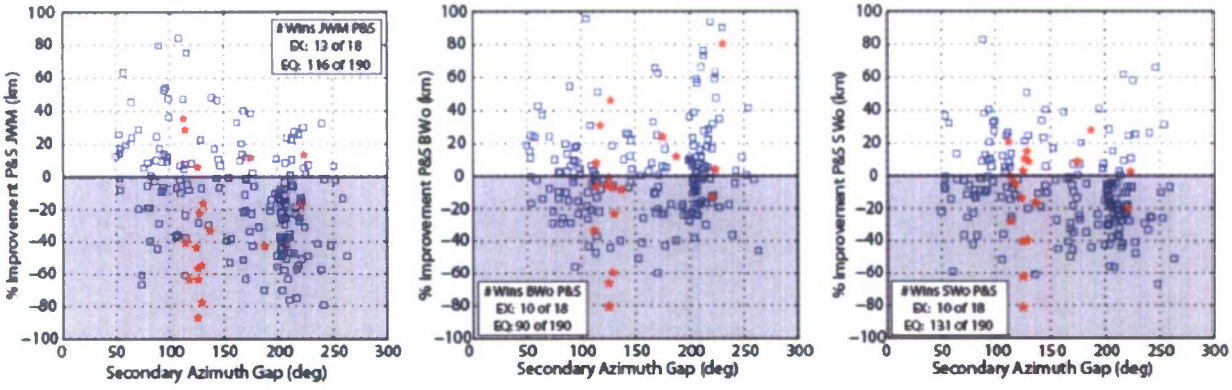

Figure 7. Epicenter mislocation comparisons when all *P* & *S* arrivals are included in the locations versus those with *P* arrivals only; JWM (left), BWo (middle) and SWo (right). The relocated events are listed in column $N_{relocate\ S>10\%}$ in Table 2. Results plotted as a function of secondary azimuth gap; inset boxes show the number of wins with *P* & *S* arrivals for explosions and earthquakes.

CONCLUSIONS AND RECOMMENDATIONS

The main objective of our new research project is to develop and apply meaningful measures of the predictive capabilities of 3D Earth models. As part of this work, we will perform a comprehensive and methodical evaluation of a set of regional velocity models for Asia based on data misfit and event mislocation metrics. These metrics are already generally accepted as informative, if not complete, measures of the prediction capability of Earth models. In this paper we applied two of these metrics to three inversion models that we derived from a 3D nonlinear inversion approach and a large database of *P*-wave travel times and Rayleigh-wave group velocities.

The primary motivation for computing the three different models, even though we plan to publish only **JWM**, was to explore the effects of a joint inversion approach under highly controlled conditions. When we hold out either the body-wave or surface-wave data from the inversion and use simple scaling relations to convert *P* to *S*, or *S* to *P*, the models that result differ in significant ways. It is clear from our analysis that using a simple scaling method will not yield the most successful model with respect to predicting travel times. This is especially true when converting an *S* model constructed with surface-wave data to a *P* model. Using our chosen constraints and velocity bounds, the group velocities produced an *S* model (and consequently a *P* model) that is too slow in the upper mantle.

A second motivation behind this work was to examine the regional *S* data available in the Group2 REL and EHB bulletin. We found that the scatter in the *S* data is significantly stronger than in the *P* data, which is disappointing but not unexpected. We also found that the GT *S* residuals with respect to the 3D models often reveal what seem to be phase identification errors. This phenomenon could be used to help identify incorrect associations and outliers along particular paths, if there are enough high-quality observations available. A subsequent benefit would be reduced picking errors in the *S* data, and further, their more successful incorporation into the location of small, regionally observed events.

REFERENCES

Bassin, C., G. Laske, and G. Masters (2000). The current limits of resolution for surface wave tomography in North America, *EOS Trans AGU* 81: F897.

Bondár, I., E. R. Engdahl, X. Yang, H. A. A. Ghalib, A. Hofstetter, V. Kirichenko, et al. (2004a). Collection of a reference event set for regional and teleseismic location calibration, *Bull. Seismol. Soc. Am.* 94: 1528–1545.

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. Seismol. Soc. Am.* 88: 722–743.

Engdahl, E. R., 2006. Application of an improved algorithm to high precision relocation of ISC test events, *Phys. Earth. Planet In.* 158: 14–18.

Flanagan, M. P., S. C. Myers, and K. D. Koper (2007). Regional travel-time uncertainty and seismic location improvement using a three-dimensional a priori velocity model, *Bull. Seismol. Soc. Am.* 97: 804–825.

Kennett, B. L .N., E. R. Engdahl, and R. Buland (1995). Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.* 122: 108–124.

Murphy, J.R., W. Rodi, M. Johnson, D. D. Sultanov, T. J. Bennett, M. N. Toksöz, et al. (2005). Calibration of International Monitoring System (IMS) stations in Central and Eastern Asia for improved seismic event location, *Bull. Seismol. Soc. Am.* 95: 1535–1560.

Ritzwoller, M. H., N. M. Shapiro, A. L. Levshin, E. A. Bergman, and E. R. Engdahl (2003). Ability of a global three-dimensional model to locate regional events, *J. Geophys. Res.* 108: 2353, doi:10.1029/2002JB002167.

Rodi, W. (2006). Grid-search event location with non-Gaussian error models, *Phys. Earth. Planet. Int.* 158: 55–66.

Rodi, W., and S. C. Myers (2007). Modeling travel-time correlations based on sensitivity kernels and correlated velocity anomalies, in *Proceedings of the 29th Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies*, LA-UR-07-5613, Vol. 1, pp. 463–471.

Reiter, D. T., and W. L. Rodi (2009). Validated 3D velocity models in Asia from joint regional body- and surface-wave tomography, *Weston Geophysical Corporation*, Final Report, AFRL-RV-HA-TR-2009-1009.

Yang, X., I. Bondár, J. Bhattacharyya, M. Ritzwoller, N. Shapiro, M. Antolik, et al. (2004). Validation of regional and teleseismic travel-time models by relocating ground-truth events, *Bull. Seismol. Soc. Amer.* 94: 897–919.