

Time-dependent seismic tomography

B. R. Julian¹ and G. R. Foulger²

¹*U. S. Geological Survey, 345 Middlefield Rd., MS 977, Menlo Park, CA, USA. E-mail: bruce.r.julian@gmail.com*

²*Department of Geological Sciences, Durham University, Durham DH1 3LE, UK*

Accepted 2010 May 17. Received 2010 May 17; in original form 2009 September 11

SUMMARY

Of methods for measuring temporal changes in seismic-wave speeds in the Earth, seismic tomography is among those that offer the highest spatial resolution. 3-D tomographic methods are commonly applied in this context by inverting seismic wave arrival time data sets from different epochs independently and assuming that differences in the derived structures represent real temporal variations. This assumption is dangerous because the results of independent inversions would differ even if the structure in the Earth did not change, due to observational errors and differences in the seismic ray distributions. The latter effect may be especially severe when data sets include earthquake swarms or aftershock sequences, and may produce the appearance of correlation between structural changes and seismicity when the wave speeds are actually temporally invariant. A better approach, which makes it possible to assess what changes are truly required by the data, is to invert multiple data sets simultaneously, minimizing the difference between models for different epochs as well as the rms arrival-time residuals. This problem leads, in the case of two epochs, to a system of normal equations whose order is twice as great as for a single epoch. The direct solution of this system would require twice as much memory and four times as much computational effort as would independent inversions. We present an algorithm, *tomo4d*, that takes advantage of the structure and sparseness of the system to obtain the solution with essentially no more effort than independent inversions require.

Key words: Inverse theory; Hydrothermal systems; Seismic tomography; Crustal structure; Volcano monitoring.

1 INTRODUCTION

A variety of natural and anthropogenic processes can cause the seismic wave speeds in the Earth to vary with time. Since the late 1960s, temporal changes in wave speeds have been detected in many geological environments and attributed to a variety of causes, including tidal strains (De Fazio *et al.* 1973), seasonal variations in groundwater hydrology (Sens-Schönfelder & Wegler 2006), tectonic stress build up (Furumoto *et al.* 2001), static stress changes caused by earthquakes, dynamic shaking caused by earthquakes, (Rubinstein & Beroza 2004, 2005; Peng & Ben-Zion 2006; Sawazaki *et al.* 2006; Wu *et al.* 2009) damage to rocks caused directly by faulting (Li *et al.* 2007), stress changes caused by deformation around volcanoes (Ratdomopurbo & Poupinet 1995; Nishimura *et al.* 2000), migration of magmatic fluids accompanying volcanic activity (Foulger *et al.* 2003; Patane *et al.* 2006), dilatancy and fluid migration before earthquakes (Aggarwal *et al.* 1973), drying of clay minerals caused by geothermal exploitation (Boitnott & Boyd 1996), CO₂ flooding of hydrocarbon reservoirs (Wang *et al.* 1998; Daley *et al.* 2007), and pore-pressure decreases in exploited geothermal reservoirs (Julian *et al.* 1998; Gunasekera *et al.* 2003). Such observations of temporal variations of wave speeds have important applications to fields such as earthquake prediction, volcano monitoring

(Foulger 2006), geothermal-reservoir exploitation, oil and gas reservoir assessment, and CO₂ sequestration.

Technical advances, particularly the increasing spatial density of seismometer networks and the transition from analogue to digital recording, have made available a broad range of analysis techniques and have increased by more than an order of magnitude the sensitivity with which wave-speed changes can be detected. These methods include using fixed vibrator and explosion sources (Clymer & McEvelly 1981), measuring the time differences between the two quasi-shear waves ('split shear waves') that result from propagation through anisotropic media (Crampin *et al.* 1990), and comparison of scattered waves in the codas of repeating local earthquakes having similar locations and mechanisms (e.g. Schaff & Beroza 2004; Taira *et al.* 2008; Zhao & Peng 2009). Most of the available techniques, however, have limited spatial resolution. The only technique that offers high spatial resolution is local-source seismic tomography.

Tomographic investigations of temporal changes in Earth structure have until now been conducted using conventional tomography programs such as those of the SIMUL family (Thurber 1993; Evans *et al.* 1994) to invert seismic-wave arrival time data sets for different epochs separately, assuming that differences in the resulting models arise from real temporal variations. Foulger *et al.* (1997), for example, using this method found decreases of about

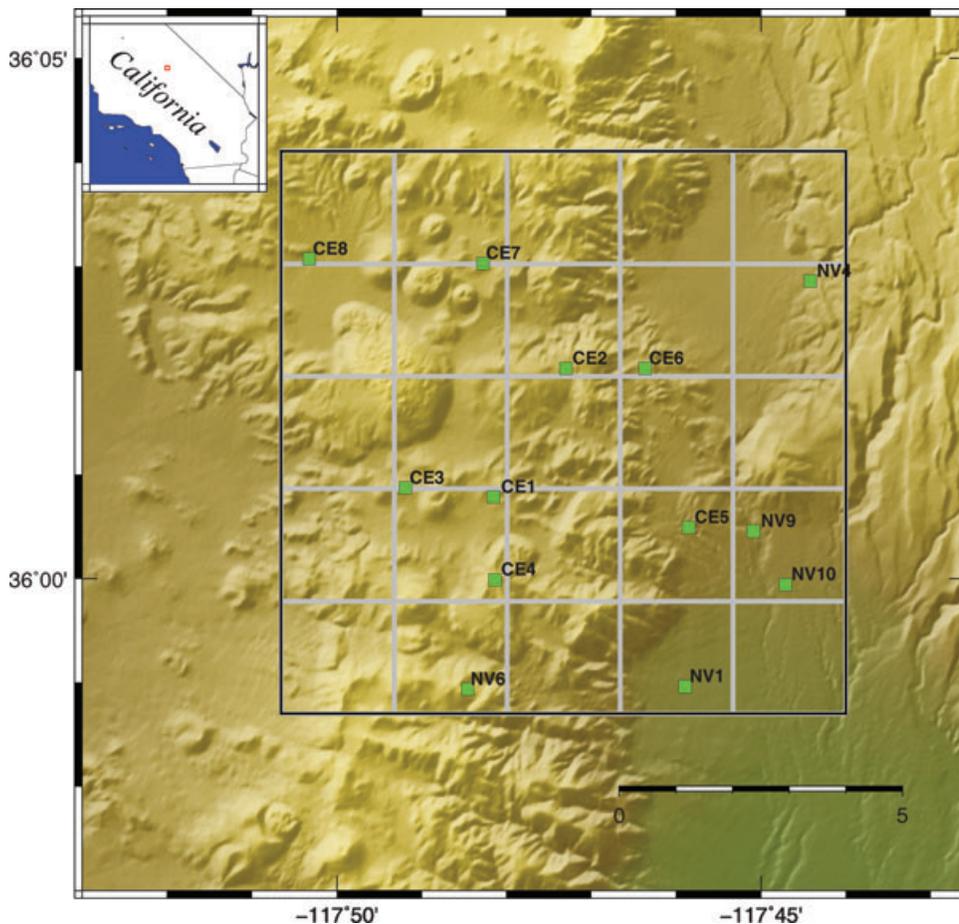


Figure 1. Map of the Coso geothermal area, eastern California, which is used as the basis for the synthetic data sets inverted in Section 3. Green squares: Stations of the U. S. Navy's permanent telemetered seismometer network. The tomographic grid has nodes spaced by 2 km horizontally and 1 km vertically, and is geometrically congruent with the grids used in Figs 3–11.

4 per cent in the ratio V_P/V_S between 1991 and 1994 at The Geysers geothermal reservoir in northern California, and Gunasekera *et al.* (2003) confirmed the change by studying data from 1993, 1996 and 1998, and showed that the change was caused primarily by a decrease in V_P , an expected consequence of an increase in pore-fluid compressibility caused by decreasing pressure in the reservoir. Using the same method, Patane *et al.* (2006) found that in certain places at Mt Etna, V_P/V_S increased by about 4.5 per cent at the time of the 2002–2003 eruption. The assumption that differences in the results of independent tomographic inversions represent true temporal variations is questionable, however, because the results of repeated tomography experiments would differ even if the structure did not change, because of variations in the seismic ray distributions caused by natural variation in earthquake locations. Even if the source locations did not change (if explosion data were used, for example) and the seismometer distribution were held fixed, differences in the derived models would be expected because of random observational errors. The reality of the temporal changes mentioned above is not subject to serious doubt, because of their large magnitude and their correlation with likely causes in the forms of intensive geothermal exploitation and volcanic activity, but some weaker reported changes presently remain open to question. Among these are possible changes between 1989 and 1997 in V_P/V_S associated with CO_2 emissions at Mammoth Mountain in Long Valley caldera, California (Foulger *et al.* 2003) and

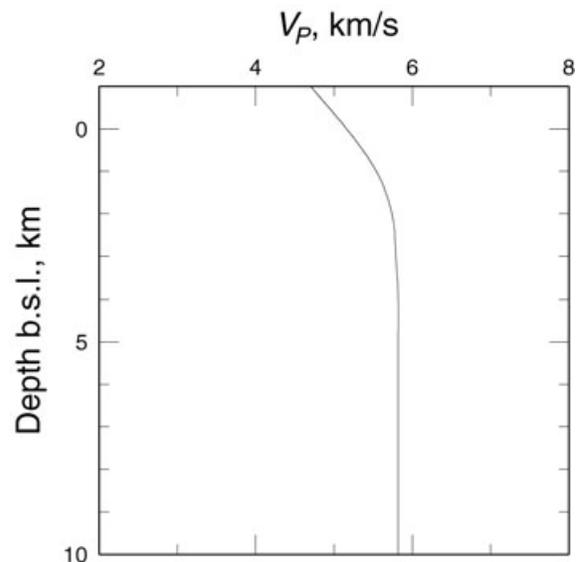


Figure 2. The depth-dependent distribution of the compressional-wave speed V_P that is used as the deterministic component of the pseudo-random models for which theoretical data are computed and inverted in Section 3. The increase of V_P with depth is based on the layered model used routinely by the U.S. Navy to locate earthquakes at the Coso geothermal area, California.

possible recent changes at the Coso geothermal area, California. Some studies, such as that by Foulger *et al.* (2003) of Long Valley caldera, have attempted to deal with this sampling problem by performing a series of inversions, alternating between two epochs and using the model derived in each inversion as the starting model in the next inversion of data from the other epoch, but this procedure

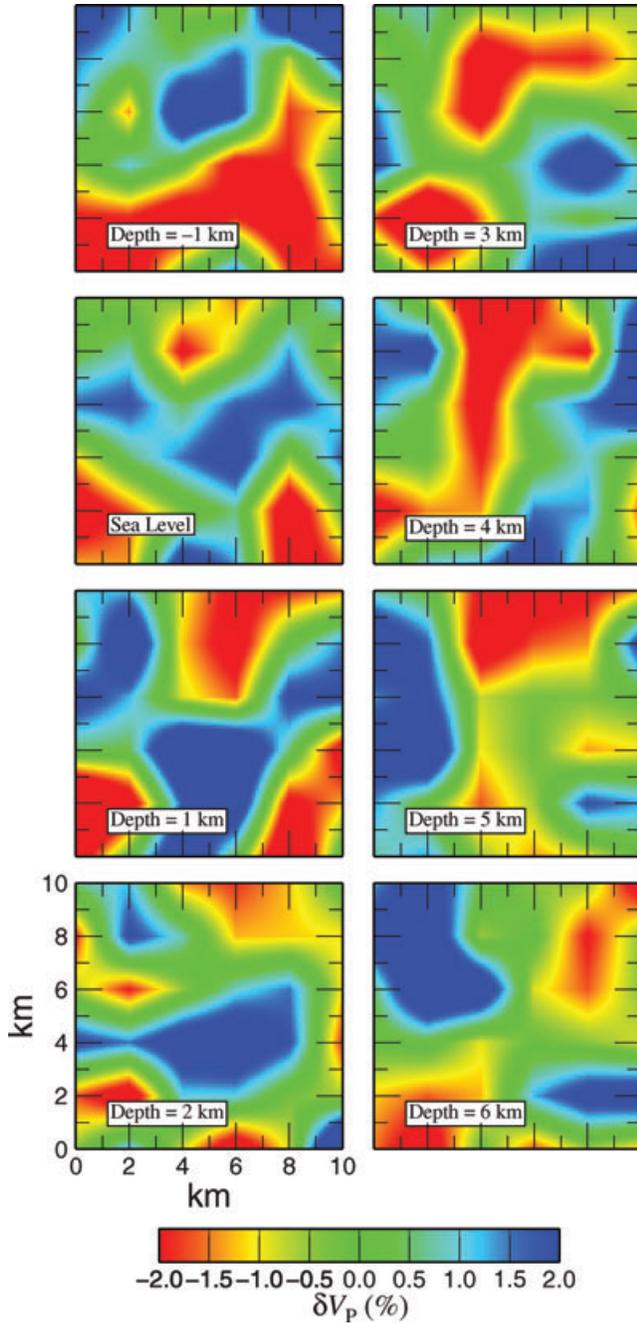


Figure 3. Maps showing spatial variations in the compressional wave speed V_P at different depths in the random time-independent model for which theoretical travel times are inverted in Section 3.1. The deterministic component of the model is the depth-dependent function shown in Fig. 2. The covariance matrix of the random component of V_P is given by eq. (1) with a standard deviation $\sigma_v = 0.1 \text{ km s}^{-1}$ and correlation distances $a_x = a_y = 2 \text{ km}$ and $a_z = 1 \text{ km}$. The grid nodes are spaced 2 km apart horizontally and 1 km apart vertically. The values shown are percent deviations from the mean at each depth.

is awkward and time-consuming, and furthermore offers no help in distinguishing between true temporal variations and the effects of random observational errors.

2 METHOD

We propose an alternative approach, which makes it possible to determine what changes are truly required by the data, namely to invert multiple data sets simultaneously, seeking to minimize the difference between the models for different epochs as well as the misfit between the observed and predicted arrival times. This problem is similar to that of seeking models consistent with an *a priori* assumed model, and Appendix A shows how it can be solved using a technique similar to the ‘damped least squares’ method of Marquardt (1963). Solving for two models simultaneously requires the determination of twice as many parameter values as solving for a single model does, and the most straightforward solution method would require substantially more than twice the computational labour. Solving the problem by means of a system of ‘normal equations’ (Whittaker & Robinson 1967, section 108), for example, requires eight times as much labour as solving for a single epoch, and thus quadruples the labour compared to solving for two epochs independently. In most cases the system of normal equations for the two-epoch problem is sparse, however, and Appendix A presents a method (Algorithm 1) that takes advantage of this fact to solve the equations with little more labour than is needed to solve for each epoch independently.

Section 3 below shows results of inverting synthetic data sets using *tomo4d*, a local-earthquake tomography computer program that uses the methods presented in Appendix A, including Algorithm 1, to invert arrival-time measurements from two epochs

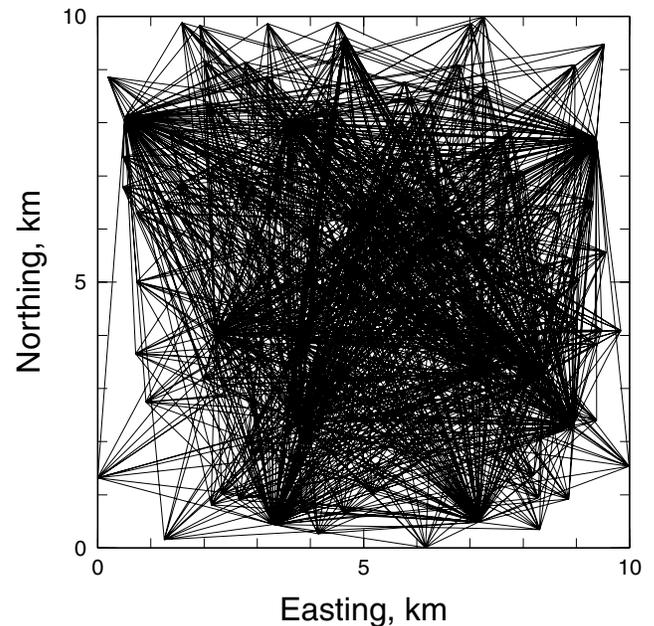


Figure 4. Surface projections of ray paths (represented as straight lines) for a sparse pseudo-random theoretical arrival-time data set of the kind inverted in Section 3.1.1, intended to represent spatially uniform background seismicity. In this example, rays connect each of 100 pseudo-random hypocenter locations with the 13 seismometer locations of the permanent U.S. Navy seismometer network at the Coso geothermal area, shown in Fig. 1. The earthquake hypocenter locations have a uniform probability distribution over the $10 \text{ km} \times 10 \text{ km} \times 10 \text{ km}$ volume.

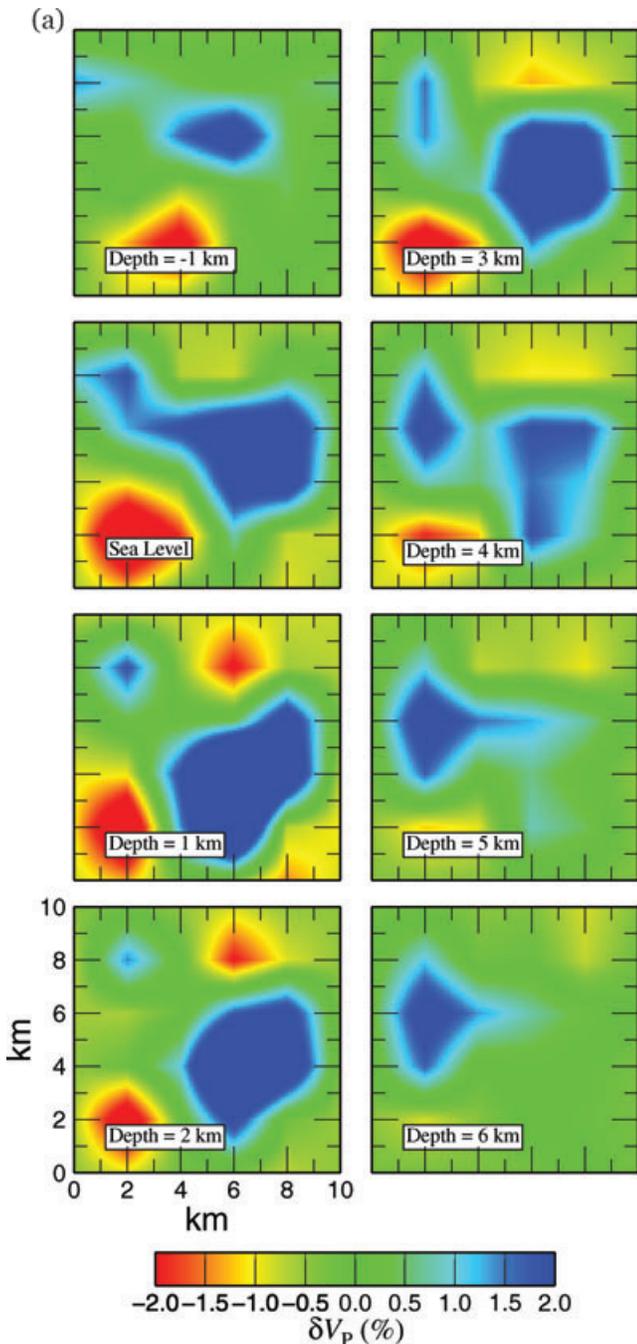


Figure 5. Maps for different depths showing the results (a and b) of independent inversions of a pair of theoretical arrival-time data sets intended to represent spatially uniform background seismicity, as described in Section 3.1.1. The data sets have a sparse statistically uniform seismicity distribution like that shown in Fig. 4 (100 events and 1300 rays per data set). The models exhibit differences because of different spatial sampling of the 3-D structure, which could be mistaken for temporal variations.

to estimate the hypocenter locations and the wave-speed changes between epochs. It uses the azimuthal-equidistant Earth-flattening approximation (Julian *et al.* 2000) to map a small region of a spherical Earth into a local Cartesian coordinate system. Many programs in common use simply identify longitude and latitude with local Cartesian coordinates x and y , an approximation that is seriously inaccurate at high latitudes or for large regions. The

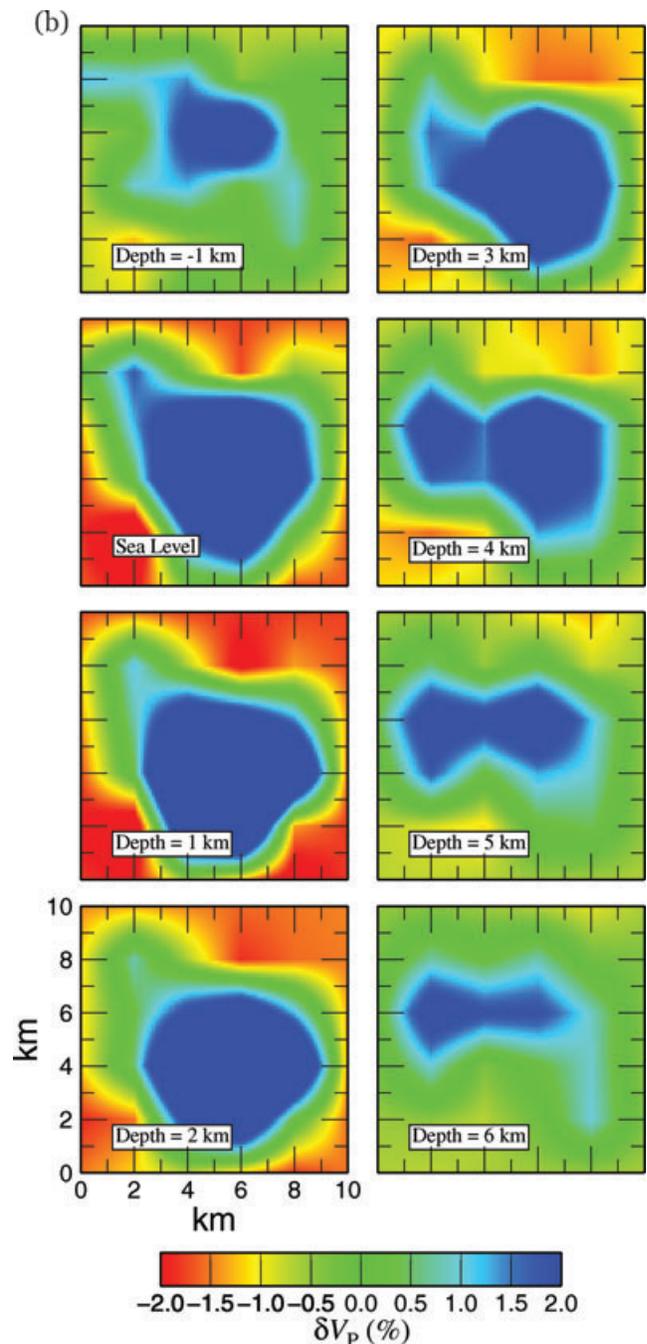


Figure 5. (Continued.)

Earth-flattening method used in *tomo4d* is invariant with respect to geographic location and is accurate over much larger regions. Models are specified by the values of seismic wave slowness (inverse of wave speed) at the nodes of a rectilinear grid, with 3-D tricubic interpolation (Press *et al.* 2007, Section 3.6) used to compute slowness values and their spatial derivatives elsewhere. The interpolated wave slownesses (and the wave speeds) are smooth, that is, they are continuous and have continuous first spatial derivatives. Ray paths are computed using the bending method of Julian & Gubbins (1977). The problems of simultaneously determining hypocenter locations and Earth structure are separated using the method of Spencer & Gubbins (1980), which requires the solution

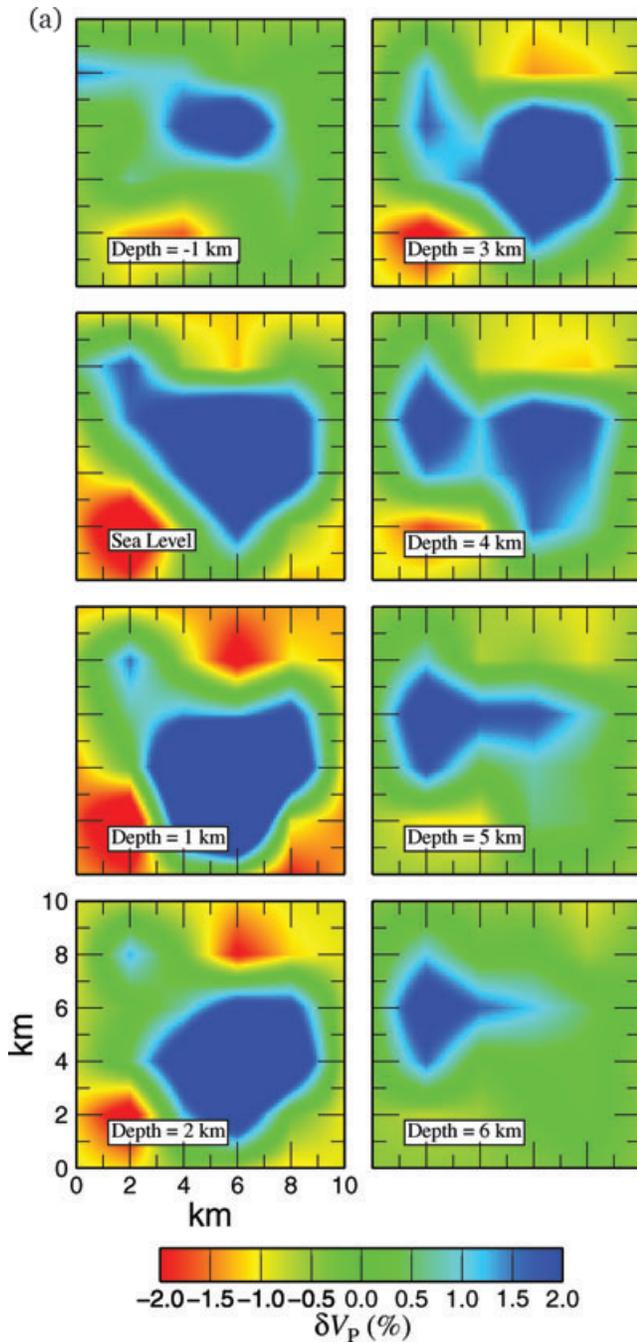


Figure 6. Maps like those of Fig. 5 showing the results of inverting the same theoretical data sets, but imposing a weak constraint to minimize differences that are not truly required by the data, as described in Section 3.1.1. Artificial differences between the results for different epochs are practically eliminated.

of only $P \times P$ and 4×4 linear systems, where P is the number of adjustable parameters in the model, and is much more efficient than solving the much larger unseparated system of equations.

3 TESTS WITH SYNTHETIC DATA

To test and illustrate the behaviour of this method, we compute theoretical arrival-time data sets for pseudo-random local Earth models

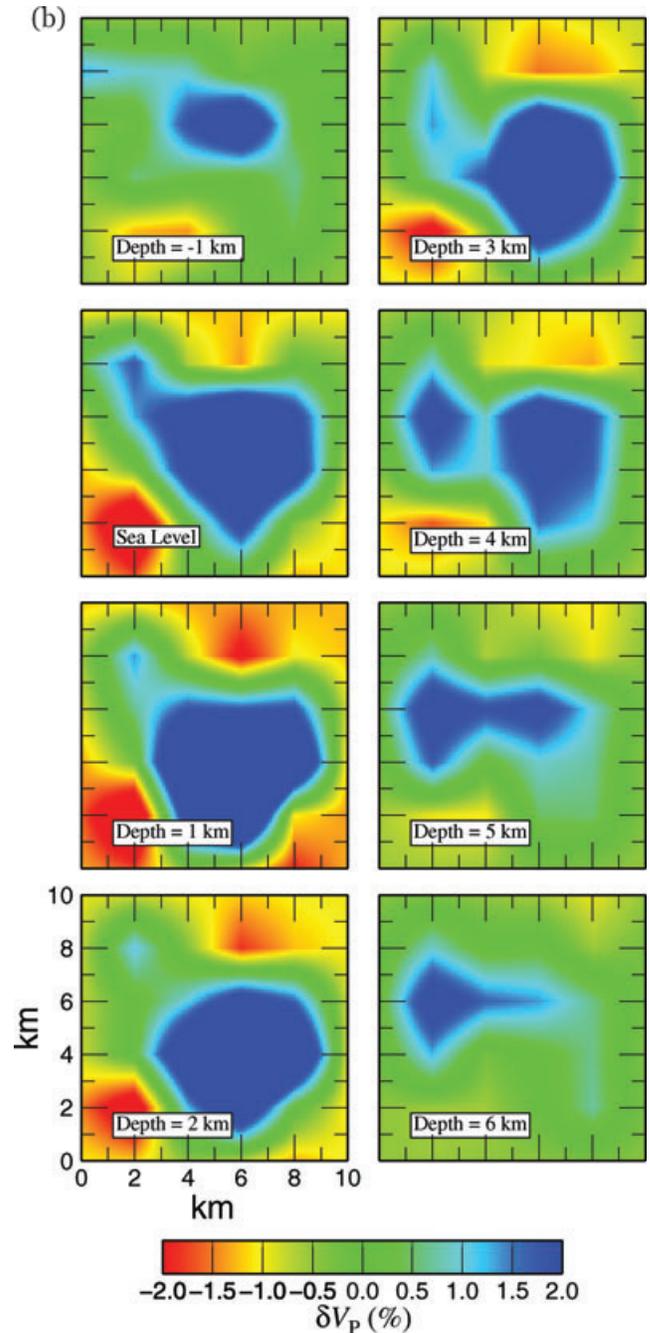


Figure 6. (Continued.)

and hypocenter locations and an actual seismometer distribution (Fig. 1), and invert these data sets both independently and using Algorithm 1 of Appendix A. In these models, the seismic wave speeds at the grid nodes are pseudo-random normal deviates computed using the method in section 7.4 of Press *et al.* (2007), having mean values ('deterministic components') that may be specified in any desired manner (here, they are purely depth-dependent, and are shown in Fig. 2) and covariances given by Gaussian functions of the form

$$\Sigma_{ij} = \sigma_v^2 \exp \left(- \left[\frac{(x_i - x_j)^2}{a_x^2} + \frac{(y_i - y_j)^2}{a_y^2} + \frac{(z_i - z_j)^2}{a_z^2} \right] \right). \quad (1)$$

Here (x_i, y_i, z_i) are the Cartesian coordinates of the i th grid node, σ_V is the standard deviation of the random variations, and a_x, a_y and a_z are correlation distances in the x, y and z directions.

We perform this exercise using both time-independent and temporally varying earth models and different sets of assumed earthquake locations, which give both dense and sparse distributions of rays and both random and systematic differences in their sampling of the structure with time.

3.1 Apparent temporal variations caused by changes in ray distribution

Because of natural changes in the locations of earthquakes, repeated tomographic inversions would be expected to give different results even if the wave speed in the Earth did not vary with time. To investigate this phenomenon, we invert theoretical P -wave arrival-time data sets generated by tracing rays through the pseudo-random model shown in Fig. 3. We investigate two types of situation. Statistically uniform spatial distributions of earthquakes, such as might represent natural background seismicity, and strongly clustered earthquakes, such as might occur in an earthquake swarm or aftershock sequence. In all cases we use a real (non-random) distribution of seismometers, namely the innermost 13 stations of the U.S. Navy’s permanent network at the Coso geothermal area, California (see Fig. 1).

In order to reduce potentially complicating effects, we made several simplifications of reality. We inverted exact theoretical arrival times, without attempting to simulate observational errors, and we did not attempt to solve for hypocenter locations, but rather used the true locations as inputs to the inversion process. In reality, observational errors and uncertainty about hypocentral locations would complicate the task of identifying artefacts such as spuri-

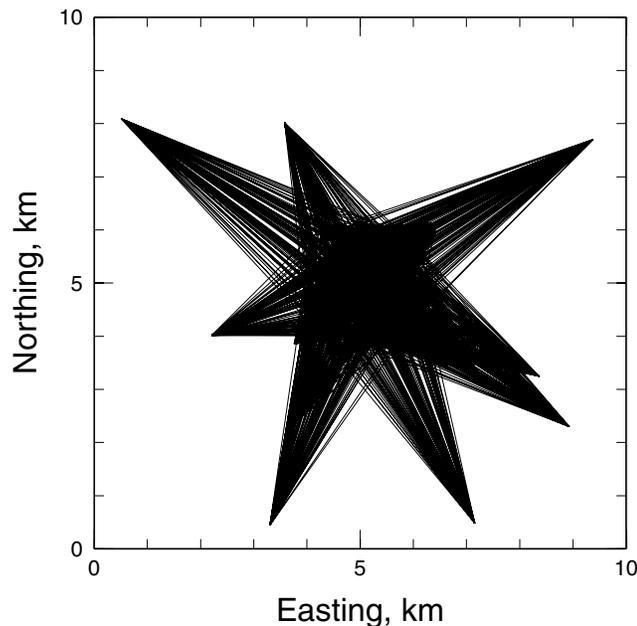


Figure 7. Surface projections of computed ray paths (represented as straight lines) for a pseudo-random theoretical arrival-time data set of the kind inverted in Section 3.1.2, intended to represent data from an earthquake swarm or aftershock sequence. Rays connect each of 100 pseudo-random hypocenter locations, uniformly distributed in a $2\text{ km} \times 2\text{ km} \times 2\text{ km}$ cube centered at $x = y = z = 5\text{ km}$, to each seismometer.

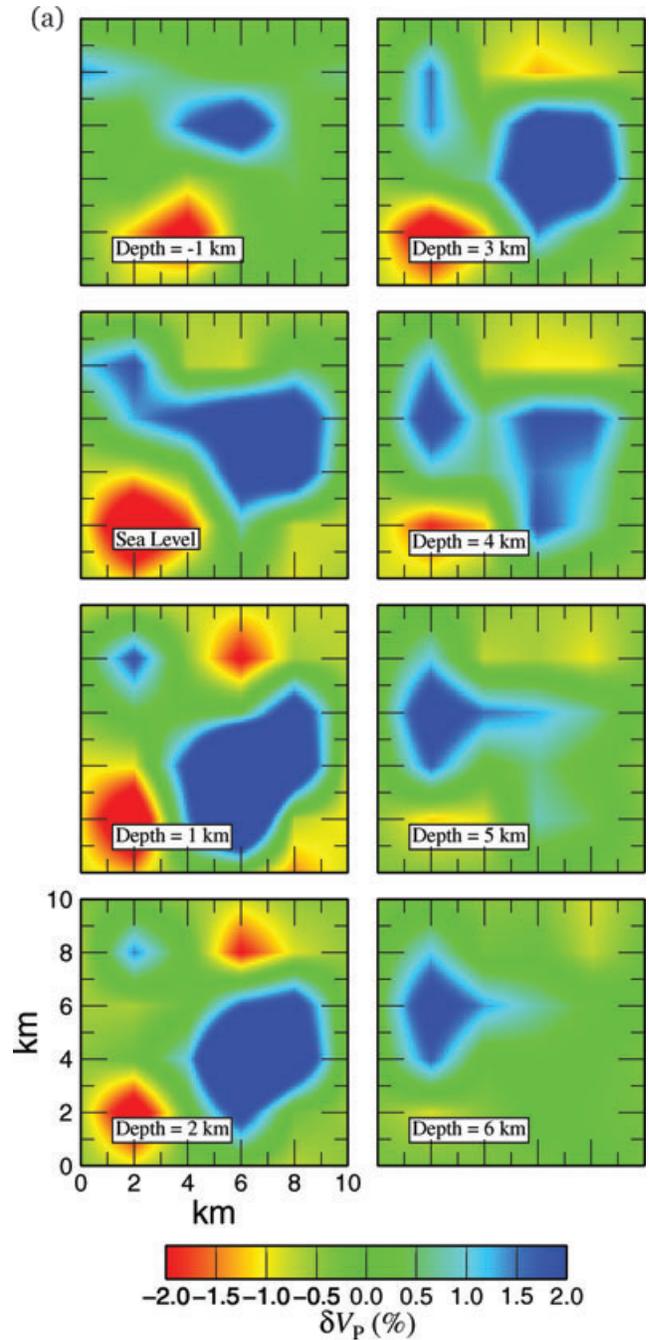


Figure 8. Maps for different depths showing the results of independent inversions of a pair of theoretical arrival-time data sets intended to represent the onset of an earthquake swarm or aftershock sequence, as described in Section 3.1.2 (a) The result derived from the earthquake distribution shown in Fig. 4, representing pre-main shock background seismicity; (b) the result derived from the earthquake distribution shown in Fig. 7, representing localized swarm seismicity. The systematic difference in the sampling of the structure by the two ray distributions produces strong differences in the results, especially in the seismically active volume, which could easily be mistaken for temporal variations in the wave speed.

ous temporal variations in wave speeds. We inverted theoretical arrival times for seismic P phases only, and thus solved for only the compressional-wave speed V_P . In reality, the shear wave speed V_S is of great intrinsic interest, and S -phase times are particularly valuable for constraining hypocentral locations. To minimize non-linear

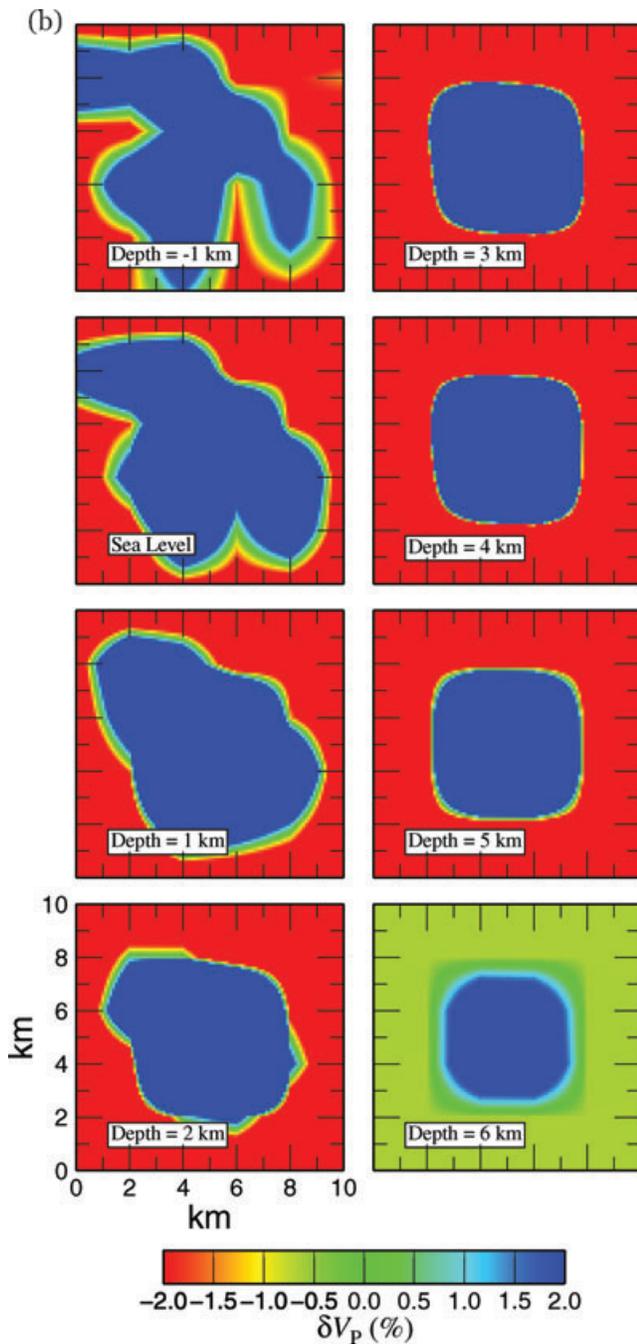


Figure 8. (Continued.)

effects, we used a pseudo-random model (Fig. 3) with comparatively weak lateral variations in wave speed (about 2 per cent) and performed only a single tomographic iteration. This restriction, combined with the use of step-size damping, means that our derived models resolve only the strongest of the pseudo-random variations.

3.1.1 Random variations in ray distribution

Fig. 4 shows an example of ray paths for a sparse, statistically uniform distribution of random hypocenter locations intended to represent unclustered background seismicity. We inverted pairs of theoretical data sets computed for such random rays using two approaches: independent inversions and joint inversions using the

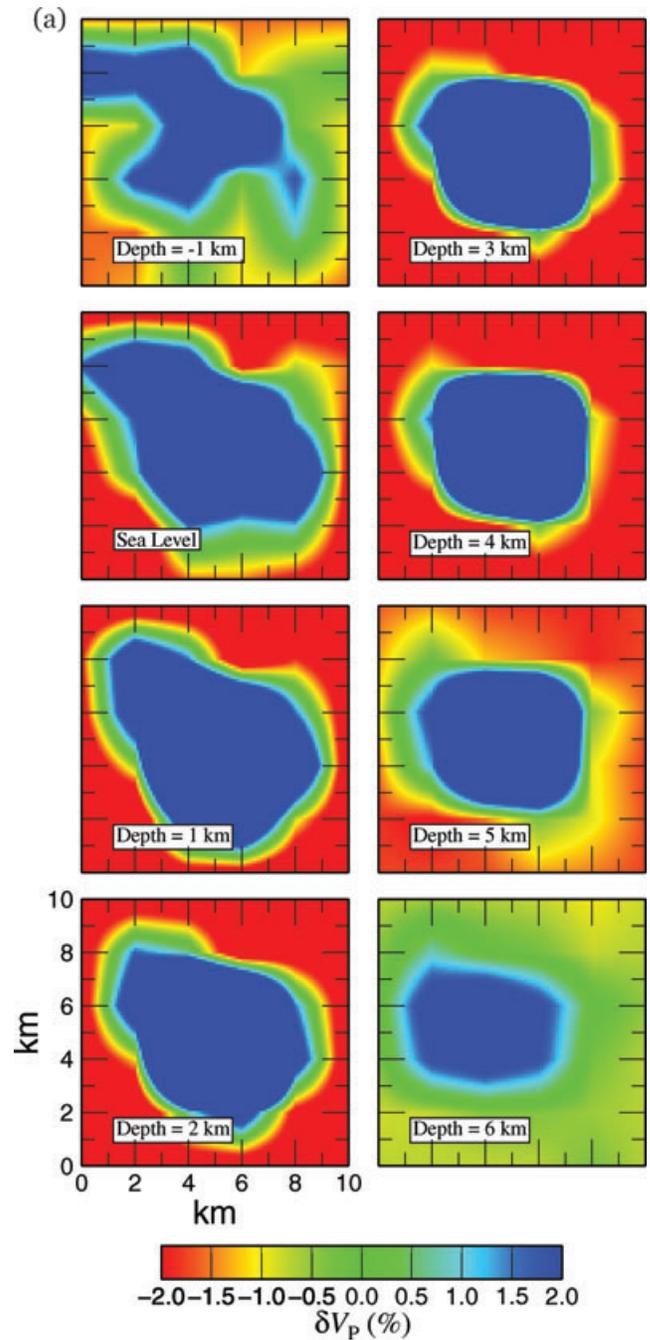


Figure 9. Maps like those of Fig. 8 showing results obtained by inverting the same theoretical data sets, but imposing a weak constraint to minimize differences that are not truly required by the data. Because of the highly non-uniform ray distribution, the results differ significantly from ones such as those shown in Fig. 6, but the spurious temporal variations that dominated Fig. 8 have almost completely disappeared.

method of Appendix A to suppress differences between derived models that are not actually required by the data.

Fig. 5 shows the results of inverting a pair of data sets independently, and Fig. 6 shows the results of inverting them jointly using *tomo4d*. All these results exhibit several large-scale anomalies, most of which correspond well to features in the pseudo-random model used to generate the theoretical data (Fig. 3), although the inevitable undersampling of the structure by rays causes many other features of the models not to be imaged well. The results for the different

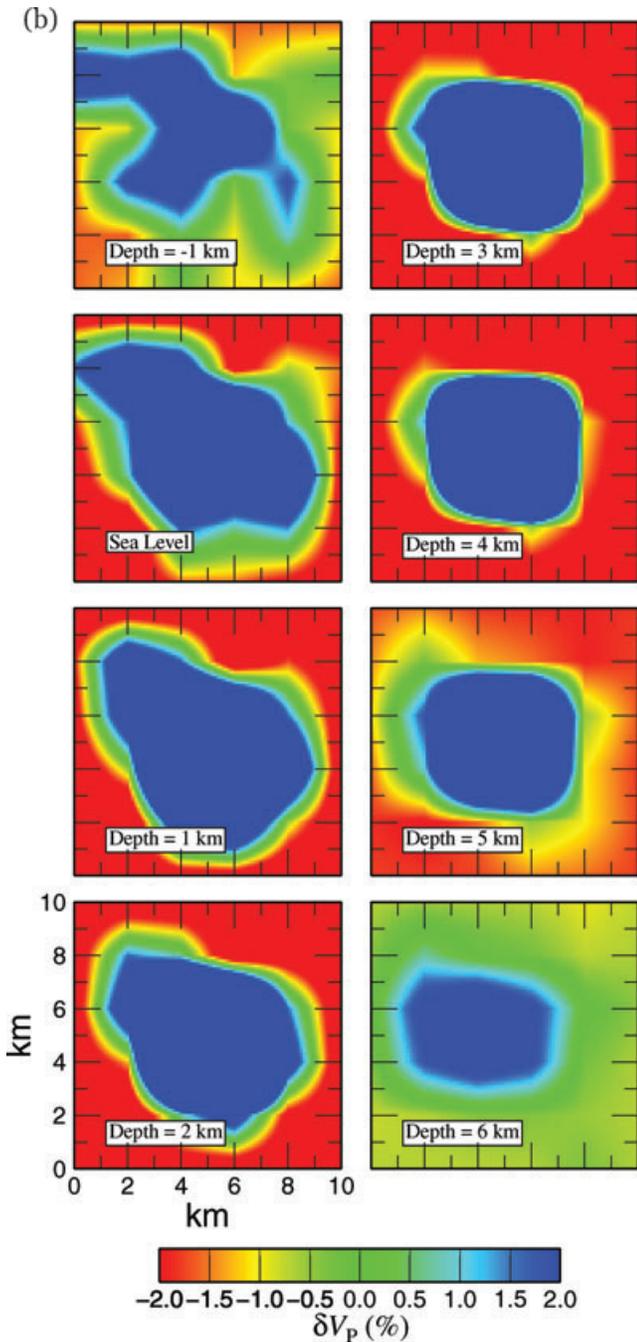


Figure 9. (Continued.)

epochs, though, have many features that differ significantly. When a mild constraint is applied to suppress differences not required by the data, all these artificial temporal changes disappear (Fig. 6).

3.1.2 Systematic aftershock-like variations in ray distribution

In reality, temporal changes in earthquake locations are often more systematic than those considered in Section 3.1.1. An extreme but important example is afforded by earthquake swarms and aftershock sequences, which often produce tens of thousands of arrival-time data densely sampling a volume that previously was sampled sparsely or not at all.

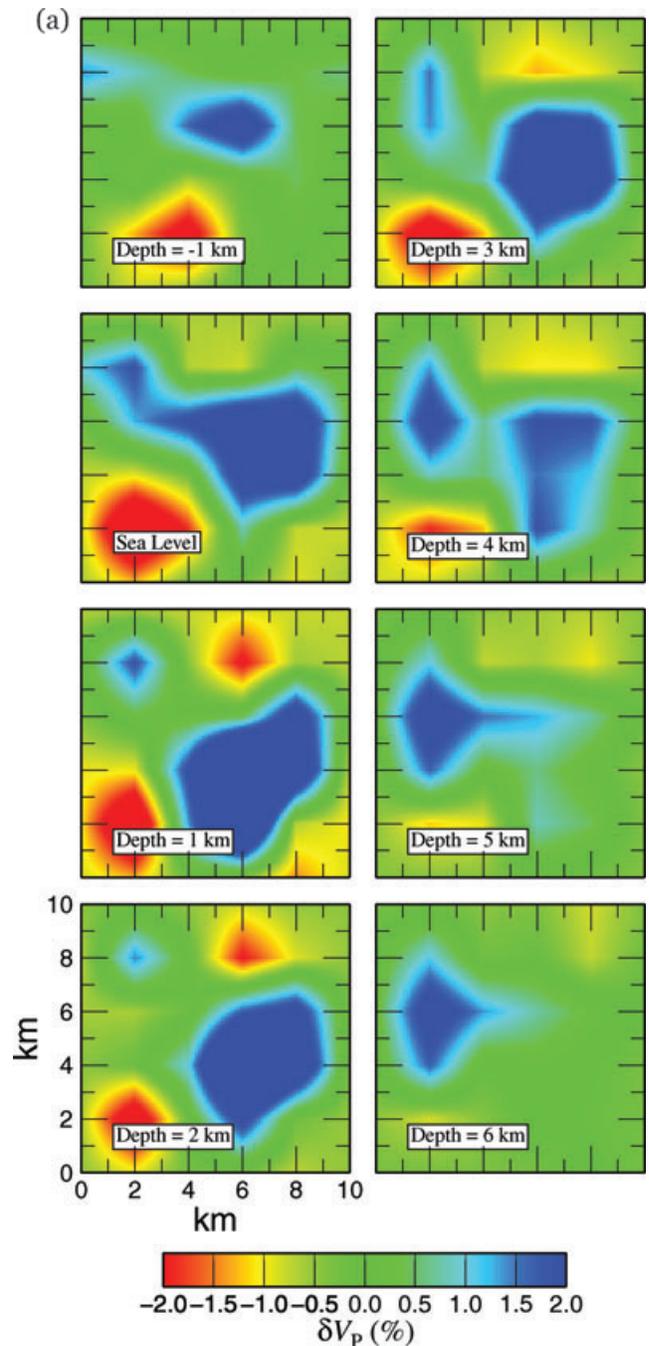


Figure 10. Maps showing wave speeds obtained by independently inverting synthetic data sets computed for a weak, localized temporal change in wave speed. The model in (a) is derived from the same data set as that used for Figs 5(a) and 8(a). That in (b) is based on the same pseudo-random set of hypocenters as Fig. 5(b) and a model with a -0.1 km s^{-1} V_P anomaly from sea level to 1 km below sea level, north coordinate from 6 to 8 km, and east coordinate from 2 to 4 km. The artificial apparent temporal changes completely obscure all evidence of the true change.

Fig. 7 shows the ray distribution for 100 pseudo-random hypocenters intended to mimic such an earthquake sequence. The hypocenters are distributed with uniform probability density throughout a $2 \text{ km} \times 2 \text{ km} \times 2 \text{ km}$ cube located in the centre of the region. This ray distribution is quite different from ones like that shown in Fig. 4, and it samples the structure poorly. Inverting

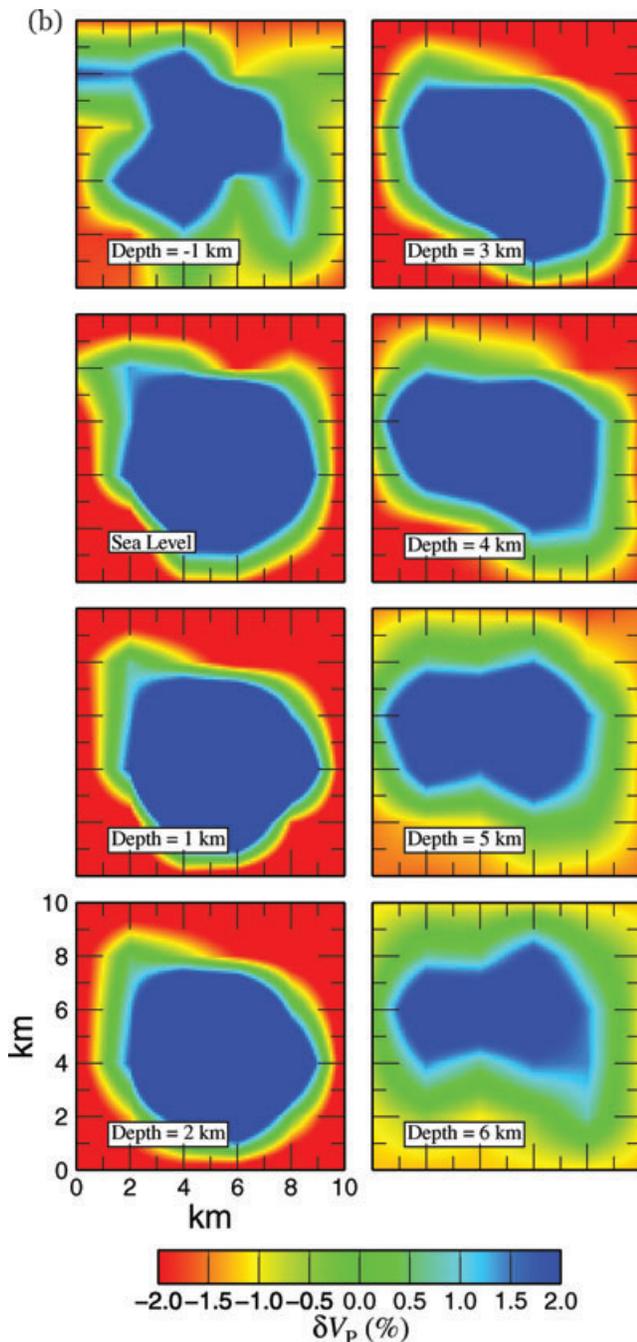


Figure 10. (Continued.)

theoretical data from it produces models with significant structural features concentrated mainly near the region of seismicity (Fig. 8b). Such models differ greatly from ones obtained from inverting a more uniform (e.g. pre-main shock) seismicity distribution (Fig. 8a) and such differences might well be mistaken for real temporal variations related to the earthquake-generation process.

Jointly inverting such an aftershock-like data set and one with uniformly distributed rays, however, eliminates the misleading apparent temporal change (Fig. 9). The strong concentration of rays in the neighbourhood of the aftershocks strongly biases the derived models (compare Fig. 9 with Fig. 6, for example), but it does not produce spurious temporal variations in wave speed.

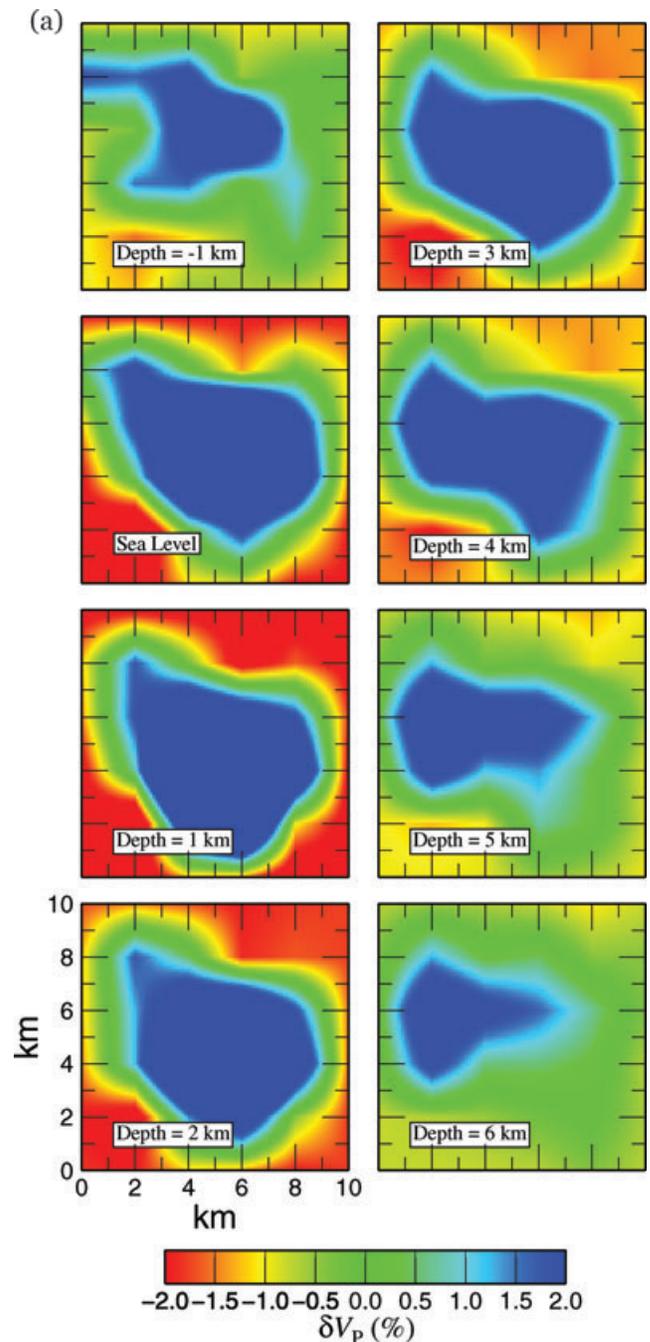


Figure 11. Maps showing wave speeds in models obtained by inverting the same data sets as in Fig. 10, but using the *tomo4d* algorithm to suppress artificial apparent temporal changes. The change in wave speed near sea level in the northwest, which involves values at the eight nodes shown as white dots, is now clear.

3.2 True temporal variations in wave speeds

To demonstrate that the *tomo4d* algorithm not only suppresses artificial temporal variations in seismic-wave speed but also can reveal true temporal variations, we conducted inversions like those described in Section 3.1.1, but using synthetic data generated using models that vary slightly between epochs. We use uniform ‘background seismicity’ distributions like the example shown in Fig. 4. The structure for the first epoch is the same as the one used in all the previous examples, but for the second epoch we introduce a

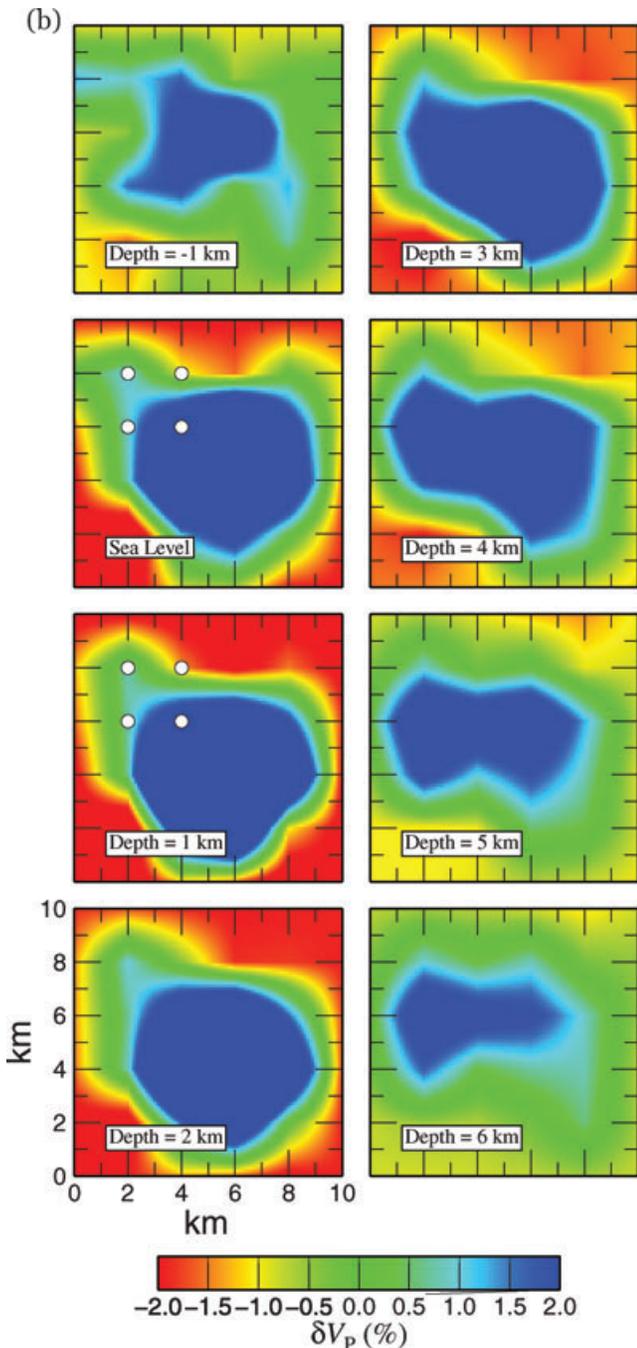


Figure 11. (Continued.)

0.1 km s^{-1} negative V_P anomaly in a $2 \times 2 \times 1 \text{ km}$ region just below sea level in the northwestern part of the model.

Fig. 10 shows the results of independent inversions. As in the case shown in Fig. 5, the difference in the sampling of the structure provided by the two pseudo-random hypocenter distributions produces differences between the derived structures that completely obscure the temporal change in the northwest. Fig. 11 shows the results of applying the *tomo4d* algorithm. The results for the two epochs are now quite similar except in the shallow volume in the northwest, where the negative V_P anomaly that we introduced for the second epoch is clearly detectable. Because of the inevitably imperfect sampling of the structure by the rays, the anomaly is smeared

vertically, and is visually detectable on Fig. 11 from the surface to depths of about 2 km below sea level.

Of course, seismic tomography is a non-linear inverse problem, because of the dependence of the seismic ray paths on the 3-D wave speed structure that is being sought. The only practical solution methods, including the variant described here, involve iterative refinement of initial estimated models, which might sometimes not converge or might converge to different solutions depending on the initial model chosen.

4 CONCLUSIONS

Repeat seismic tomography has the potential to detect and resolve spatially temporal changes in the seismic wave speeds in the Earth, but it is subject to bias caused by temporal variations in ray paths caused by changes in earthquake locations or seismometer-network geometry, and also by random observational errors in measured arrival times. These effects can produce spurious and potentially misleading apparent temporal changes in derived tomographic models. Artefacts of this kind can be eliminated by inverting data sets from multiple epochs simultaneously, imposing constraints to minimize interepoch differences between models, as described in Appendix A. Direct application of this method requires solution of a large system of linear equations, which is expensive in terms of both storage requirements and numerical labour. The particular structure of the equations, however, and their sparseness, make it possible to simultaneously invert data from two epochs with about the same storage requirements and computational effort as inverting the data sets independently. The algorithm for doing this (Algorithm 1, Appendix A) is equally applicable to any linear or linearized inverse problem, such as gravity, electrical, or magnetotelluric interpretation, in addition to seismic tomography.

ACKNOWLEDGMENTS

Many colleagues have helped with suggestions for the improvement of this material. We particularly thank John R. Evans of the U.S. Geological Survey for many valuable discussions of seismic tomography over the years. Jeanne Hardebeck and John R. Evans of the U.S. Geological Survey reviewed the manuscript.

REFERENCES

- Aggarwal, Y.P., Sykes, L.R., Armbruster, J. & Sbar, M.L., 1973. Premonitory changes in seismic velocities and prediction of earthquakes, *Nature*, **241**, 101–104.
- Boitnott, G.N. & Boyd, P.J., 1996. Permeability, electrical impedance, and acoustic velocities of reservoir rocks from The Geysers geothermal field, in *Twenty-first Workshop on Geothermal Reservoir Engineering*, Stanford, California, Stanford University, pp. 343–350.
- Clymer, R.W. & McEvelly, T.V., 1981. Travel-time monitoring with vibroseis, *Bull. seism. Soc. Am.*, **71**, 1903–1927.
- Crampin, S., Booth, D.C., Evans, R., Peacock, S. & Fletcher, J.B., 1990. Changes in shear-wave splitting at Anza near the time of the North Palm Springs earthquake, *J. geophys. Res.*, **95**, 11197–11212.
- Daley, T.M., Solbau, R.D., Ajo-Franklin, J.B. & Benson, S.M., 2007. Continuous active-source seismic monitoring of CO₂ injection in a brine aquifer, *Geophysics*, **72**, A57–A61.
- De Fazio, T.L., Aki, K. & Alba, J., 1973. Solid Earth tide and observed change in the in situ seismic velocity, *J. geophys. Res.*, **78**, 1319–1322.
- Evans, J.R., Eberhart-Phillips, D. & Thurber, C.H., 1994. User's manual for SIMULPS12 for imaging V_P and V_P/V_S , a derivative of the Thurber

- tomographic inversion SIMUL3 for local earthquakes and explosions, U.S. Geol. Surv. Open File Rep. 94-431, 142 pp.
- Foulger, G.R., 2006. Toward “supervolcano” technology, *Science*, **313**, 768–769.
- Foulger, G.R., Grant, C.C., Ross, A. & Julian, B.R., 1997. Industrially induced changes in Earth structure at The Geysers geothermal area, California, *Geophys. Res. Lett.*, **24**, 135–137.
- Foulger, G.R., Julian, B.R., Pitt, A.M., Hill, D.P., Malin, P.E. & Shalev, E., 2003. Three-dimensional crustal structure of Long Valley caldera, California, and evidence for the migration of CO₂ under Mammoth Mountain, *J. geophys. Res.*, **108**, 2147, doi:10.1029/2000JB000041.
- Furumoto, M., Ichimori, Y., Hayashi, N., Hiramatsu, Y. & Satoh, T., 2001. Seismic wave velocity changes and stress build-up in the crust of the Kanto-Tokai region, *Geophys. Res. Lett.*, **28**, 3737–3740.
- Gunasekera, R.C., Foulger, G.R. & Julian, B.R., 2003. Reservoir depletion at The Geysers geothermal area, California, shown by four-dimensional seismic tomography, *J. geophys. Res.*, **108**, 2134, doi:10.1029/2001JB000638.
- Julian, B.R. & Gubbins, D., 1977. Three-dimensional seismic ray tracing, *J. Geophys.*, **43**, 95–113.
- Julian, B.R., Pitt, A.M. & Foulger, G.R., 1998. Seismic image of a CO₂ reservoir beneath a seismically active volcano, *Geophys. J. Int.*, **133**, F7–F10.
- Julian, B.R., Evans, J.R., Pritchard, M.J. & Foulger, G.R., 2000. A geometrical error in some computer programs based on the Aki-Christofferson-Husebye (ACH) method of teleseismic tomography, *Bull. seism. Soc. Am.*, **90**, 1554–1558.
- Li, Y.G., Chen, P., Cochran, E.S. & Vidale, J.E., 2007. Seismic velocity variations on the San Andreas fault caused by the 2004 M6 Parkfield Earthquake and their implications, *Earth Planets Space*, **59**, 21–31.
- Marquardt, D.W., 1963. An algorithm for least-squares estimation of non-linear parameters, *J. Soc. Indust. Appl. Math.*, **11**, 431–441.
- Nishimura, T., Uchida, N., Sato, H., Ohtake, M., Tanaka, S. & Hamaguchi, H., 2000. Temporal changes of the crustal structure associated with the M6.1 earthquake on September 3, 1998, and the volcanic activity of Mount Iwate, Japan, *Geophys. Res. Lett.*, **27**, 269–272.
- Patane, D., Barberi, G., Cocina, O., De Gori, P. & Chiarabba, C., 2006. Time-resolved seismic tomography detects magma intrusions at Mount Etna, *Science*, **313**, 821–823.
- Peng, Z. & Ben-Zion, Y., 2006. Temporal changes of shallow seismic velocity around the Karadere-Düzce branch of the north Anatolian fault and strong ground motion, *Pure Appl. Geophys.*, **163**, 567–600.
- Press, W.H., Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T., 2007. *Numerical Recipes—The Art of Scientific Computing*, 3rd edn, Cambridge University Press, Cambridge.
- Ratdomopurbo, A. & Poupinet, G., 1995. Monitoring a temporal change of seismic velocity in a volcano—application to the 1992 eruption of Mt. Merapi (Indonesia), *Geophys. Res. Lett.*, **22**, 775–778.
- Rubinstein, J.L. & Beroza, G.C., 2004. Evidence for widespread nonlinear strong ground motion in the M_W 6.9 Loma Prieta earthquake, *Bull. seism. Soc. Am.*, **94**, 1595–1608.
- Rubinstein, J.L. & Beroza, G.C., 2005. Depth constraints on nonlinear strong ground motion from the 2004 Parkfield earthquake, *Geophys. Res. Lett.*, **32**, doi:10.1029/2005GL023189.
- Sawazaki, K., Sato, H., Nakagara, H. & Nishimura, T., 2006. Temporal change in site response caused by earthquake strong motion as revealed from coda spectral ratio measurement, *Geophys. Res. Lett.*, **33**, doi:10.1029/2006GL027938.
- Schaff, D.P. & Beroza, G.C., 2004. Coseismic and postseismic velocity changes measured by repeating earthquakes, *J. geophys. Res.*, **109**, doi:10.1029/2004JB003011.
- Sens-Schönfelder, C. & Wegler, U., 2006. Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia, *Geophys. Res. Lett.*, **33**, 229–236.
- Spencer, C. & Gubbins, D., 1980. Travel-time inversion for simultaneous earthquake location and velocity structure determination in laterally varying media, *Geophys. J. R. astr. Soc.*, **63**, 95–116.
- Taira, T., Silver, P.G., Niu, F. & Nadeau, R.M., 2008. Detecting seismogenic stress evolution and constraining fault zone rheology in the San Andreas Fault following the 2004 Parkfield earthquake, *J. geophys. Res.*, **113**, doi:10.1029/2007JB005151.
- Thurber, C., 1993. Local earthquake tomography: velocities and V_P/V_S theory, in *Seismic Tomography: Theory and Practice*, pp. 563–583, eds Iyer, H.M. & Hirahara, K., Chapman and Hall, London.
- Wang, Z., Cates, M.E. & Langan, R.T., 1998. Seismic monitoring of a CO₂ flood in a carbonate reservoir: a rock physics study, *Geophysics*, **63**, 1604–1617.
- Whittaker, E. & Robinson, G., 1967. *The Calculus of Observations*, Dover Publications, New York.
- Wu, C., Peng, Z. & Ben-Zion, Y., 2009. Non-linearity and temporal changes of fault zone site response associated with strong ground motion, *Geophys. J. Int.*, **176**, 265–278.
- Zhao, P. & Peng, Z., 2009. Depth extent of damage zones around the central Calaveras fault from waveform analysis of repeating earthquakes, *Geophys. J. Int.*, **179**, 1817–1830.

APPENDIX A: INDIRECT MEASUREMENT OF SECULAR CHANGES

Suppose seismological or other indirect measurements of Earth structure are made at two epochs and we wish to estimate what changes, if any, have occurred during the intervening period. Because of uncontrollable factors, the two data sets do not correspond exactly to each other. The numbers of data in the sets are not likely to be the same, the earthquakes are generally not in the same places, and the numbers, locations, and characteristics of the seismometers may differ. Even if the two data sets do not differ in these respects (if timed explosion sources are used, for example), inverting them is expected to give somewhat different results because of random measurement errors. We seek an inversion method that will distinguish between such sampling fluctuations and true changes in Earth structure.

A1 Conventional least-squares method

Let the Earth model at an epoch be described by n parameters arranged in a column vector, let the model vectors for the two epochs be \mathbf{x} and \mathbf{y} , and let the observed arrival-time residuals (observed minus predicted times) similarly be arranged in column vectors \mathbf{b} and \mathbf{d} , of dimensions m_b and m_d . The functional relation between a model and the corresponding predicted arrival times is non-linear, because of the dependence of the ray paths on the model, and the only currently practical solution method consists of iteratively solving a linearized problem in which small changes in the models, $\delta\mathbf{x}$ and $\delta\mathbf{y}$, are related to predicted changes $\delta\mathbf{b}$ and $\delta\mathbf{d}$ in the residuals by linear operators (the first terms of Taylor-series expansions). These operators take the form of matrices \mathbf{A} and \mathbf{C} , of dimensions $m_b \times n$ and $m_d \times n$, such that, to first order, predicted changes in the residual vectors are given by the ‘design equations’

$$\delta\mathbf{b} = -\mathbf{A}\delta\mathbf{x} \text{ and } \delta\mathbf{d} = -\mathbf{C}\delta\mathbf{y}. \quad (\text{A1})$$

The conventional least-squares method [e.g. Whittaker & Robinson (1967), chapter IX] seeks at each iteration to optimize the fit between the predicted and observed arrival times by minimizing the ‘objective function’

$$\chi^2 \stackrel{\text{def}}{=} (\mathbf{A}\delta\mathbf{x} - \mathbf{b})^T \mathbf{S}_b^{-1} (\mathbf{A}\delta\mathbf{x} - \mathbf{b}) + (\mathbf{C}\delta\mathbf{y} - \mathbf{d})^T \mathbf{S}_d^{-1} (\mathbf{C}\delta\mathbf{y} - \mathbf{d}), \quad (\text{A2})$$

where the superscript T indicates matrix transposition and $\mathbf{S}_b \stackrel{\text{def}}{=} E[\mathbf{b}\mathbf{b}^T]$ and $\mathbf{S}_d \stackrel{\text{def}}{=} E[\mathbf{d}\mathbf{d}^T]$ are the symmetric $m_b \times m_b$ and $m_d \times m_d$ covariance matrices of the observational errors for the two data sets. The two terms on the right-hand side of eq. (A2) are the sums of the squares of the variance-normalized data residuals for the two epochs. In most cases, the errors in different observations are statistically independent and the covariance matrices \mathbf{S}_b and \mathbf{S}_d are diagonal, the diagonal elements being the squares of the standard errors of the observations, or equivalently the inverses of the weights to be given to the observations.

The solution to this minimization problem is found by differentiating χ^2 with respect to each of the components of $\delta\mathbf{x}$ and $\delta\mathbf{y}$ and setting the resulting expressions to zero. In this simple case, the problems for the two epochs are independent and we obtain two independent $n \times n$ systems of linear ‘normal’ equations,

$$\mathbf{A}^T \mathbf{S}_b^{-1} \mathbf{A} \delta\mathbf{x} = \mathbf{A}^T \mathbf{S}_b^{-1} \mathbf{b} \quad \text{and} \quad \mathbf{C}^T \mathbf{S}_d^{-1} \mathbf{C} \delta\mathbf{y} = \mathbf{C}^T \mathbf{S}_d^{-1} \mathbf{d}, \quad (\text{A3})$$

to solve for $\delta\mathbf{x}$ and $\delta\mathbf{y}$.

A2 Regularization

In most least-squares problems (and virtually all seismic-tomography problems) the design-equation matrices \mathbf{A} and \mathbf{C} have ranks less than n , because of inadequate sampling of the Earth by the available rays, and consequently the normal-equation matrices $\mathbf{A}^T \mathbf{S}_b^{-1} \mathbf{A}$ and $\mathbf{C}^T \mathbf{S}_d^{-1} \mathbf{C}$ are singular, so additional information must be supplied to ‘regularize’ the problem. This is conventionally done by using the modified objective function

$$\chi^2 + \alpha (\delta\mathbf{x}^T \mathbf{Q}_x^{-1} \delta\mathbf{x} + \delta\mathbf{y}^T \mathbf{Q}_y^{-1} \delta\mathbf{y}). \quad (\text{A4})$$

Here α is a Lagrange multiplier and $\mathbf{Q}_x \stackrel{\text{def}}{=} E[\delta\mathbf{x}\delta\mathbf{x}^T]$ and $\mathbf{Q}_y \stackrel{\text{def}}{=} E[\delta\mathbf{y}\delta\mathbf{y}^T]$ are $n \times n$ *a priori* covariance matrices of the perturbations to the models. The modified objective function (A4) leads to modified normal equations

$$(\mathbf{A}^T \mathbf{S}_b^{-1} \mathbf{A} + \alpha \mathbf{Q}_x^{-1}) \delta\mathbf{x} = \mathbf{A}^T \mathbf{S}_b^{-1} \mathbf{b} \quad \text{and} \quad (\text{A5})$$

$$(\mathbf{C}^T \mathbf{S}_d^{-1} \mathbf{C} + \alpha \mathbf{Q}_y^{-1}) \delta\mathbf{y} = \mathbf{C}^T \mathbf{S}_d^{-1} \mathbf{d}.$$

The simplest useful form for \mathbf{Q}_x and \mathbf{Q}_y is the $n \times n$ identity matrix \mathbf{I} , for which choice

$$\delta\mathbf{x}^T \mathbf{Q}_x^{-1} \delta\mathbf{x} \equiv \delta\mathbf{x}^2 \quad \text{and} \quad \delta\mathbf{y}^T \mathbf{Q}_y^{-1} \delta\mathbf{y} \equiv \delta\mathbf{y}^2. \quad (\text{A6})$$

and the regularization attempts to make the model perturbations $\delta\mathbf{x}$ and $\delta\mathbf{y}$ small. This ‘damped least-squares’ method, introduced by Marquardt (1963), requires simply adding the constant α to each diagonal element of the simple normal-equation matrix (A3). Other useful choices for \mathbf{Q} can impose constraints on other characteristics, such as the smoothness, of the model perturbations.

A3 Differential least-squares method

We wish to invert data sets for two epochs simultaneously, in a way that suppresses any tendency of the derived models for the two epochs to differ unless doing so significantly improves the fit to the data. We therefore modify the objective function (A4) still further

and seek to minimize the quantity

$$\chi^2 + \alpha (\delta\mathbf{x}^T \mathbf{Q}_x^{-1} \delta\mathbf{x} + \delta\mathbf{y}^T \mathbf{Q}_y^{-1} \delta\mathbf{y}) + \beta (\delta\mathbf{y} - \delta\mathbf{x})^T \mathbf{R}^{-1} (\delta\mathbf{y} - \delta\mathbf{x}), \quad (\text{A7})$$

where β is a new Lagrange multiplier and \mathbf{R} is a new $n \times n$ *a priori* covariance matrix.

Differentiating the quantity (A7) with respect to the components of $\delta\mathbf{x}$ and $\delta\mathbf{y}$ and setting the resulting derivatives to zero, we obtain a $2n \times 2n$ system of normal equations, which can be written in partitioned form as

$$\begin{bmatrix} \mathbf{E} + \beta \mathbf{R}^{-1} & -\beta \mathbf{R}^{-1} \\ -\beta \mathbf{R}^{-1} & \mathbf{G} + \beta \mathbf{R}^{-1} \end{bmatrix} \begin{bmatrix} \delta\mathbf{x} \\ \delta\mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{h} \end{bmatrix}, \quad (\text{A8})$$

where $\mathbf{E} \stackrel{\text{def}}{=} \mathbf{A}^T \mathbf{S}_b^{-1} \mathbf{A} + \alpha \mathbf{Q}_x^{-1}$, $\mathbf{G} \stackrel{\text{def}}{=} \mathbf{C}^T \mathbf{S}_d^{-1} \mathbf{C} + \alpha \mathbf{Q}_y^{-1}$, $\mathbf{f} \stackrel{\text{def}}{=} \mathbf{A}^T \mathbf{S}_b^{-1} \mathbf{b}$ and $\mathbf{h} \stackrel{\text{def}}{=} \mathbf{C}^T \mathbf{S}_d^{-1} \mathbf{d}$ are shorthand notations for quantities appearing in the regularized normal eq. (A5).

As before, the simplest useful choice for \mathbf{R} is the $n \times n$ identity matrix, in which case

$$(\delta\mathbf{y} - \delta\mathbf{x})^T \mathbf{R}^{-1} (\delta\mathbf{y} - \delta\mathbf{x}) = (\delta\mathbf{y} - \delta\mathbf{x})^2, \quad (\text{A9})$$

and we seek to minimize the sum of the squares of the element-by-element differences of the vectors $\delta\mathbf{x}$ and $\delta\mathbf{y}$.

Most useful choices of the covariance matrix \mathbf{R} consist mostly of zeros, and dealing with the full $2n \times 2n$ system (A8) is wasteful of computer memory and processor time. When \mathbf{R} is diagonal the storage inefficiency is nearly a factor of two and the computational inefficiency is nearly a factor of four. (This statement refers to the task of solving the equations only. Generating the equations in the first place is often more time consuming.) It is possible, however, to solve the system of eq. (A8) for $\delta\mathbf{x}$ and $\delta\mathbf{y}$ with little more labour than it takes to solve two independent sets of normal equations separately, by the following algorithm:

Algorithm 1.

(1) Invert the $n \times n$ matrix $\mathbf{G}' \stackrel{\text{def}}{=} \mathbf{G} + \beta \mathbf{R}^{-1}$ and compute the n -vector $\mathbf{G}'^{-1} \mathbf{h}$.

(2) Solve the $n \times n$ system

$$(\mathbf{E}' - \beta^2 \mathbf{R}^{-1} \mathbf{G}'^{-1} \mathbf{R}^{-1}) \delta\mathbf{x} = \mathbf{f} + \beta \mathbf{R}^{-1} \mathbf{G}'^{-1} \mathbf{h} \quad (\text{A10})$$

for $\delta\mathbf{x}$, where $\mathbf{E}' \stackrel{\text{def}}{=} \mathbf{E} + \beta \mathbf{R}^{-1}$.

(3) Compute $\delta\mathbf{y} = \mathbf{G}'^{-1} (\mathbf{h} + \beta \mathbf{R}^{-1} \delta\mathbf{x})$.

Step 3 follows from the bottom half of the system of normal eq. (A8). Steps 1 and 2 follow from the top half of the inverse of the system (A8),

$$\begin{bmatrix} \delta\mathbf{x} \\ \delta\mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{E}'^{-1} & \beta \mathbf{E}'^{-1} \mathbf{R}^{-1} \mathbf{G}'^{-1} \\ \beta \mathbf{G}'^{-1} \mathbf{R}^{-1} \mathbf{E}'^{-1} & \mathbf{G}'^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \mathbf{h} \end{bmatrix}, \quad (\text{A11})$$

whose validity can be verified directly by substitution. Here we have introduced the shorthand symbols

$$\mathbf{E}'' \stackrel{\text{def}}{=} \mathbf{E}' - \beta^2 \mathbf{R}^{-1} \mathbf{G}'^{-1} \mathbf{R}^{-1} \quad \text{and} \quad (\text{A12})$$

$$\mathbf{G}'' \stackrel{\text{def}}{=} \mathbf{G}' - \beta^2 \mathbf{R}^{-1} \mathbf{E}'^{-1} \mathbf{R}^{-1}.$$

Steps 1 and 2 each require the inversion of an $n \times n$ system, while step 3 requires only a matrix-vector multiplication, so the total computational labour is comparable to solving two independent $n \times n$ systems.

This algorithm is implemented in the C-language code *tomo4d*.