Evaluation of Model Resolution with Various Forward Theory and Parameterization: A Synthetic Study

A. Abstract

Whether different forward theories (data rules) and parameterizations employed in tomographic imaging lead to the improvement of the resulting Earth structures has been a focus of attention in the seismological community. Recent advance in tomographic theory has gone beyond ray theory and incorporated the 3-D sensitivity kernels of frequency-dependent travel-time data into probing the mantle velocity heterogeneity. On the other hand, the idea of multiscale parameterization has been introduced to deal with naturally uneven data distribution and spatially-varying model resolution in the inversion. The multi-resolution model automatically built through the wavelet decomposition and synthesis results in the non-stationary spatial resolution and data-adaptive resolvable scales. Because the Gram matrix of sensitivity kernels that relates observed data to seismic velocity variations is usually too large to be practically inverted by singular value decomposition (SVD), the iterative LSQR algorithm is instead used in the inversion which inhibits the calculation of model resolution and covariance to assess the model performance. With the increasing computing power, the SVD of the large Gram matrix becomes viable by the parallel PROPACK solver. In this study, we test the resolvability of a synthetic 3-D random model using ground-truth travel-time residuals and various data rules and parameterizations. The source-receiver configuration mimics the borehole tomography. The tradeoff relations between model covariances (errors) and model spreads for the models with the same data rule and parameterization are used

to determine the optimal models. The optimal finite-frequency models always yield larger model norms and lower model misfits regardless of the parameterization and regularization adopted in the inversion. The optimal multi-scale models also yield larger model norms and longer-wavelength structures and have smaller model covariances obtained with damping regularization. The inversion based on finite-frequency theory, multi-scale parameterization, and damping regularization leads to the best-fitting model.

B. Introduction

Seismic tomographic inversion always encounters the problems of model non-uniqueness and resolution resulting from imperfect source-station distribution, finite parameterization of continuous velocity model and finite-bandwidth seismic data. Recent advance in finite-frequency theory has taken into account 3-D volumetric sensitivity kernels of frequency-dependent seismic travel-time data in the tomographic imaging (e.g., Hung et al., 2004). On the other hand, multiscale parameterization has been advocated to obtain the models with spatially-varying and data-adaptive resolutions without employing a priori smoothness constraints (Chiao and Kuo, 2001; Chiao and Liang, 2003). Because the Gram matrix that links observed data to model parameters is usually too vast to be inverted directly, a computationally- economic iterative LSQR algorithm (Paige and Saunders, 1982) is instead utilized to solve for the model parameters. To avoid the under- or over-interpretation of the resolved velocity structures, the tradeoff analysis between data fits and model norms or variances are performed to determine the optimal model solution. The checkerboard and/or spike tests limited to stationary variations and a

small amount of model parameters are alternatively used to estimate the resolution of the resulting models. With the great leap in computing power, a PROPACK solver for singular value decomposition of sparse matrices (Larsen, 1998) has made it feasible to invert the large Gram matrix in travel-time tomography and calculate the resolution matrix exactly. In this study, we use a set of synthetic ground-truth travel-time data generated in known velocity structures to invert the input structure and assess how the model resolution and covariance are influenced by the chosen data rule and parameterization.

C. Model configuration

Following the previous study (Yang and Hung, 2005), we generate synthetic finite-frequency waveforms in 3-D Gaussian random acoustic media with the characteristic scalelength, 1.5λ, and strength of velocity perturbation, 2% (Fig.1a). The model is a box with the dimension of 5740×5740×5120 km in x, y, and z directions, respectively (Fig.1b). The data are the travel-time shifts between the acoustic wave arrivals in the 3-D heterogeneous structure and those in the homogeneous background medium (Fig. 2) measured by waveform cross correlation and adding one standard deviation of random errors. We parameterize the model space into equally-spaced grid nodes at which the slowness (the inverse of velocity) perturbation are inverted for (Fig. 3).

D. Singular Value Decomposition : Simple-grid vs. Multi-scale

Assume d represents the data vector that comprises *N* measured travel-time data and m the model vector composed of *M* discretized model parameters at nodes, the tomographic inversion problem can be written in a concise matrix form

$$\mathbf{Gm = d}$$

where **G** is the N-by-M Gram matrix, which can be factorized into three matrices through the singular value decomposition (SVD),

$$\mathbf{G} = \mathbf{U}\Lambda\mathbf{V}^{\mathbf{T}}$$

where $\mathbf{\Lambda}$ is an *N*-by-*M* diagonal matrix with all the singular values of **G**, **U** an *N*-by-*N* unitary matrix, and $\mathbf{V}^{T}$ the transpose of an *M*-by-*M* unitary matrix. The estimated model under simple grid parameterization is then obtained from,

$$\hat{\mathbf{m}} = \mathbf{G}^{-1}\mathbf{d} = \mathbf{V}(\Lambda^{-1})\mathbf{U}^{\mathbf{T}}\mathbf{d}$$

The corresponding resolution matrix **R** is determined from

$$\mathbf{R} = \mathbf{G}^{-1}\mathbf{G}$$

The model parameters, **m'**, under the wavelet-based multiscale parameterization can be obtained by applying a transforming operator, **W**, on m for those parameterized at grid nodes, that is,

$$\mathbf{m'} = \mathbf{Wm}$$

Similarly, the resolution matrix for the multi-scale model, **R'**, can be derived from that for grid parameterization,

$$\mathbf{R'} = \mathbf{WRW}^{-1}$$

where $\mathbf{W}^{-1}$ is the inverse transforming operator of **W**.

Assuming each data is independent and has the same error, the model covariance matrix in the simple grid parameterization is calculated by

$$\mathbf{C}_{\mathbf{m}} = \mathbf{G}^{-1}(\mathbf{G}^{-1})^{\mathbf{T}}$$

Likewise, the model covariance matrix for the multi-scale model, C_m', can be obtained by

$$\mathbf{C_m' = WC_m W^T}$$

For each type of parameterization, we employ two data rules: linearized ray and finite-frequency theory (Fig. 4). In the following, we conduct the tradeoff analysis between model covariance and model spread to determine the optimal models for each type of the data rule, parameterization, and regularization adopted in the inversion. The model spread is estimated by the summation of (1-diag(R))2, where diag(R) is the diagonal value of the resolution matrix, R, while the model covariance is calculated by the summation of the squares of the diagonal values of the model covariance matrix. Both values are determined at the grid-basis domain and wavelet-basis domain for the simple grid and multi-scale models, respectively.

E. Result

Figure 6 shows the tradeoff relation between model covariance (error) and model spread for the models with various data rules, parameterizations, and regularizations. (a) With damping regularization, the optimal multi-scale models yield significantly smaller model covariances than the simple grid models. (b) With truncation regularization, the model covariances for all the types of the optimal models are similar.

Figure 7 shows the 2-D cross sections of the resulting velocity structures of the optimal models. The multi-scale, finite-frequency models yield longer-wavelength structures with larger velocity perturbations, while the simple grid, ray-based models are rougher and recover smaller amplitudes of velocity perturbations.

By comparing the tradeoffs between model covariance and variance reduction (data fit) and model norm and variance reduction shown in Fig. 8, the finite-frequency

models give larger model covariances and norms under the same data fit, regardless of the adopted parameterization and regularization. The multi-scale models obtained with damping lead to smaller model covariances but those with truncation has larger model covariances. Under the same data rule and regularization, all the multi-scale models have larger model norms. Fig. 9 compare the model misfit varying with the variance reduction and model covariance. The model misfit is estimated by , M is total model parameters, where mhat is the inverted model and mtrue the input random model. Among all the types of the solutions, the optimal multi-scale, finite-frequency model with damping regularization best fits to the true model.

F. Discussion and Conclusions

1.Our synthetic experiments of borehole tomography indicate that the models obtained with singular value truncation yield smaller model covariances. Under the same data rule with damping, the multi-scale models has smaller model covariances than the simple grid models.

2.From the tradeoff analysis between model covariance and model spread, the optimal finite-frequency models yield larger model norms and covariances than the ray-based models regardless of the parameterization and regularization. The optimal multi-scale models lead to smaller model covariances than the simple grid models, particularly under the damping regularization and have always larger norms and longer-wavelength structures.

3.Overall, the finite-frequency theory results in the models with better fits to the input structure, regardless of the parameterization and regularization. The model misfits for the damping solutions are smaller than the truncation solutions. The optimal multi-

scale, finite-frequency model obtained with damping regularization gives the best fitting to the true model.

G. Reference

Baig, A.M., F.A. Dahlen & S.-H. Hung, 2003., Geophys. J. Int., 153, 467-482.

Chiao, L.-Y., & B.-Y. Kuo, 2001., Geophys. J. Int., 145, 517-527.

Chiao, L.-Y & W.-T. Liang, 2003., Geophysics, 68, 199-209

Hung, S.-H., Y. Shen, & L.-Y. Chiao, 2004. , J. Geophys. Res., 109, B08305, doi: 10.1029/2003JB002889, 1-16.

Larsen, R.M., 1998. Lanczos bidiagonalization with partial reorthogonalization, Department of Computer Science, Aarhus University, Technical Report, DAIMI PB-357, September 1998. The PROPACK package realizing this algorithm is available at http://soi.stanford.edu/~rmunk/PROPACK/ .

Paige, C.C. & M.A. Saunders, 1982., Trans. Mart. Software, 8,43–71.

Yang, H.-Y., & S.-H. Hung, 2005. , Geophys. Res. Lett., 32, L20302, doi: 10.1029/2005GL023501

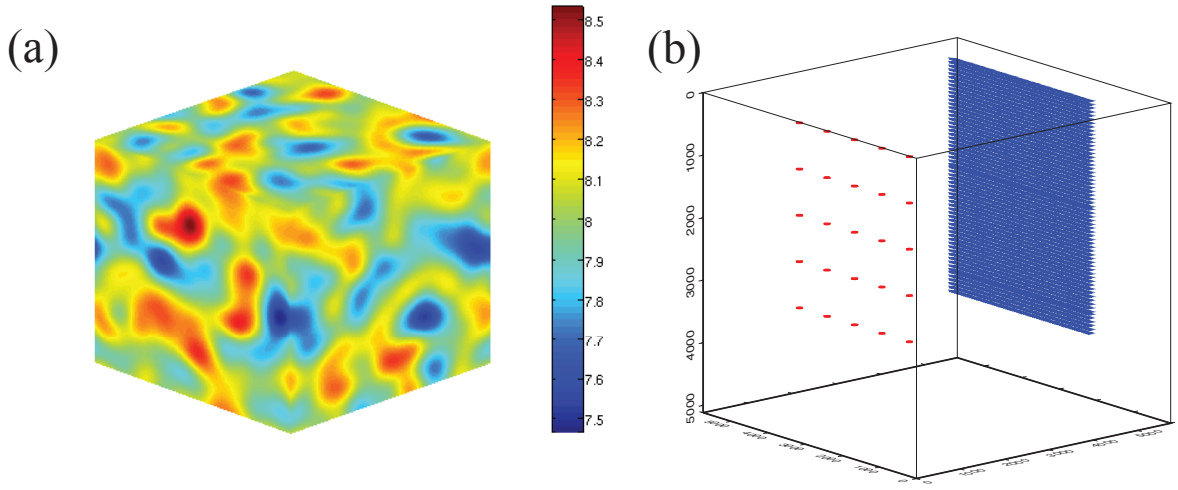

Fig.1 (a) A 3-D random acoustic medium used to generate synthetic travel-time data. (b) Source-station distribution and model space used in the inversion. There are 25 point sources denoted by red circles placed on the vertical plane at x=1000 km, and each of them is spaced by 380 km in the y and z direction. Likewise, 3969 stations shown by blue triangulars spaced 60 km apart are located on the vertical plane at x=4740 km. The model space for the inversion experiments is confined in the region between the sources and stations, discretized into 33x33x33 nodes with equal spacings of ~117 km in each dimension.

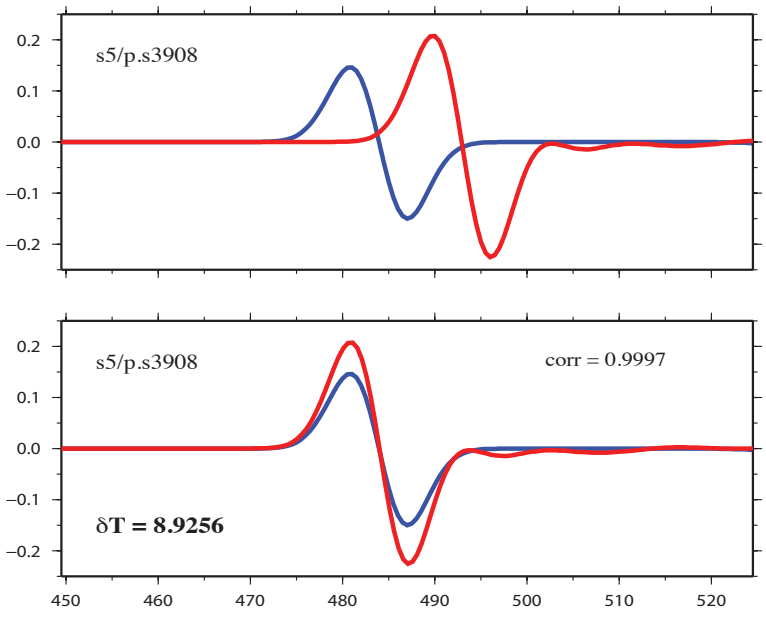

Fig.2. Example of a finite-frequency travel-time shift measured by waveform cross correlation used for the tomographic inversion. The synthetic acoustic wave arrivals in the heterogeneous and homogeneous structure before and after cross correlation are shown on the top and bottom panel, respectively.

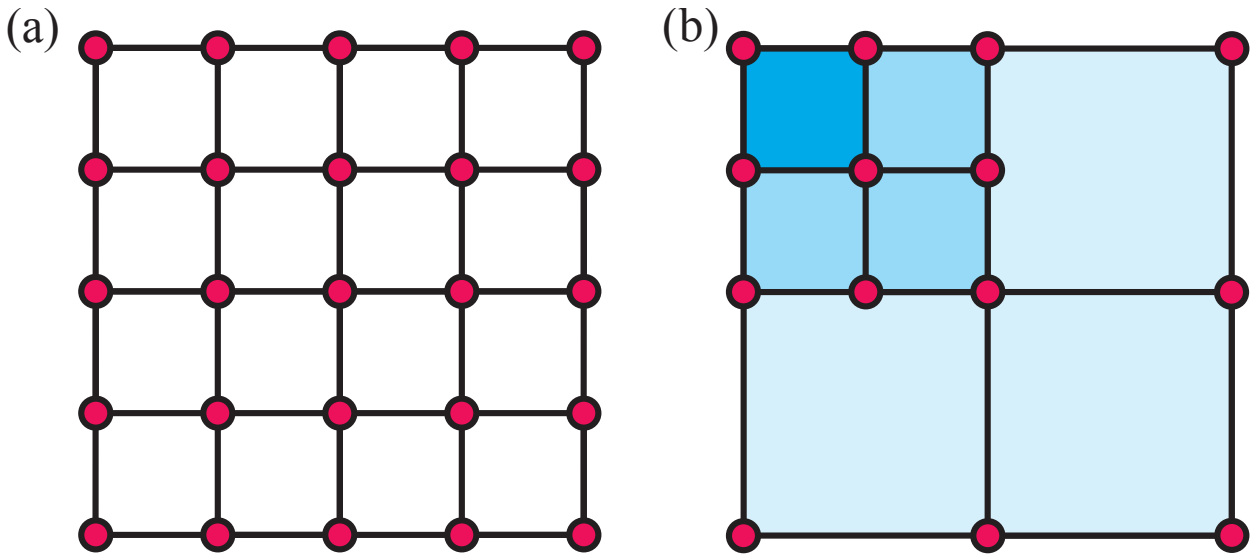

Fig.3. Cartoon shows (a) simple grid parameterization where the model parameters of slowness perturbations are solved directly at discretized, equally-spaced nodes; (b) multi-scale parameterization built through five successive refinements from the coarsest level comprising the entire model to the finest leaf level equivalent to the grid spacing in simple grid parameterization.

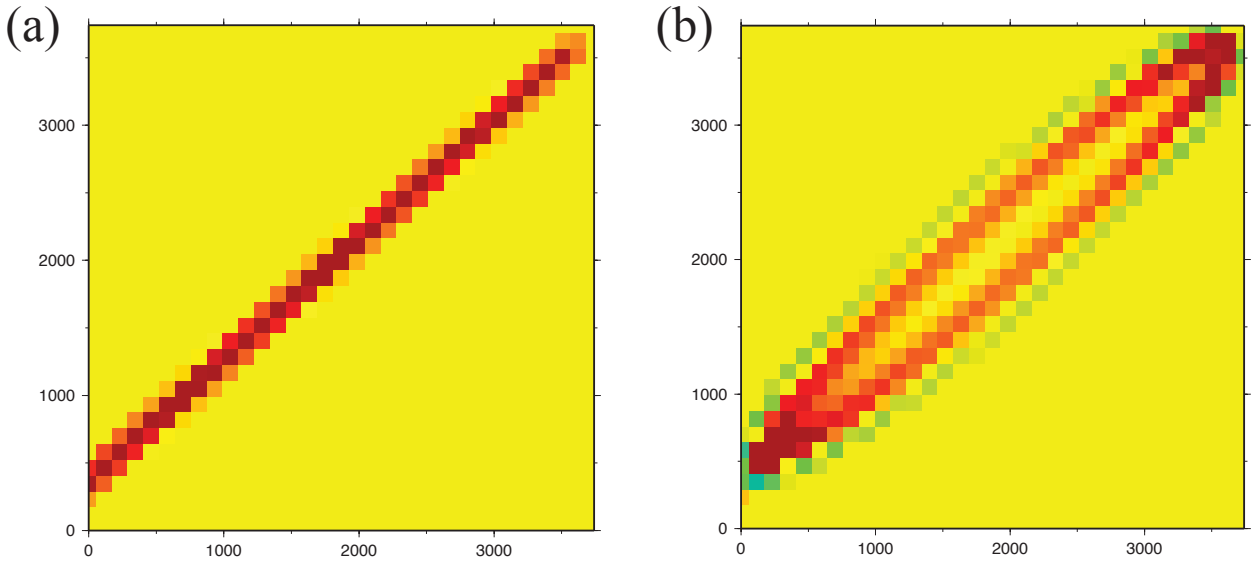

Fig.4 2-D cross sectional views of travel-time kernels based on (a) linearized ray and (b) finite-frequency theory. The sensitivity of a ray-theoretical travel-time shift in a homogeneous background medium is simply a straight-line fat ray between the

source and station, while the sensitivity of a finite-frequency travel-time shift resembles a banana shape. The velocity structure located in the red regions contribute mostly to the observed travel-time shifts.

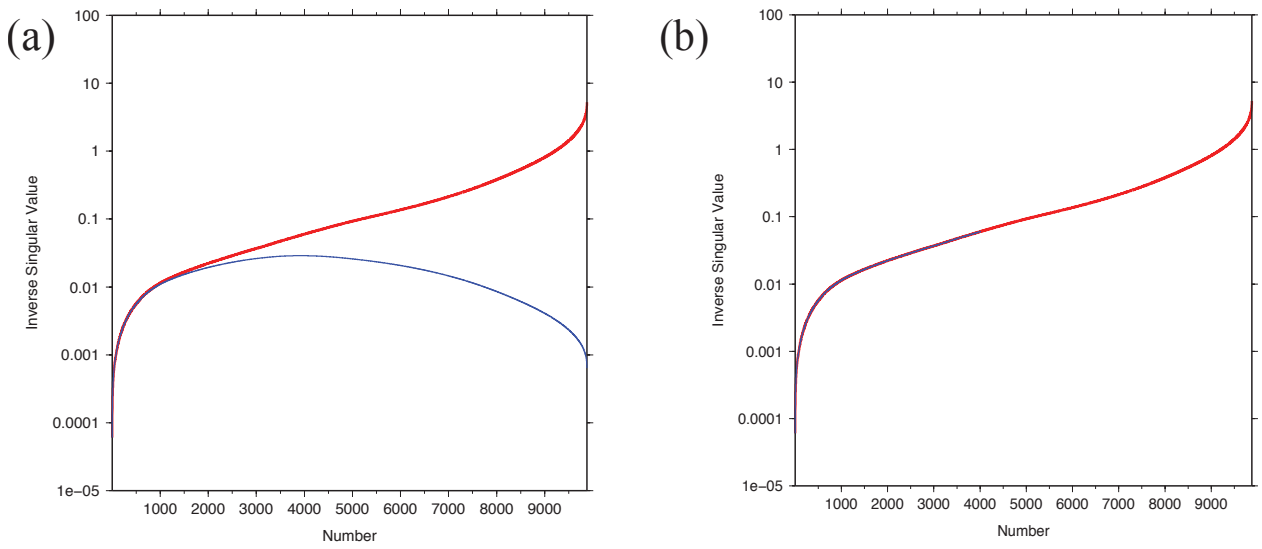

Fig.5 The inverse singular values of the Gram matrix, i.e., diag(Λ-1) based on ray theory in an ascending order shown by the red line. (a) Regularization employed by adding a damping value lamda to all the singular values, i.e., Λ/(Λ2+λ) which would change all the singular values shown by the blue line, particularly for those smaller ones. (b) Regularization employed by truncating smaller singular values below a threshold value shown by the blue line.

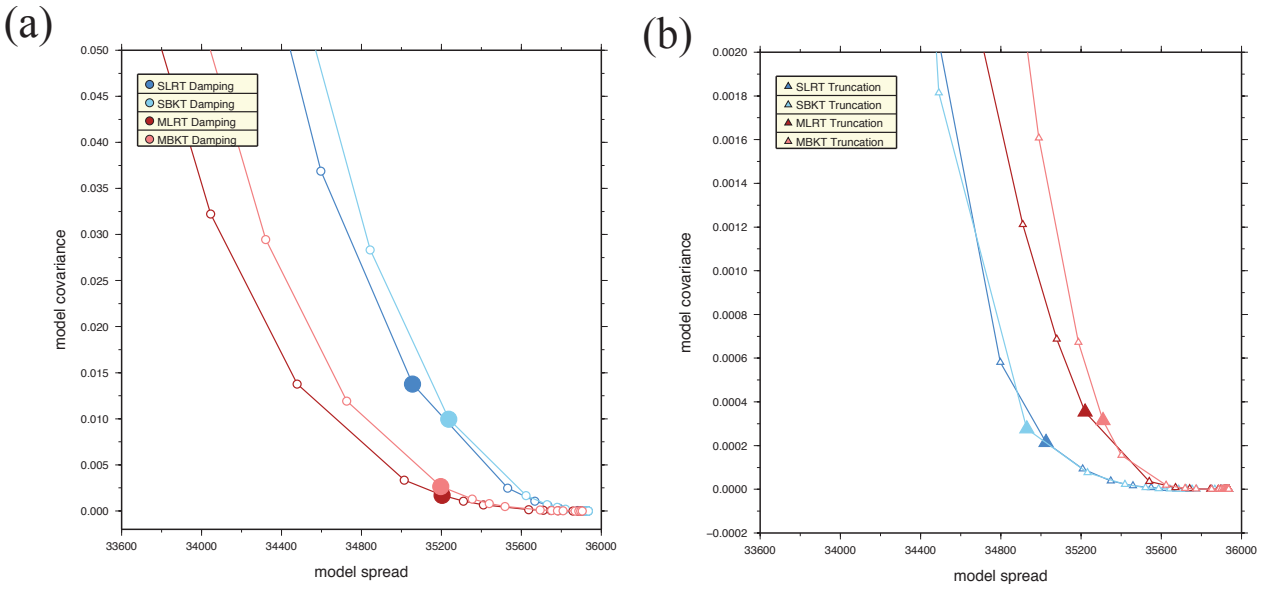

Fig. 6. Tradeoff relation between model covariance and model spread for the models obtained with singular value (a) damping and (b) truncation. The light- and dark-colored symbols represent the models based on finite-frequency theory and linearized ray theory, respectively. The red and blue colors indicate the models obtained with multi-scale and simple grid parameterizations, respectively. The optimal model solutions are chosen in the pivot regimes of the curves denoted by solid circles and triangles.

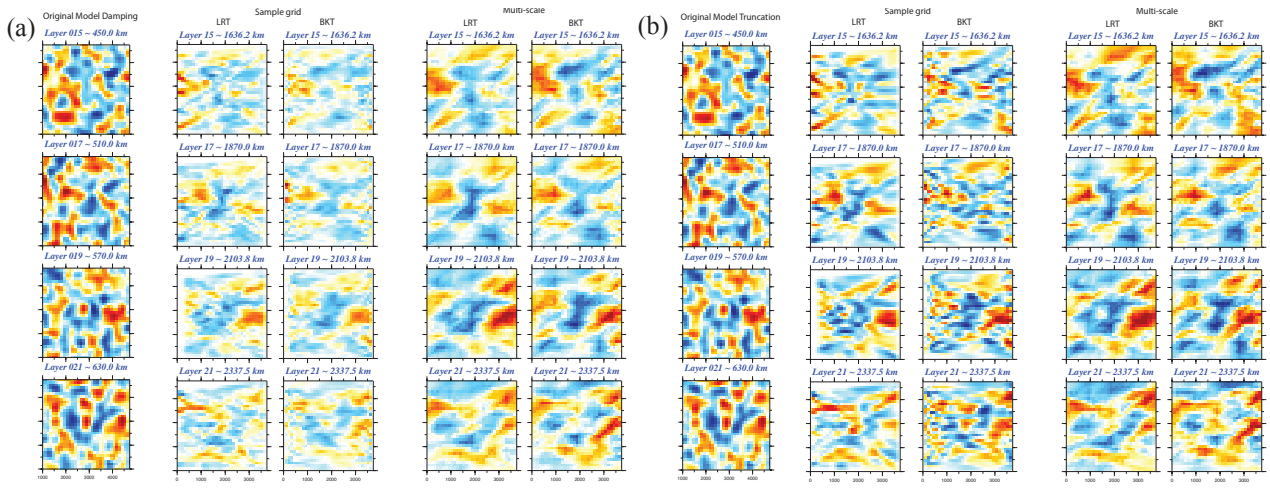

Fig.7. Four constant-depth slices through the resulting 3-D velocity structures in the optimal (a) damping and (b) truncation models.

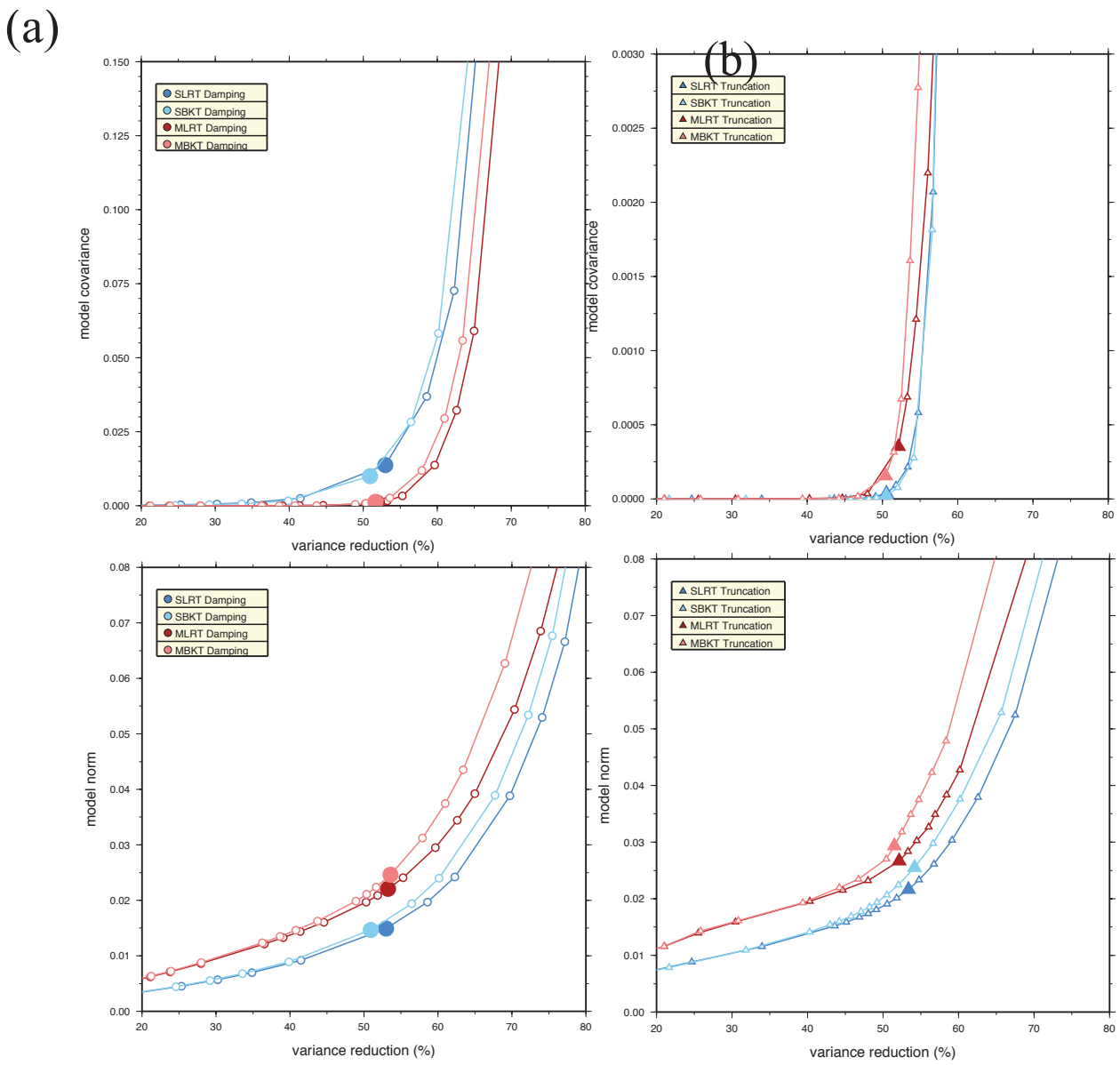

Fig.8. Comparision of the tradeoffs between model covariance and variane reduction shown on the left and between model norm and variance reduction shown on the right for the models obtained with (a) damping and (b) truncation.

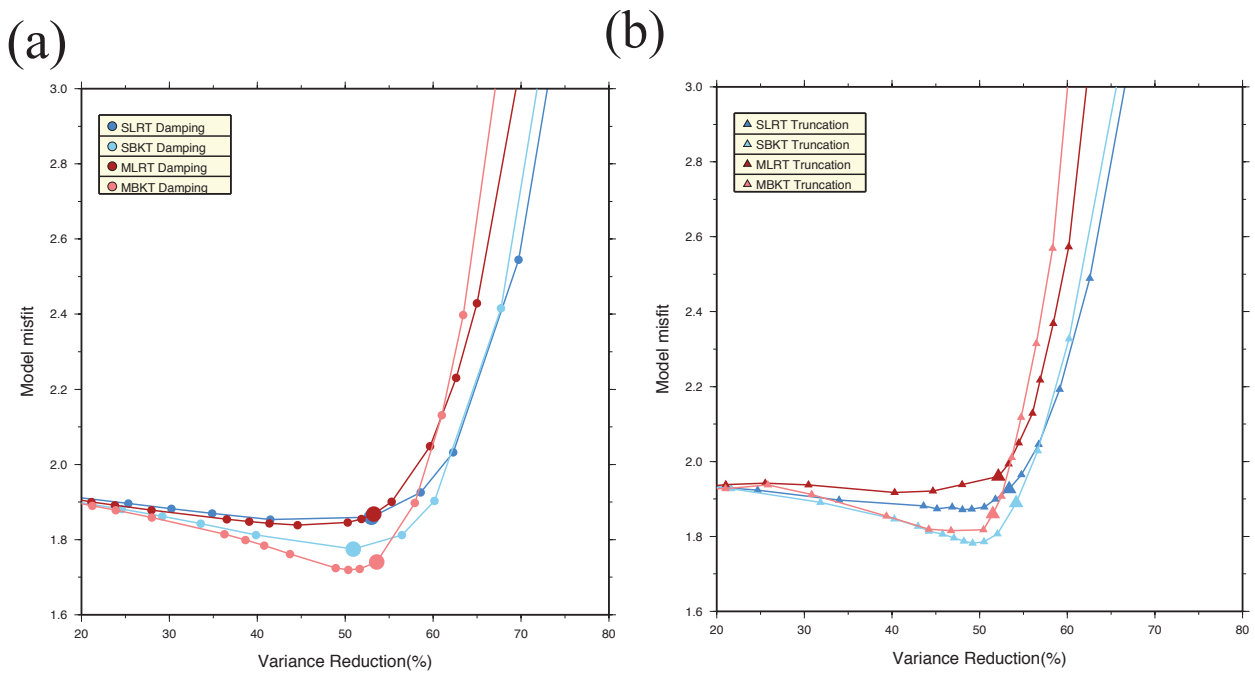

Fig. 9. Comparision of the model misfits obtained with (a) damping and (b) truncation regularization. On the left shows the model misfits varying with variance reduction and on the right are those varying with model covariance.