MODELING TRAVEL-TIME CORRELATIONS BASED ON SENSITIVITY KERNELS AND CORRELATED VELOCITY ANOMALIES

William Rodi[1] and Stephen S. Myers[2]

Massachusetts Institute of Technology[1] and Lawrence Livermore National Laboratory[2]

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

Contract No. DE-FC52-05NA26603[1] and W-7405-ENG-48[2]

ABSTRACT

The errors in the travel times predicted by a reference Earth model are correlated between different event-station paths when the paths are sufficiently close to one another, reflecting the fact that the velocity anomalies not represented in the reference model are correlated over a finite spatial extent. Previous studies have shown that ignoring these error correlations can induce large biases in seismic event locations when the observing network of stations is far from uniform. Event location algorithms have begun recently to account for travel-time correlations with the use of non-diagonal covariance matrices in the optimality criteria they employ, but the covariance values used have been either empirical or calculated from simple geometrical considerations. The goal of this project is to develop an approach to modeling travel-time correlations based on the physics of seismic travel times and geostatistical descriptions of the Earth's velocity heterogeneity. Our approach involves a two-step process: calculating travel-time sensitivity kernels for various event-station paths, and integrating the kernels, pairwise, with a specified correlation function for velocity anomalies. These steps yield the covariance matrix for the travel-time predictions among the given paths. This paper describes and illustrates the numerical techniques we are developing to perform the required calculations. Our initial focus has been on infinite-frequency travel times, but our approach will also address finite-frequency travel times to determine how the finite width of sensitivity kernels affects travel-time correlations.

OBJECTIVE

Seismic event location algorithms typically assume that errors in the seismic arrival-time data are statistically independent, or uncorrelated. In the case of regional and teleseismic arrivals, however, a significant component of the data errors are prediction errors in the model-based travel times used by the location algorithm—or "model errors"—rather than the "pick errors" committed in the measurement process. Model errors are attributable to velocity anomalies in the Earth that are not rendered in the velocity model used for travel-time prediction. When the observing stations are sufficiently close to one another, their model errors will be correlated. While it is now standard practice in nuclear monitoring applications for event location algorithms to assign error *variances* that include the contribution from model prediction errors, the algorithms generally set the error *covariances* to zero; correlations are ignored. Doing so can seriously degrade location accuracy when the distribution of seismic stations is far from uniform, as was shown, for example, by Chang et al. (1983) and Yang et al. (2004).

This project is using a model-based approach to investigate correlations between the travel-time prediction errors that are induced by velocity heterogeneities in the Earth. We are developing techniques to calculate covariance matrices of travel-time model errors by integrating travel-time sensitivity kernels with plausible correlation functions describing the velocity heterogeneity of the crust and upper mantle. Our initial focus is on first Pn arrivals (epicentral distances between about 2° and 20°) and infinite-frequency travel times, whose sensitivities are concentrated along geometrical rays. Later, the project will address direct teleseismic arrivals (to about 90°) and finite-frequency travel times, whose sensitivities are given by spatially distributed banana-doughnut kernels (e.g., Zhao et al., 1999; Dahlen et al., 2000; Hung et al., 2000). An important issue we will consider is whether the spatial extent of banana-doughnut kernels causes a significant increase in model-error correlations.

This paper formulates our approach to modeling travel-time correlations and describes some of the numerical techniques we have developed thus far, illustrated with an example involving Pn correlations between stations in Asia.

RESEARCH ACCOMPLISHED

Formulation of Approach

Given a set of n arrival time data from an event, one can decompose the data errors as (for $i = 1,\ldots,n$)

$$e_i = e_{\mathrm{p},i} + e_{\mathrm{m},i} \tag{1}$$

where the first term is the *pick*, or measurement, error and the second term is the *model* error, or error in the model-based travel-time prediction. Location algorithms such as EvLoc assume the above decomposition and set the data variances accordingly as

$$\sigma_i^2 = \sigma_{\mathrm{p},i}^2 + \sigma_{\mathrm{m},i}^2. \tag{2}$$

However, the errors for different i are assumed uncorrelated. The problem we are addressing is the calculation of a full covariance matrix for model errors, having components $\sigma_{\mathrm{m},ij}$ defined by

$$\sigma_{\mathrm{m},ij} = \mathrm{E}[e_{\mathrm{m},i}\, e_{\mathrm{m},j}]. \tag{3}$$

E[] denotes the expectation operator.

To calculate the model error covariance matrix, we are making two simplifying assumptions: (1) that the travel-time dependence on velocity is well-approximated as a linear departure from the reference Earth model used for travel-time prediction, and (2) that 3D velocity heterogeneity can be described by a Gaussian random field. Considering only P-wave arrivals, let $v_0(\mathbf{x})$ denote the P velocity function for the reference model, and let $v_E(\mathbf{x})$ denote the Earth's true (and unknown) velocity function. Then, model errors can be linked to the slowness difference, which we denote $m(\mathbf{x})$:

$$m(\mathbf{x}) = v_E^{-1}(\mathbf{x}) - v_0^{-1}(\mathbf{x}). \tag{4}$$

Linearization of the travel-time dependence on slowness implies that the model error in the *i*th datum is given by

$$e_{\mathrm{m},i} = \int a_i(\mathbf{x})\, m(\mathbf{x})\, d\mathbf{x} \tag{5}$$

where $a_i(\mathbf{x})$ is the first-order sensitivity kernel of the ith travel-time functional, as evaluated at v_0. In the high-frequency limit, this kernel is concentrated along the geometrical ray connecting the event and station locations. For finite frequency, it is spatially distributed around this ray. In either case, we point out that the model error is a function of the event *and* station locations.

Now consider a geostatistical description of $m(\mathbf{x})$, assuming it to be a Gaussian random field with zero mean at all points and having a specified covariance between any two points, as described by a covariance function $C(\mathbf{x},\mathbf{x}')$:

$$\mathrm{E}\left[m(\mathbf{x})\right] = 0 \tag{6}$$

$$E\left[m(\mathbf{x})m(\mathbf{x}')\right] = C(\mathbf{x},\mathbf{x}'). \tag{7}$$

The covariance between two model errors is then given by

$$\sigma_{\mathrm{m},ij} = \int d\mathbf{x}\, a_i(\mathbf{x}) \int d\mathbf{x}'\, C(\mathbf{x},\mathbf{x}')\, a_j(\mathbf{x}'). \tag{8}$$

The core task of this project is the evaluation of the double volume integral in this equation.

Equation (8) implies that the extent to which model errors are correlated depends on how strongly correlated the velocity field is *and* on how spatially distributed the sensitivity kernels are. Our project is considering both effects.

Velocity Model Covariance Functions

Geostatistical methods for data interpolation (kriging, e.g. Deutsch and Journel, 1998; Schultz et al., 1998) typically specify the covariance function, $C(\mathbf{x},\mathbf{x}')$, directly and, assuming stationarity, allow it to be a function only of the separation between the points $\mathbf{x}$ and $\mathbf{x}'$:

$$C(\mathbf{x},\mathbf{x}') = \rho(\mathbf{x}-\mathbf{x}'), \tag{9}$$

where ρ is called the *correlation* function. Simple examples of correlation functions are ones of the *Gaussian* type,

$$\rho(\mathbf{x}-\mathbf{x}') = \sigma_0^2 \exp\left[-\frac{|\mathbf{x}-\mathbf{x}'|^2}{2\lambda^2}\right], \tag{10}$$

which characterize smoothly varying velocity fluctuations, and ones of the *exponential* type,

$$\rho(\mathbf{x}-\mathbf{x}') = \sigma_0^2 \exp\left[-\frac{|\mathbf{x}-\mathbf{x}'|}{\lambda}\right], \tag{11}$$

for rougher fluctuations. (Note: the random field described by a correlation function has a Gaussian probability distribution—i.e., is a *Gaussian* random field—regardless of the type of correlation function.) In these expressions, σ_0^2 is the variance of the slowness field and λ is a correlation length. These examples define *isotropic* fields, having the same spatial statistics in all directions.

An indirect way to specify a model covariance function is through the operator inverse of $C(\mathbf{x},\mathbf{x}')$, which we denote *D*, such that

$$DC(\mathbf{x},\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}'). \tag{12}$$

If we take *D* to be a differential operator, then $C(\mathbf{x},\mathbf{x}')$ is its Green's function. Rodi et al. (2003) implemented this approach with *D* as

$$D = \frac{\text{const}}{\lambda_1^2\lambda_2\sigma_0^2}\left[\delta(\mathbf{x}) - \frac{1}{(2l-3)}\left(\frac{\lambda_1^2}{r^2}\nabla_1^2 + \lambda_2^2\frac{\partial^2}{\partial z^2}\right)\right]^l, \quad l \geq 2. \tag{13}$$

Here, l is an integer specifying the order of the operator; ∇_1^2 is the horizontal Laplacian operator; λ_1 and λ_2 are horizontal and vertical correlation lengths, respectively; and σ_0^2 is a prior variance. The Gaussian and exponential type correlation functions correspond to $l = \infty$ and $l = 2$, respectively. An advantage of this approach is that one can introduce non-stationarity by allowing (smooth) spatial variations in σ_0 and $\lambda_{1,2}$, while preserving the essential properties of $C(\mathbf{x},\mathbf{x}')$ as a proper covariance function (e.g., positive-definiteness).

The indirect method of specifying the covariance operator is the basis of numerical techniques we have developed thus far for computing model-error covariance matrices.

Numerical Techniques

The numerical techniques we have developed apply to a discrete model parameterization. We thus replace the model function $m(\mathbf{x})$ with a *k*-dimensional vector **m**; the sensitivity kernels $a_i(\mathbf{x})$ with sensitivity vectors $\mathbf{a}_i$; and the operators *C* and *D* with matrices **C** and **D**. We require that

$$\mathbf{DC} = \mathbf{I}, \tag{14}$$

where **I** is the $k \times k$ identity matrix. This requirement implies that **C** and **D** cannot be independently generated as discrete approximations to *C* and *D*.

In the discrete formulation, the double integral of equation (8) becomes the quadratic matrix expression

$$\sigma_{\mathrm{m},ij} = \mathbf{a}_i^T\mathbf{C}\mathbf{a}_j. \tag{15}$$

Let us also define the $n \times k$ sensitivity matrix having the $\mathbf{a}_i$ as its rows:

$$\mathbf{A} = \left(\mathbf{a}_1 \;\; \mathbf{a}_2 \;\; \cdot \;\; \mathbf{a}_n\right)^T. \tag{16}$$

If we let the $\sigma_{\mathrm{m},ij}$ be the components of an $n \times n$ matrix $\mathbf{\Sigma}_\mathrm{m}$, we can rewrite equation (15) as

$$\Sigma_\mathrm{m} = \mathbf{ACA}^T. \tag{17}$$

$\mathbf{\Sigma}_\mathrm{m}$ is the covariance matrix of the *n* model errors. We can now say that the core task of this project is the evaluation of the matrix product in equation (17).

We have formulated a numerical method for computing $\mathbf{\Sigma}_\mathrm{m}$ given **D** and **A** as input. It applies to the situation where the matrix **D** is given explicitly as a finite-difference approximation to the differential operator *D*, with **C** being defined implicitly as the inverse of **D** as per equation (14). However, our technique avoids the step of generating **C** from **D** with matrix inversion.

Instead, our approach solves a minimization problem for each datum, $i = 1, 2, n$, stated as

$$\left(\mathbf{n}_i - \mathbf{A}\mathbf{m}_i\right)^T \Sigma_\mathrm{p}^{-1}\left(\mathbf{n}_i - \mathbf{A}\mathbf{m}_i\right) + \mathbf{m}_i D\mathbf{m}_i = \text{minimum} \tag{18}$$

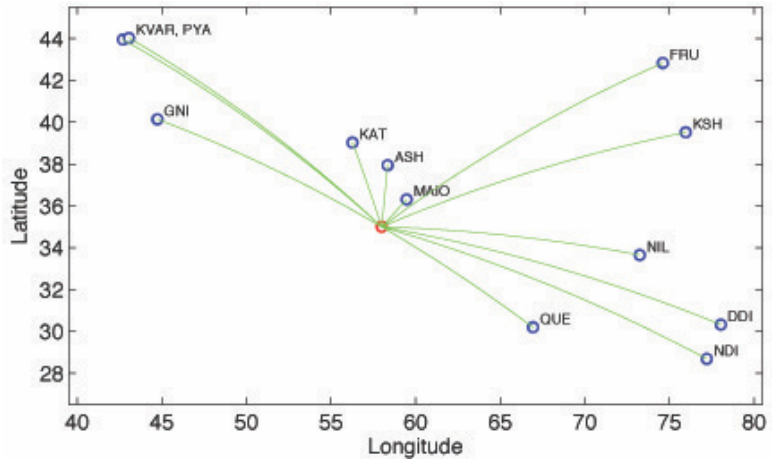

Figure 1. Event/station geometry used for testing our travel-time covariance modeling algorithm. The event location (red circle) is for one of the summary events formed by Reiter and Rodi (2006, these Proceedings) for their body-wave tomography application in south/central Asia. The summary event had Pn arrivals for 12 stations (blue circles).

where $\mathbf{\Sigma}_\mathrm{p}$ is the diagonal covariance matrix for pick errors, and $\mathbf{n}_i$ is the ith column of the $n \times n$ identity matrix. We compute the solution, $\mathbf{m}_i$, of each minimization problem iteratively using the conjugate gradients method, since $\mathbf{D}$ and $\mathbf{A}$ are sparse matrices. To see how $\mathbf{\Sigma}_\mathrm{m}$ can be generated from the resulting solutions, let us consider analytical expressions for the solutions, given by

$$\mathbf{m}_i = \mathbf{C}\mathbf{A}^T \left(\mathbf{A}\mathbf{C}\mathbf{A}^T + \Sigma_\mathrm{p}\right)^{-1} \mathbf{n}_i. \tag{19}$$

The residual vectors can then be written:

$$\mathbf{r}_i \equiv \mathbf{n}_i - \mathbf{A}\mathbf{m}_i = \Sigma_\mathrm{p}\left(\mathbf{A}\mathbf{C}\mathbf{A}^T + \Sigma_\mathrm{p}\right)^{-1} \mathbf{n}_i. \tag{20}$$

Arranging all n residual vectors into a matrix **R,**

$$\mathbf{R} = \begin{pmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \cdots & \mathbf{r}_n \end{pmatrix}, \tag{21}$$

we find

$$\mathbf{R} = \Sigma_\mathrm{p}\left(\mathbf{A}\mathbf{C}\mathbf{A}^T + \Sigma_\mathrm{p}\right)^{-1}. \tag{22}$$

A location algorithm uses the inverse of the covariance matrix for the total errors (pick plus model). This is given as

$$\left(\Sigma_\mathrm{p} + \Sigma_\mathrm{m}\right)^{-1} \equiv \left(\Sigma_\mathrm{p} + \mathbf{A}\mathbf{C}\mathbf{A}^T\right)^{-1} = \Sigma_\mathrm{p}^{-1}\mathbf{R}. \tag{23}$$

The model-error covariance matrix, which is not needed explicitly for location but is the object of our study, is obtained with the additional calculation:

$$\Sigma_\mathrm{m} \equiv \mathbf{A}\mathbf{C}\mathbf{A}^T = \left(\Sigma_\mathrm{p}^{-1}\mathbf{R}\right)^{-1} - \Sigma_\mathrm{p}. \tag{24}$$

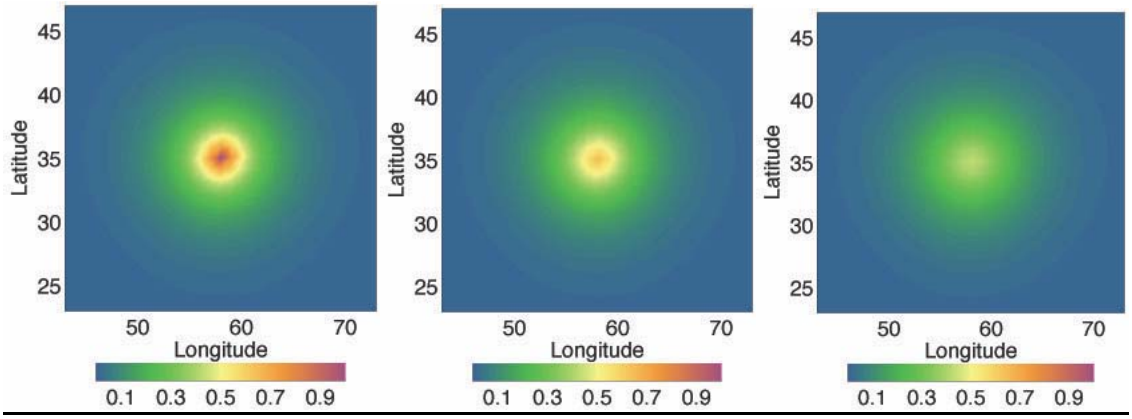

Figure 2. Horizontal slices of a 3D correlation function at depths of approximately 40 km (left panel), 80 km (center) and 120 km (right). The value at each point is the correlation coefficient between the slowness at that point and the slowness at the fixed point: 35°N, 58°E, z = 40 km. The correlation function was computed numerically by solving a differential equation using the differential operator of equation (13). The operator was assigned order l = 2, horizontal correlation length λ_1 = 300 km, and vertical correlation length λ_2 = 100 km.

It is informative to recognize that each minimization problem solved with this method, equation (18), is essentially *single-event* tomography, i.e. velocity tomography using data from only one event.

Numerical Results

We have tested our numerical approach on the event/station geometry shown in Figure 1. The example is extracted from the body-wave tomography data set developed by Reiter and Rodi (2006). The event is one of the summary events they generated from the EHB (Engdahl et al., 1998) data base for the years 1988–2004. The summary event—located at 35°N, 58°E (northeastern Iran) with a depth of 15 km—is associated with Pn arrivals from 12 stations, ranging from 1.8° to 17.5° in epicentral distance and spanning 180° in azimuth. We performed our calculations using the sensitivity vectors (discretized kernels) generated by Reiter and Rodi (2006, these Proceedings), which they computed using finite difference raytracing in a 3D reference Earth model.

We defined the velocity model covariance function indirectly with the differential operator D of equation (13), represented as a differencing matrix (**D**) in spherical coordinates. The order of the differential operator (l) was set to 2, implying a correlation function of the exponential type. The horizontal and vertical correlation lengths (λ_1 and λ_2) were set to 300 km and 100 km, respectively. The slowness standard deviation (σ_0) was set to 2% of the slowness itself at each point of the reference model. We constructed the differential operator to omit vertical derivatives across the Moho discontinuity, implying that velocity anomalies in the crust and mantle are uncorrelated. This crust/mantle decoupling, and the variation of σ_0 with position, are two mechanisms for non-stationarity that are easily accomplished with the differential operator approach.

To demonstrate the equivalence of specifying the model covariance function directly and indirectly, we computed $C(\mathbf{x},\mathbf{x}_0)$ for a fixed point $\mathbf{x}_0$. We chose this point as 35°N, 58°E (the summary event epicenter) with a depth at the top of the mantle. The computation entailed solving the linear system

$$\mathbf{D}\mathbf{c}_0 = \mathbf{n}_0 \tag{25}$$

with the vector $\mathbf{n}_0$ being an appropriate column of the identity matrix and the solution vector $\mathbf{c}_0$ being a discrete approximation to the desired covariance function. The solution function is zero in the crust (because of the crust/mantle decoupling). Figure 2 shows horizontal slices of the function for three depths in the mantle. The function has been scaled by the inverse variance so that it becomes a correlation function, with a peak value of one

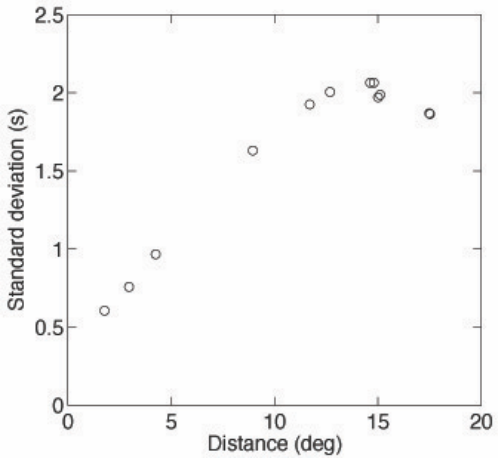

Figure 3. Standard deviation of model error as a function of epicentral distance.

at $\mathbf{x} = \mathbf{x}_0$. We see that the numerically computed correlation function displays spatial behavior consistent with the input parameters, including both its lateral decay from the peak point and its decay with depth.

Our main test was to calculate the 12×12 covariance matrix, $\mathbf{\Sigma}_m$, for the model errors at the 12 stations. This involved solving the single-event tomography problems of equation (18) and then evaluating equation (24) with the resulting matrix of data residuals.

We show the results in two parts. The standard deviations of the model errors, obtained as the square root of the diagonal elements of the covariance matrix, are listed in Table 1 and plotted as a function of epicentral distance in Figure 3. The plot displays the expected increase with epicentral distance, but leveling off beyond about 12°. The model-error standard deviation at the two farthest stations (17.5°) is slightly less than some closer stations, which may reflect the deeper sampling of Pn rays at far regional distances or, possibly, azimuthal variation due to the use of a 3D reference model. We point out that Pn model-error standard deviations of 2 s and less, as these are, are not very large; this reflects our use of 2% for the Earth velocity standard deviation, which underestimates at least the crustal component of the heterogeneity.

The second part of the model-error covariance results consists of the correlation coefficients between model errors at different stations. Table 2 lists the full 12×12 correlation matrix, obtained by dividing each row and column of the covariance matrix ($\mathbf{\Sigma}_m$) by the appropriate standard deviation. The correlation matrix is symmetric and has unit diagonal, so there are 66 independent coefficients in the table, corresponding to the number of possible station pairs. The 66 values are plotted in Figure 4 as a function of the azimuth difference between the two stations in a pair. Looking at Table 2, we see that the highest correlations are between stations at similar locations, e.g. 0.99 for KVAR/PYA and 0.91 for DDI/NDI. The three closest stations (ASH, KAT, MAIO) also have highly correlated errors. From Figure 4 we see that the correlation coefficient generally decreases with the azimuth difference between stations. One would expect a monotonic decrease if the stations were equidistant from the event and the reference model were 1D, neither of which is the case. Stations MAIO and FRU, for example, are at similar azimuths but at very different distances and have a correlation of only 0.24. We also note that the correlation coefficient between any two stations is never below 0.1, which must be due to the fact that all the raypaths sense the structure within one correlation length of the event hypocenter.

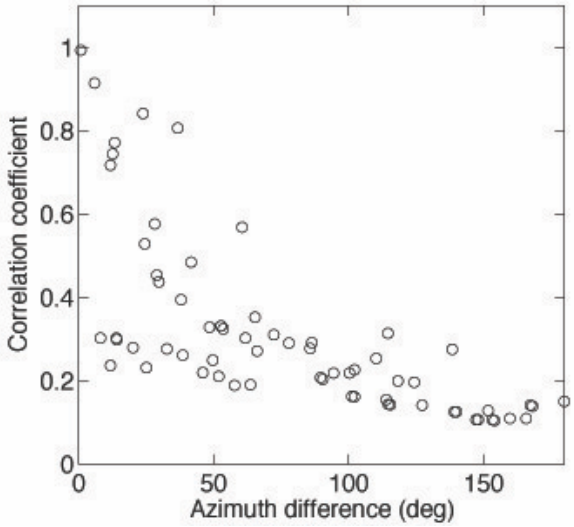

Figure 4: Correlation coefficient between model-error pairs as a function of azimuth difference.

CONCLUSIONS AND RECOMMENDATIONS

Our initial efforts in this project confirm the feasibility of using a model-based approach to investigate the correlation structure of travel-time model errors, as a function of event location and seismic network geometry. The need to account for error correlations in seismic event location algorithms has been demonstrated in several previous studies (e.g. Yang et al., 2004) but the estimation of correlations has been restricted primarily to empirical studies in data rich areas (e.g., Myers and Schultz, 2000). We expect the approach we are developing to provide a useful tool for extrapolating such studies to data poor (aseismic) areas, taking the principles of seismic wave propagation and our knowledge of Earth structure heterogeneity into account.

Table 1. Standard deviation of model errors

Station	*Dist. (deg)*	*Az. (deg)*	*Std. dev. (s)*
ASH	2.96	5.4	0.76
DDI	17.51	99.8	1.86
FRU	15.10	53.9	1.99
GNI	11.70	299.9	1.93
KAT	4.25	341.5	0.97
KSH	15.00	67.2	1.97
KVAR	14.80	311.7	2.07
MAIO	1.77	42.1	0.61
NDI	17.49	105.8	1.87
NIL	12.68	91.8	2.00
PYA	14.62	312.6	2.07
QUE	8.94	112.0	1.63

Table 2. Correlation coefficient between model errors

	ASH	DDI	FRU	GNI	KAT	KSH	KVAR	MAIO	NDI	NIL	PYA	QUE
ASH	1.00	.22	.33	.35	.84	.30	.32	.81	.22	.29	.33	.31
DDI	.22	1.00	.22	.11	.20	.28	.11	.19	.91	.30	.11	.28
FRU	.33	.22	1.00	.16	.31	.77	.16	.24	.21	.40	.16	.27
GNI	.35	.11	.16	1.00	.48	.14	.72	.23	.11	.13	.74	.15
KAT	.84	.20	.31	.48	1.00	.28	.44	.57	.20	.25	.45	.27
KSH	.30	.28	.77	.14	.28	1.00	.14	.23	.26	.53	.14	.33
KVAR	.32	.11	.16	.72	.44	.14	1.00	.20	.11	.12	.99	.14
MAIO	.81	.19	.24	.23	.57	.23	.20	1.00	.19	.25	.21	.29
NDI	.22	.91	.21	.11	.20	.26	.11	.19	1.00	.30	.11	.30
NIL	.29	.30	.40	.13	.25	.53	.12	.25	.30	1.00	.13	.58
PYA	.33	.11	.16	.74	.45	.14	.99	.21	.11	.13	1.00	.14
QUE	.31	.28	.27	.15	.27	.33	.14	.29	.30	.58	.14	1.00

ACKNOWLEDGMENTS

We are very grateful to Prof. Vernon Cormier of the University of Connecticut for collaborating with us on seismic-wave propagation theory and the future extension of our approach to finite-frequency travel times.

REFERENCES

Chang, A. C., R.H. Shumway, R. R. Blandford and B. W. Barker (1983). Two methods to improve location estimates—preliminary results, *Bull. Seism. Soc. Am.* 73: 281–295.

Dahlen, F.A., S.-H. Hung and G. Nolet (2000). Fréchet kernels for finite-frequency traveltimes I. Theory, *Geophys. J. Int.* 141: 157–174.

Deutsch, C. V. and A. G. Journel (1998). *GSLIB: Geostatistical Software Library and User's Guide*, 2nd ed., Oxford University Press, Inc., New York, 369 pp.

Engdahl E. R., R. van der Hilst and R. Buland (1998). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seism. Soc. Am.* 88: 722–743.

Hung, S.-H., F. A. Dahlen and G. Nolet (2000). Fréchet kernels for finite-frequency traveltimes II. Examples, *Geophys. J. Int.* 141: 175–203.

Myers, S. C., and C. A. Schultz (2000). Improving sparse network seismic locations with Bayesian kriging and teleseismically constrained calibration events, *Bull. Seism. Soc. Am.* 90: 199–211.

Reiter, D. and W. Rodi (2006). Crustal and upper-mantle P- and S-velocity structure in central and southern Asia from joint body- and surface-wave inversion, in current Proceedings.

Rodi, W., S. C. Myers and C. A. Schultz (2003). Grid-search location methods for ground-truth collection from local and regional seismic networks, in *Proceedings of 25th Seismic Research Review—Nuclear Explosion Monitoring: Building the Knowledge Base*, LA-UR-03-6029, Vol. 1, pp. 311–319.

Schultz, C.A., S.C. Myers, J. Hipp and C.J. Young (1998). Nonstationary Bayesian kriging: a predictive technique to generate spatial corrections for seismic detection, location, and identification, *Bull. Seism. Soc. Am.* 88: 1275–1288.

Yang, X., I. Bondár, J. Bhattacharyya, M. Ritzwoller, N. Shapiro, M. Antolik, G. Ekström, H. Israelsson and K. McLaughlin (2004). Validation of regional and teleseismic travel-time models by relocating ground-truth events, *Bull. Seism. Soc. Am. 94:* 897–919.

Zhao, L., T.H. Jordan and C.H. Chapman (2000). Three-dimensional Fréchet differential kernels for seismic delay times, *Geophys. J. Int., 141*, 558–576.