Ryberg, T., Haberland, C. (2019): Bayesian simultaneous inversion for local earthquake hypocentres and 1-D velocity structure using minimum prior knowledge. - Geophysical Journal International, 218, 2, 840-854.

https://doi.org/10.1093/gji/ggz177

Geophysical Journal International

Geophys. J. Int. (2019) **218,** 840–854
Advance Access publication 2019 April 16
GJI Seismology

doi: 10.1093/gji/ggz177

Bayesian simultaneous inversion for local earthquake hypocentres and 1-D velocity structure using minimum prior knowledge

T. Ryberg and Ch. Haberland

Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences, Telegrafenberg, Potsdam, 14473 *Potsdam, Germany.*
E-mail: trond@gfz-potsdam.de

Accepted 2019 April 15. Received 2019 April 11; in original form 2018 December 11

SUMMARY

We present a Bayesian approach to solve the problem of simultaneous inversion for optimal hypocentre parameters and 1-D velocity models as well as station corrections for a given set of local earthquakes utilizing a hierarchical, transdimensional Markov chain Monte Carlo (McMC) algorithm. The simultaneous inversion is necessary because of the velocity–hypocentre coupling inherent to the problem.

Tests with synthetic arrival time data indicate an excellent performance of the approach, at the same time benefiting from all the advantages related to the McMC algorithm. These advantages are that only minimum prior knowledge is used (i.e. regarding starting focal coordinates, initial velocity model, which are set to random initial values), no regularization parameters (e.g. damping) have to be selected, and the parametrization of the velocity model (i.e. model nodes/layers) is automatically set and adjusted according to the quality of the data, that is noise level. By minimizing the amount of pre-inversion assumptions, which are regularly not available at the required precision or often only available after very careful and time-consuming assessment, the inversion results are therefore almost exclusively data-driven. On output, we obtain a suite of well fitting models which can statistically be analysed and provide direct estimates of the posterior uncertainties of the models.

Tests with real arrival time data from a temporary local network deployed in South-Central Chile in 2004 and 2005 show a very good agreement with the results obtained with a conventional inversion method.

Key words: Inverse theory; Statistical methods; Body waves; Crustal imaging; Seismic tomography.

1 INTRODUCTION

Determining hypocentre coordinates (spatial coordinates x_0, y_0, z_0 and origin time t_0) of earthquakes is a fundamental task in seismology. This task is usually solved by picking the phase onset times (e.g. of first arrival *P* and *S* waves) on the waveform data of the seismic network stations and feeding them into an inversion process. Traditionally, this is done in an iterative way by least-squares inversion after the inversion problem has been linearised (e.g. Geiger 1910, 1912; Eaton 1970; Lee & Lahr 1975; Lienert *et al.* 1988; Lahr 1989; Lienert & Havskov 1995). These approaches have the advantage of being very fast and effective. However, they require *a priori* information on initial values which are often not available (i.e. velocity model which should be close to the 'real' values). More recent attempts involve grid search techniques (Sambridge & Kennett 1986; Bai *et al.* 2009), Oct-Tree Importance sampling (Husen *et al.* 2003), Monte Carlo methods (Lomax *et al.* 2000, 2009; Lomax 2005; Myers *et al.* 2007; Theunissen *et al.* 2018) or genetic algorithms (Billings *et al.* 1994; Ružek & Kvasnička 2001). Most of these approaches thoroughly search the model space and have the advantage of providing uncertainty estimates of the derived focal parameters.

In either case, the hypocentre coordinates are intrinsically tied to the seismic velocities in the subsurface (hypocentre–velocity coupling; see e.g. Kissling 1988; Thurber 1992; Kissling *et al.* 1994), which can be highly variable, particularly at the local scale. While the location methods listed above (single event location) rely on fixed and previously known (often 1-D) velocity models (input), other attempts invert simultaneously for the hypocentre parameters of a whole set of earthquakes and the velocity structure (and/or station corrections; see also joint hypocentre determination; Douglas 1967; Pujol 2000, 2003). This can, in particular, be applied to earthquakes occurring within seismic networks on a local scale (local events). It can be done assuming horizontally layered (1-D) models (e.g. *velest*; Kissling *et al.* 1994) or 3-D models (Aki & Lee 1976, 'simultaneous inversion for hypocentres and 3-D velocity

Table 1. Model space and selection of starting models.

Model parameter	Range (min/max)		Starting model m_0
Hypocentres (x, y)	–200	200 km	Uniform in range
Quake depth (z)	0	200 km	Uniform in range
Velocity, *Vp*	2	12 km s^{-1}	Normal with mean/sigma 6/0.5 km s^{-1}
		Markov chain Monte Carlo	
Vp/*Vs* ratio	1	2.5	Normal with mean/sigma 1.732/0.2
Layer depth	0	200 km	Uniform in range
Number of layers	1	200	Normal with mean/sigma 5/3
Noise σ_p	0.001	10.0 s	1 s
Noise σ_s	0.001	10.0 s	1 s
Station correction τ^p	–5	5 s	0 s
Station correction τ^s	–5	5 s	0 s

Note that all prior distributions are uniform. The ranges and starting values are uncritical, that is significant changes do not change the general appearance of the derived reference model.

Table 2. Width of the Gaussian distribution in model perturbation.

Parameter	Gaussian width
Earthquake hypocentres (x,y,z)	2 km
Velocity *Vp*	0.05 km s^{-1}
Vp/*Vs* ratio	0.05
Layer depth	10 km
Noise σ_p and σ_s	0.01 s
Station corrections τ^p and τ^s	0.05 s

The exact values are not critical, that is their choice does not change the derived reference model, but instead mainly influences the acceptance rate.

structure' or 'local earthquake tomography', for example, Thurber 1983, 1993; Kissling 1988; Eberhart-Phillips 1986a; Evans *et al.* 1994; Eberhart-Phillips & Michael 1998). The advantage of these methods is that the velocity models are derived directly from the data at the same time as the earthquakes are (re-)located, in this way taking into account the hypocentre–velocity coupling. Also these (joint hypocentre/velocity structure) methods conventionally employ iterative inversion strategies based on (damped) least squares after the inversion problem has been linearised (e.g. Thurber 1983; Kissling 1988; Kissling *et al.* 1994). While these methods are very robust and have been very successfully applied now for decades (e.g. Husen *et al.* 1999; Husen & Kissling 2001; Eberhart-Phillips *et al.* 2005, 2006; Haberland *et al.* 2006, 2009), the use of the regularized inversion suffers from the need for initial values (for velocity and hypocentres), regularization parameters such as damping or smoothing, and/or a preselected model parametrization. Furthermore, estimates of uncertainties are hard to derive during this approach and synthetic tests or bootstrapping have usually to be invoked to assess the resolution and the uncertainties of both the velocity structure and the hypocentre parameters.

In this paper, we apply a Bayesian approach to the problem of jointly determining the hypocentre parameters and 1-D velocity models (for *P* and *S* waves) as well as station corrections for a given set of local earthquakes. In the Bayesian approach all information about the model **m** is represented in probabilistic terms. The aim of the Bayesian inference is to quantify the posterior probability distribution of parameters of model **m** given the observations

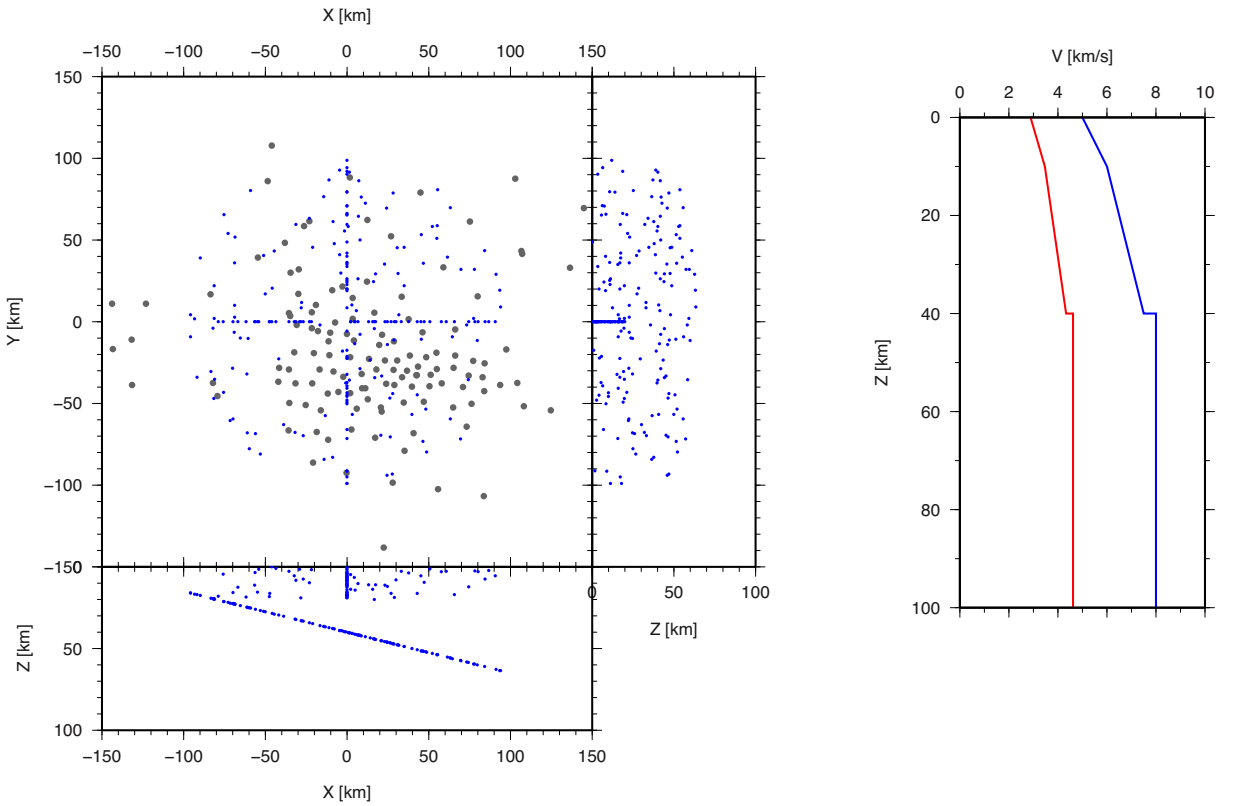

Figure 1. Station and earthquake distribution used in the synthetic test. Left-hand panel: map and side views of stations (grey circles) and earthquakes (blue circles), right-hand panel: 1-D synthetic model (blue: *Vp*, red *Vs*). The *Vp*/*Vs* ratio was $\sqrt{3}$ in the entire model.

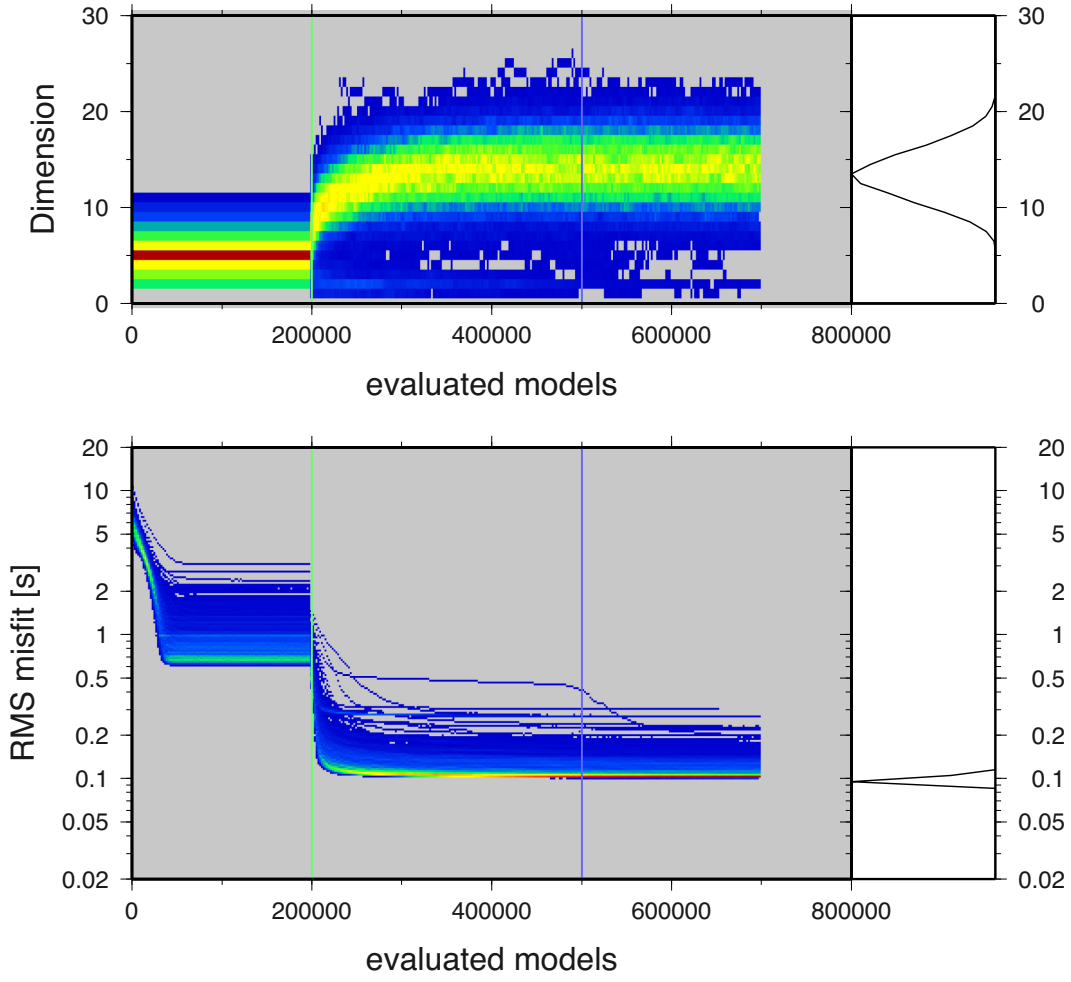

Figure 2. Convergence of 1000 Markov chains (synthetic test). Shown is the distribution of the data misfit (bottom) and model dimension (number of cells, top) during the evolution along the chains, red and blue colors correspond to high and low probabilities. For the first 200 000 models only the earthquake locations are changed along the Markov chains (green line), beyond this point the velocity model (*Vp*, *Vp*/*Vs*), the station corrections and the data noise level was allowed to change. Beyond 500 000 models the sampling was assumed to be stationary (burn-in phase finished, blue line). Note that a small number of post-burn-in chains does not reach low misfits. By taking only 90 per cent of the best-fitting models we automatically excluded these chains. Relative histogram plots of the distribution of data misfit and model dimension for the post-burn-in phase are added at the right-hand side. The best-fitting models are typically characterized by a somewhat higher model dimensionality. Note the log scale for the data misfit.

$\mathbf{d}_{\mathbf{obs}}$ which can be written as $p(\mathbf{m}|\mathbf{d}_{\mathbf{obs}})$. We do this by employing a hierarchical, transdimensional Markov chain Monte Carlo (McMC) algorithm (Metropolis *et al.* 1953) following closely the procedures outlined in Bodin *et al.* (2012a,b). The main advantages are that only minimum prior knowledge is needed (independence from starting models, regularization parameters, parametrization, assumed noise) and that the results can be statistically analysed including the determination of so-called reference models (e.g. averages) and uncertainty estimates. McMC have recently gained a lot of attention in a broad range of geophysical disciplines such as waveform fitting (Mosegaard 1998), ambient noise tomography (Bodin *et al.* 2012a) or electrical resistivity inversion (Schott *et al.* 1999). In a recent paper, we used a McMC method to derive 2-D velocity models from refraction seismic data sets (Ryberg & Haberland 2018). Gesret *et al.* (2015) applied a two-step approach of first using a Monte Carlo method to invert controlled source data to derive a velocity model, and then locate—again using a probabilistic framework—earthquakes using the previously derived velocity model. Agostinetti *et al.* (2015) applied a hierarchical, transdimensional McMC method for the simultaneous inversion of earthquake hypocentre locations and the 3-D velocity structure. However, they seem to have used rather restricted prior information (for hypocentre locations and velocity structure) based on prior knowledge (i.e. from conventional inversion runs) and only used *P*-wave arrival times. In our study we follow a strict minimum prior-information approach and start directly from the arrival time picks and random initial models.

2 METHOD

Instead of using conventional methods (based for instance on regularized inversion) we use a probabilistic (statistical) approach for joint hypocentre and velocity structure determination. In the search, randomly created models (*Vp*, *Vp*/*Vs* models; hypocentre coordinates; *P* and *S*‘ station corrections) are tested and their capability to

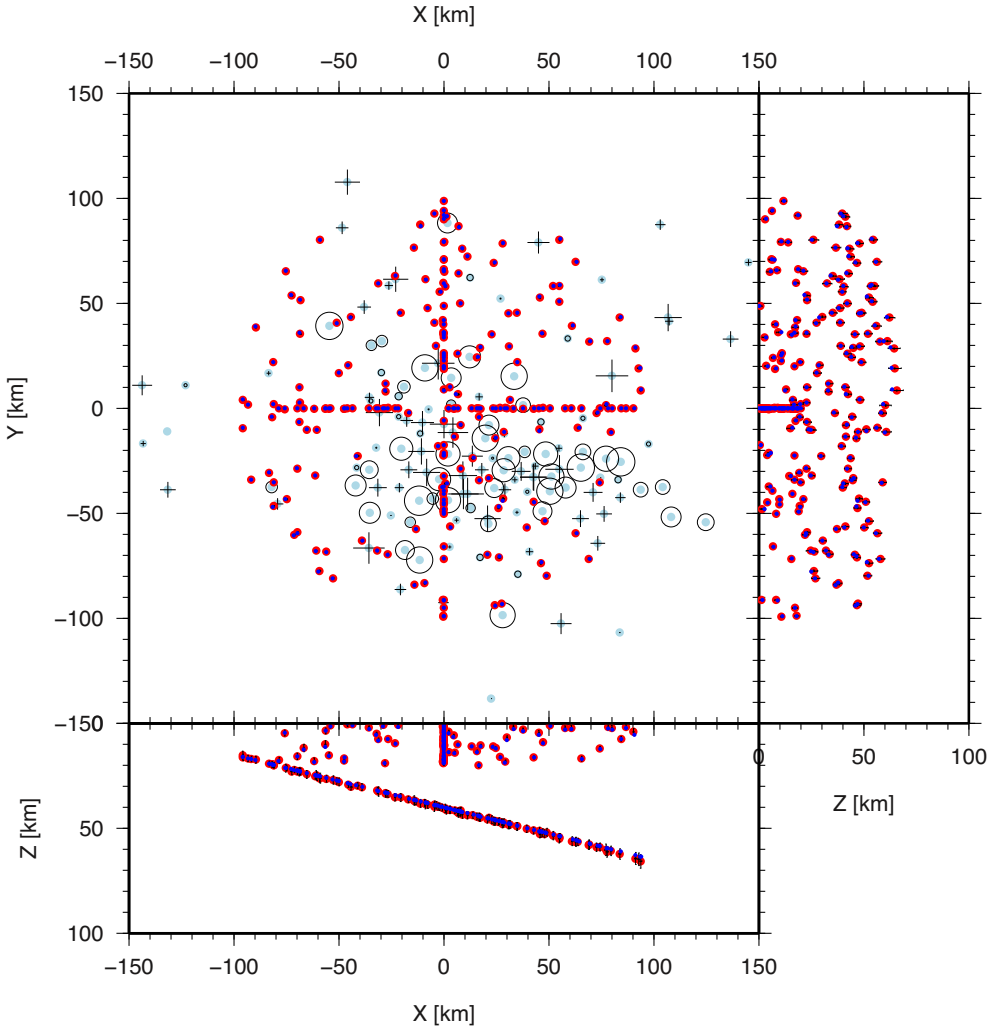

Figure 3. Results of the synthetic test. Locations of the recovered earthquakes (red circles) with error bars (one σ) compared to the exact locations (blue dots). Light blue circles indicate recovered station locations. Crosses (positive corrections) and open circles (negative corrections) show the recovered station corrections, scaled by their values.

explain the data (or misfit, respectively) is calculated. The models are accepted or rejected according to certain acceptance criteria, new models are then (randomly) created by perturbation of the previous model, and the evaluation and acceptance process starts again. After the development of a so-called 'chain' (Markov chain) of a large number of tested models, the well fitting models are explored in a statistical way. The main ingredients of the method such as the model parametrization, the forward problem, the misfit function, the details of the search procedure and—eventually—the statistical analysis will be presented in the following.

We use the transdimensional and hierarchical version of the McMC method, that is we treat the velocity model dimension (model complexity or number of layers) as unknown and let the data decide. In addition, the hierarchical version tries to invert for data noise, that is split the data into signal (part of the data which can be explained by the model) and noise. The combination of both McMC extensions can be seen as some kind of auto-regularization, where the data noise level automatically controls the velocity model complexity. For instance, very noisy data lead automatically to less complex models (i.e. with a small number of layers). No prior knowledge of the number of velocity layers and/or data noise level is necessary, thus leading to a completely data-driven inversion approach.

Model parametrization, forward problem and misfit function

The 1-D velocity models are described by a set of points $p_i = (z_i, Vp_i, Vp_i/Vs_i)$, with $0<i<K$, z_i being the depth and Vp_i (Vp_i/Vs_i) the seismic *P*-wave velocity (*Vp*/*Vs* ratio). K is the number of points/nodes forming the model—this is an unknown, variable number (see below). In addition to the *Vp* velocity and *Vp*/*Vs* ratio, the model contains the hypocentre coordinates of the earthquakes, the *P* and *S* wave station corrections, and the *P* and *S* data noise levels. Therefore, the free parameters (model space dimension) in the search are:

(i) number of earthquakes x 3 spatial hypocentre coordinates +
(ii) number of model nodes/layers (variable) x 3 (*P*-wave velocity + *Vp*/*Vs* ratio + node depth) +
(iii) number of station corrections x 2 (for *P*- and *S*-wave arrivals) +
(iv) value of data noise x 2 (for *P*- and *S*-wave arrival time picks).

For the forward travel time calculation (see below) we have to convert the velocity model which is composed of irregular points (layers or cells) into a regular mesh. We use an interpolation based

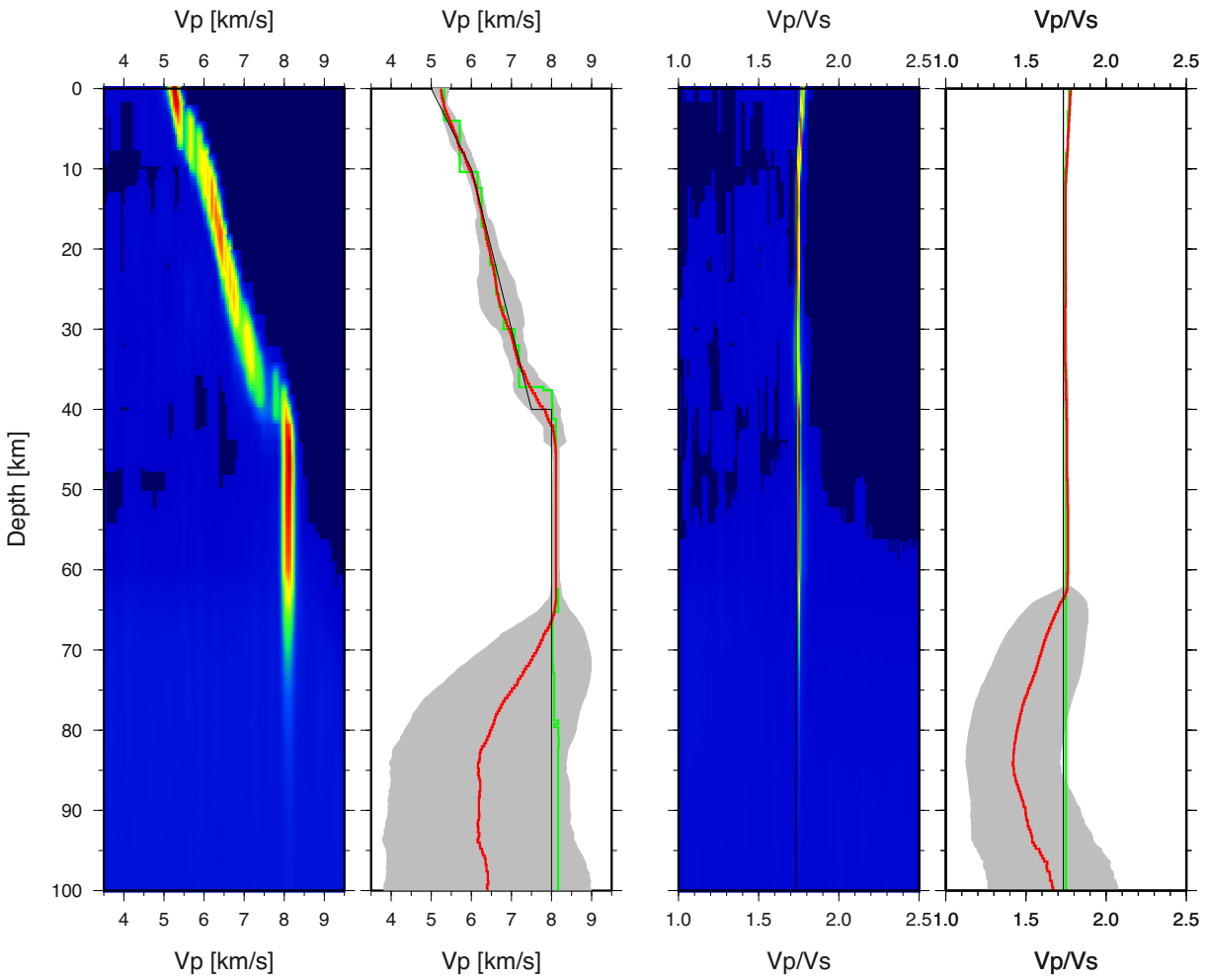

Figure 4. Results of the synthetic test: recovered *Vp* and *Vp*/*Vs* ratio models. The left two panels show the results for *Vp*: the histogram plot of the *Vp* versus depth (warm colors correspond to high probabilities) and a comparison with the synthetic model (black line). The red line shows the averaged (modified averaging procedure, see text) *Vp*, the grey region shows the standard deviation (1σ) and the green line indicates the maximum probability model (velocity of highest probability). The right-hand panels show the respective results for *Vp*/*Vs*. Down to a depth of ∼60 km the synthetic model was recovered very well, except in the depth range around 40 km (crust/mantle discontinuity) with an elevated velocity uncertainty.

on Voronoi cells, which means that the velocities at a certain depth are set to the velocity value of the nearest model point p_i.

For the forward calculation (of the travel times t^{cal}) we use the efficient 2-D finite differences (FD) solution (Eikonal solver) of Podvin & Lecomte (1991). The regular velocity mesh (with given numbers of nodes in *x*-direction (distance) and *z*-direction (depth) and given spatial increment h) needed for the forward calculation is formed from the interpolated points p_i for the required maximum epicentral distances (see above). For the travel time calculations we used a grid size of 1 km (vertically and horizontally). We tested coarser and finer grids but found no significant changes in the reference model, except for the large influence on computation time. For efficiency, the forward calculation (travel times) is pre-calculated for many sources for the complete range of depths (from the surface to some maximum depth at a certain increment) and stored in a table which can be efficiently accessed during the inversion/search. The travel time table is only re-calculated when the model has changed thus reducing the calculation time to a minimum.

In order to account for the station elevations we calculate station-specific travel time delays (separately for *P* and *S* waves) from the station elevation and the velocity v_{elev} of the shallowest node point (layer) and add these values to the travel times from the table. Furthermore, we introduce additional station-specific time delays τ^{P}_{l} and τ^{S}_{l} (for *P* and *S* waves; for station l) which account for deviations from the 1-D model and the very shallow velocity structure beneath the stations. Thus, the synthetic travel time (observation j) from source i can be formulated as:

$$t_{ij}^{syn} = t_{ij}^{cal} + \tau_{(j)} + z_{(j)}/v_{elev}$$

with (j) indicating the station index corresponding to a particular observation j.

To calculate the misfit function we assume that the data set consists of arrival times T_{ij}^{obs} of N earthquakes ($1 <= i <= N$ with $1 <= j <= M_i^P$ or $1 <= j <= M_i^S$ for *P* and *S* observations, respectively) referenced to the arrival time $T_i^{obs,\,MIN}$ of the earliest observation of an event i

$$t_{ij}^{obs} = T_{ij}^{obs} - T_i^{obs,MIN}.$$

The summed difference between observed (referenced) arrival and calculated travel times for event i for the *P* arrival j is

$$\Delta t_{ij}^{p} = t_{ij}^{P,obs} - t_{ij}^{P,syn}$$

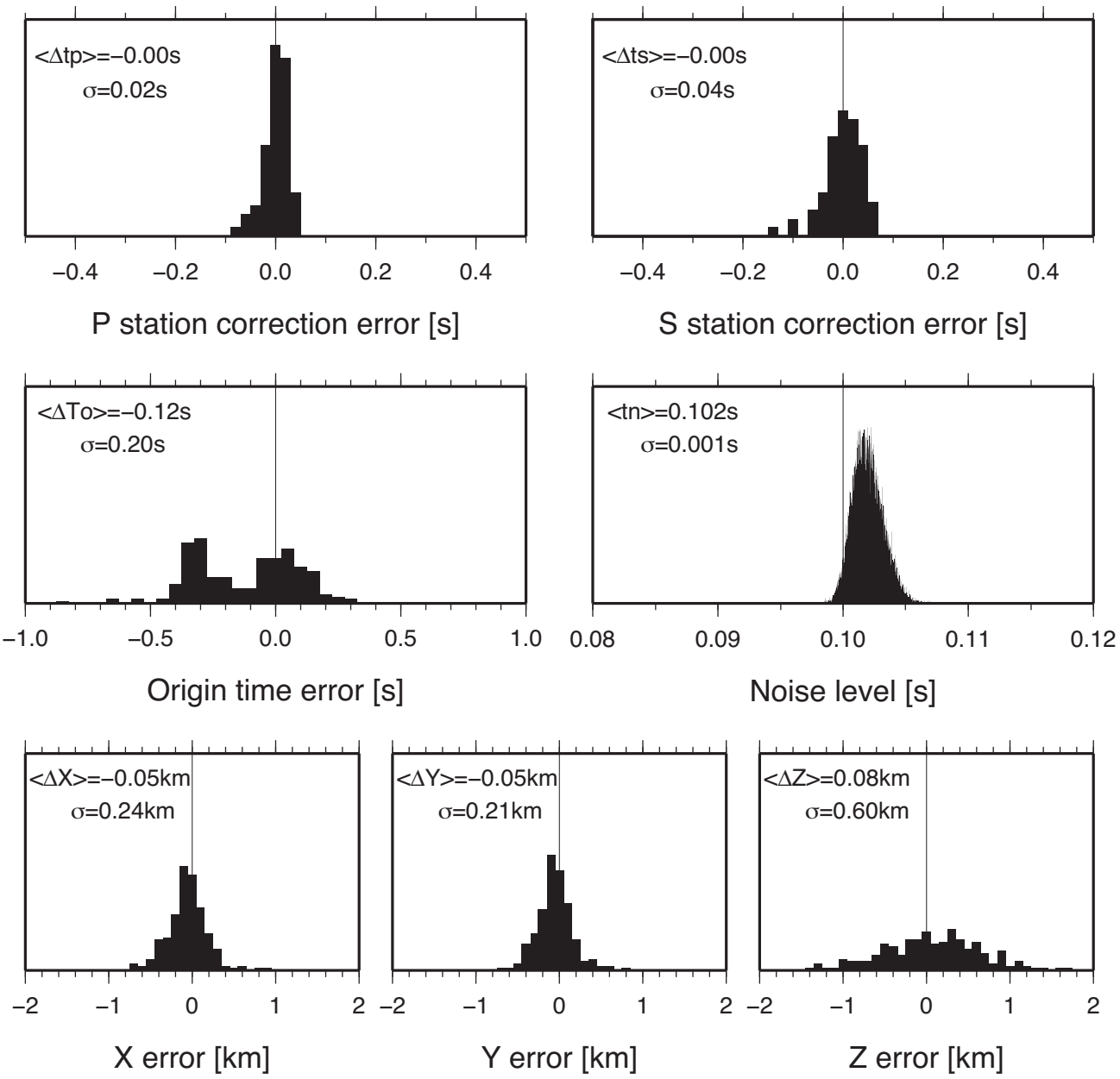

Figure 5. Misfit of recovered horizontal and vertical coordinates of the earthquake hypocentres (bottom row), the recovered quake origin time and data noise level (middle panels) and recovered station corrections for *P* and *S* phases (top; synthetic test). While the recovered epicentral locations deviate by less than a few hundred metres (RMS), the recovered quake depths deviate by several hundred meters (RMS). Both station corrections for *P* and *S* phases and the origin time could be recovered very well. The data noise (time jitter added to the synthetic travel times) is slightly higher than the 0.1 s added. Forward modelling errors might explain this small discrepancy.

and for the S arrival is

$$\Delta t_{ij}^{S} = t_{ij}^{S,obs} - t_{ij}^{S,syn}.$$

The misfit function $\Phi(m)$ of a model $\mathbf{m}$ is thus:

$$\Phi\ (m) = \sum_{i=1}^{N} \left[\sum_{j=1}^{M_i^P} \frac{\left(\Delta t_{ij}^{P} - \langle \Delta t_i \rangle\right)^2}{\sigma_p^2} + \sum_{j=1}^{M_i^S} \frac{\left(\Delta t_{ij}^{S} - \langle \Delta t_i \rangle\right)^2}{\sigma_S^2} \right],$$

with σ being data noise (for *P* and *S* travel times). The noise includes actual data noise (picking errors), inaccuracies due to approximations (i.e. treatment of station elevations) and forward travel time prediction errors (i.e. caused by the finite forward grid size) and is typically assumed to be normally distributed and uncorrelated. The term $\langle \Delta t_i \rangle$ is the average (mean over all travel time picks for event i) of the differences Δt_{ij}^{P} and Δt_{ij}^{S}, respectively, and is defined as $\langle \Delta t_i \rangle = \frac{1}{M_i^P} \sum_{j=1}^{M_i^P} \Delta t_{ij}^{P} + \frac{1}{M_i^S} \sum_{j=1}^{M_i^S} \Delta t_{ij}^{S}$. In the case of well fitting models $\langle \Delta t_i \rangle$ is the relative origin time with respect to $T_i^{obs,MIN}$. The origin time is an output of our search and can be calculated by:

$$T_i^{orig} = T_i^{obs,MIN} + \langle \Delta t_i \rangle .$$

Average values of the origin times and their uncertainties can be calculated. Minimizing the misfit function $\Phi(m)$ is equivalent to maximizing the Gaussian likelihood function which takes the form:

$$p\ (d|m) = \left(2\pi\sigma_p^2\right)^{\frac{-M_p}{2}} \left(2\pi\sigma_s^2\right)^{\frac{-M_s}{2}} e^{-\frac{\Phi}{2}}.$$

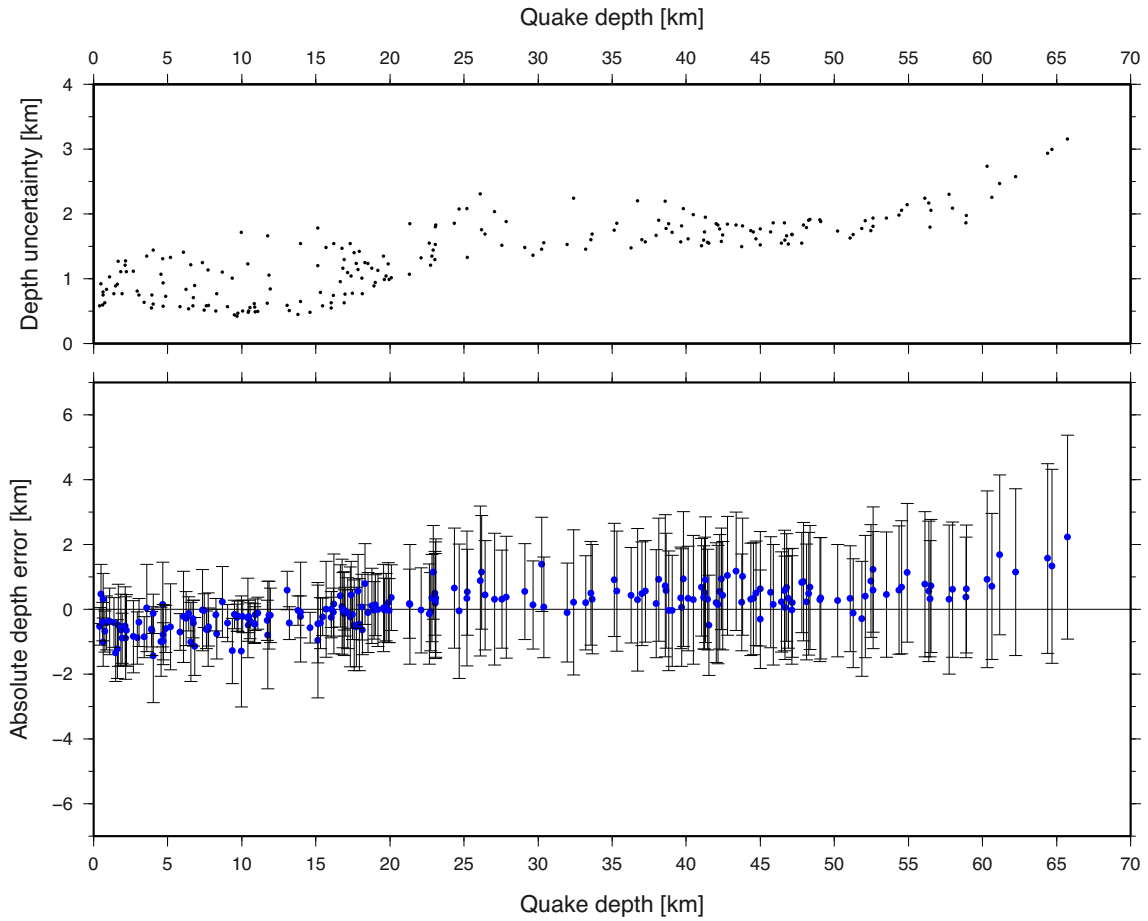

Figure 6. Dependence of depth error (misfit) on quake depth (bottom panel) in the synthetic test. Blue dots show the difference between the recovered vs. synthetic depth, bars indicate the 1σ standard deviations. The depth uncertainty of recovered quakes (top panel) shows a slight tendency of increase for quakes at greater depth.

Transdimensional, hierarchical McMC method

Starting from a randomly chosen model $\mathbf{m_0}$ (i.e. 1-D velocity for *P* and *S* waves, hypocentre coordinates, station corrections, data noise level) a new model $\mathbf{m_1}$ is drawn from a proposal distribution which depends on the current model. Data are predicted for the new model $\mathbf{m_1}$ followed by the calculation of the misfit. This model is accepted if the misfit of model $\mathbf{m_1}$ is smaller than that for model $\mathbf{m_0}$. Note that even if the misfit is larger, there is still some non-zero probability that the proposed model will be accepted. If accepted the new model is a new sample in the Markov chain, if rejected another new model will be proposed. This sequence is repeated until convergence is reached, so that the density of samples converges to the posterior probability distribution (for details see for instance Bodin *et al.* 2012a).

For our problem when inverting for hypocentres and velocity structures, there are seven types of model perturbations: adding a new layer in the velocity model, removing a layer from the velocity model, changing the *P*-wave velocity of a randomly selected layer, changing the *S*-wave velocity of a randomly selected layer, changing the location of a randomly selected earthquake, changing the correction time of a randomly selected station, changing the data noise parameters (i.e. noise for *P*-wave travel time picks and for *S-wave* travel time picks). Instead of keeping the correction times (for *P* and *S* waves) fixed at a reference station we keep the average (across the network) of *P*- and *S*-wave station correction times to be zero.

The choice of the proposal probability distribution, that is a Gaussian one, theoretically does not affect the sampling of the posterior distribution under the assumption of infinite long Markov chains. In reality, the specific choice is important to achieve fast convergence and thorough sampling of the posterior distribution. We choose a Gaussian distribution with zero-mean and standard deviation s for the proposal probability of the model perturbation. For example, in the case of very small values of s, the progress in exploring the model space will be small, getting 'stuck' in local misfit minima, while the acceptance rate is large. In the opposite case of very large values of s the acceptance rate will be decreased. Some optimal values exist to provide fast convergence of the Markov chains.

Since successive models along a Markov chain are not independent, we thinned the chain by retaining only every 2000th sample of the chain. All models before the convergence is achieved will be ignored for further analysis (burn-in phase). After the burn-in phase, we assume stationary sampling of the model space, thus generating a large set of models well-fitting the data. Statistical properties (i.e. mean, standard deviation, modes) of those models will be used to calculate final *P*- and *S*-wave models, earthquake hypocentres and correction times.

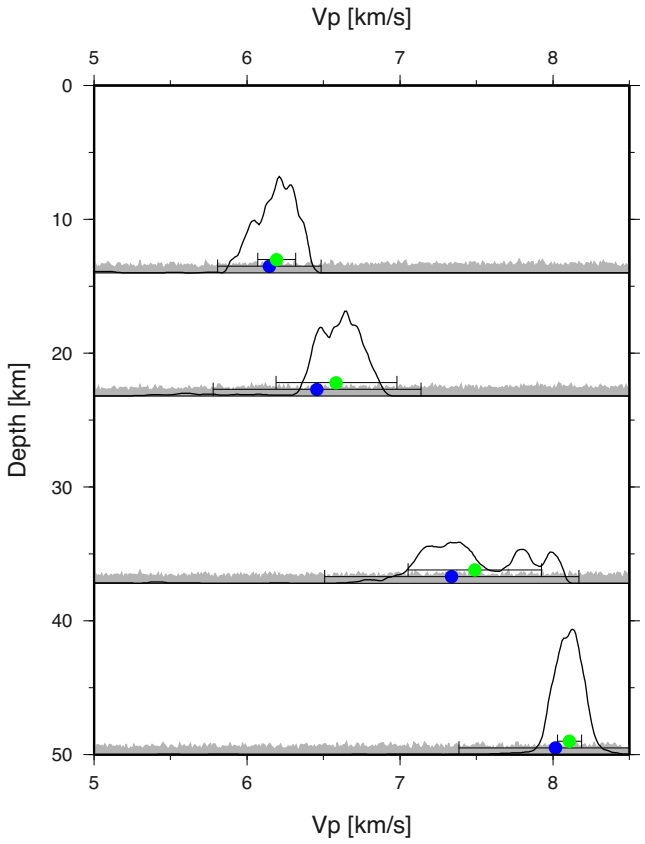

Figure 7. Posterior distribution of *Vp* at selected depths (synthetic test). Blue circles and associated error bars show the classical average values and their standard deviation. Green circles and error bars show modified averages and standard deviations. Note that the velocity distributions are not Gaussian, but are more complex (several modes). Classical averaging to derive a final model will fail. The prior distribution, a uniform distribution between 2 and 12 km s^{-1} (bounds on *Vp*), is shown in light grey. The distribution at ∼37 km depth appears to be rather blurred because it is located close to the crust–mantle transition ('Moho').

Prior information, starting a Markov chain and convergence assessment

Our original intention was to keep the amount of prior knowledge at a minimum (i.e. assumptions, known velocity models prior inversion). This choice is called non-informative prior distribution and is used by us to avoid any potential dependence of our final model on those assumptions or potentially uncertain prior information.

Thus, instead of starting a Markov chain from any prior given model (*P*- and *S*-wave velocity model, earthquake hypocentres known for the region) we started the chains with random initial models: randomly chosen number of layers with random velocity and *Vp*/*Vs* ratio, randomly distributed starting locations of hypocentres, see Table 1.

Since the choice of the proposal probability of the parameters (standard deviation in case of Gaussian, see Table 2) controls the speed of chain convergence, we tested different s to determine optimal values, finally resulting in acceptance rates between 30 and 40 per cent. In addition to the choice of the proposal distributions, we allowed for a very wide range of model properties, that is velocity values are bounded between 2 and 12 km s^{-1} for *P* waves (divided by $\sqrt{3}$ for respective *S* waves). This includes completely 'unrealistic' rock velocities and *Vp*/*Vs* ratios to avoid a truncation of the sampling of the posterior distribution. We carefully checked that the choice of these artificial boundaries does not affect the posterior.

We monitored the convergence of the Markov chains to make sure that the posterior distribution of the final models is stationary. We assumed that the sampling along a Markov chain became stationary when the misfit and model dimension (number of layers) approached an asymptotic value and only jittered around this value.

To accelerate the thorough sampling of the model space, we simultaneously started 1000 separate Markov chains, and combined their output after the burn-in phase for further statistical analysis. All of them investigated 1 million models. To accelerate the convergence of an individual Markov chain we perturbed for the first 200 thousand models of an individual chain the earthquake hypocentres only, then started the full exploration of velocity model space, noise parameters and station corrections. We found that the burn-in phase was finished starting from ∼500 thousand models. The successive models, all of them fitting the data very well, have been used to derive statistical model properties, that is earthquake hypocentres, *P*- and *S*-wave velocity models, noise hyper-parameters and station corrections.

3 APPLICATION TO SYNTHETIC DATA

We applied the proposed algorithm to a synthetic data set for a typical subduction zone scenario. The simple 1-D velocity model consisted of two gradient layers and a velocity jump at 40 km depth ('Moho'). Note that the two gradient layers, although being quite realistic in the context of crustal structures, pose a challenge to be represented with constant value Voronoi cells (layers) and therefore need a larger number of cells (layers) to be described with sufficient accuracy. To simulate a realistic station distribution, we adapted the locations of 130 stations of the TIPTEQ project (see Section 4). A synthetic data set of 220 earthquake locations was generated consisting of randomly distributed earthquakes at an artificially dipping surface ('subduction zone quakes') and in two vertical curtains ('crustal fault zones'). Station distribution, hypocentres and velocity model are shown in Fig. 1. For the hypocentres, travel times were calculated with a 3-D FD Eikonal solver (Podvin & Lecomte 1991; Tryggvasson & Bergman 2006) and 0.1 s random noise (RMS) and random stations corrections for *P*- and *S*-picks (equally distributed between –1.0 and 1.0 s) were added. The data set was limited to events located within the network (i.e. events with a so-called 'GAP'angle $<180^\circ$. Altogether 9207 *P*-wave and 5874 *S*-wave picks were used for the joint inversion. The size of the forward grid is the major factor controlling the CPU runtime for the inversion. We tested different grid increments: 1 and 2 km, resulting in grid sizes of 400 × 400 × 200 and 200 × 200 × 100, respectively. We found no significant differences in the final inversion results. Even sparse grid increments of 5 km, for which we would expect larger errors related to the forward travel time calculations, gave only slightly deteriorated final velocity models and earthquake locations.

1000 Markov chains with completely randomized starting models (velocity structure, earthquake locations, but zero station corrections) were used to investigate the model space. For the first 200 000 models only the earthquake locations have been perturbed, followed by 500 000 more models with the complete set of model perturbations. Typically, the point of stationary sampling of the model space (end of burn-in phase, when the misfit is not significantly decreasing any more) is reached after 500 000 models (see

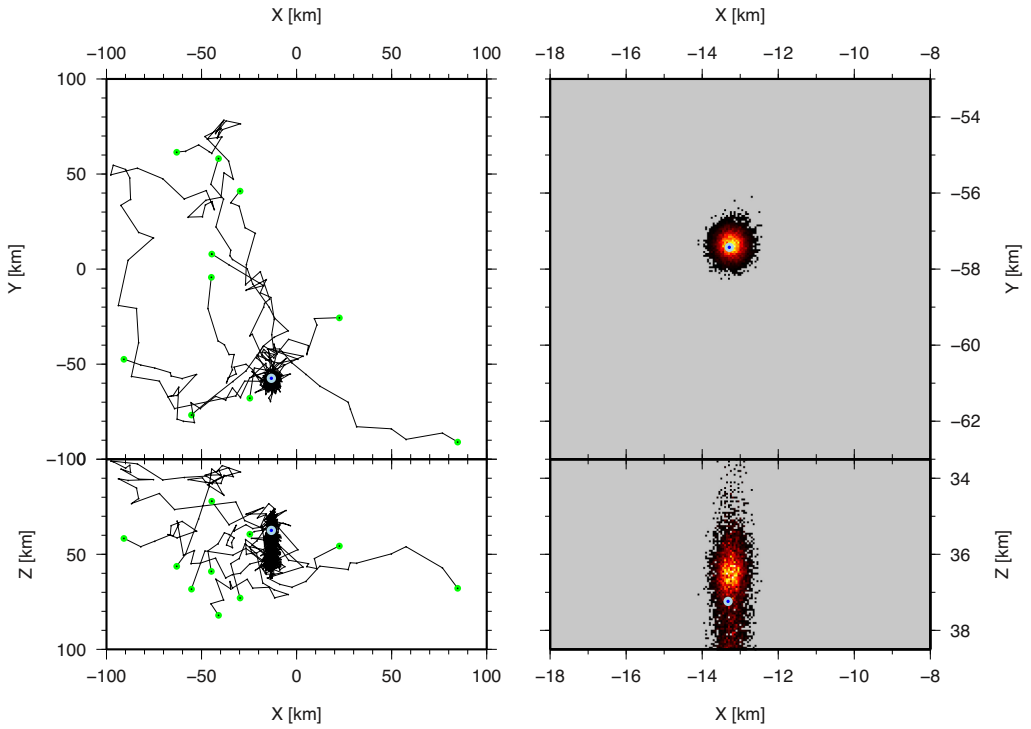

Figure 8. Evolution of earthquake locations (x–y and x–z plots) for 10 chains (left-hand panel) for one selected earthquake (synthetic test). The random starting positions are indicated by green circles. Black lines show the 'migration' of the quake locations during the evolution of a chain. The blue dot shows the exact position of the earthquake. On the right-hand side a histogram plot (zoomed-in heat map) of the final quake locations (post-burn-in) are shown. While the epicentre location (white colour) fits the exact location (blue dot) very well and is quite narrow, the depth resolution (vertical axes of distribution) is significantly poorer and the centre (white colour) is shifted by several hundred meters with respect to the exact location.

Fig. 2). The remaining 500 000 models are decimated by a factor of 2000 and all models from the 1000 individual Markov chains were combined. Since occasionally some Markov chains do not converge to small data misfits (probably get 'stuck' in local misfit minima), we decimated the models by removing 10 per cent of the worst fitting ones (travel time misfits), similar to Shen *et al.* (2013). The remaining models (∼30 000) show a generally small misfit of ∼0.1 s, which is coincident with the random noise level added to the synthetic travel times. The algorithm needs roughly between 10 and 15 layers to properly model the data (see Fig. 2, top).

From these models we derived average and standard deviation values for earthquake hypocentres, station corrections (Fig. 3) and *P*-wave velocity and *Vp*/*Vs* ratio (Fig. 4). For the *P*-wave velocity and *Vp*/*Vs* ratio models we additionally determined the maximum posterior probability model (MAP). We observed that the posterior distribution of velocity values in specific depth ranges is non-Gaussian (see Fig. 7). The distribution appears to be a composite of a Gaussian and another distribution. This leads to characteristic shifts when calculating conventional average values to determine a reference (mean) model. However, to be able to derive a reference model, Ryberg & Haberland (2018) suggested a modified averaging procedure (in the following we use the term modified averages). Essentially, only velocity value samples with a probability exceeding the prior distribution are taken into account to calculate averages and standard deviations. Details can be found in Ryberg & Haberland (2018).

Another problem when determining earthquake hypocentres might occur when a hypocentre approaches the model boundaries (e.g. close to the surface). Their posterior distribution might be clipped by the model boundaries and a conventional averaging procedure will result in biased values. Although not done in the present analysis, this could be overcome by fitting a truncated Gaussian distribution function with its respective mean and standard deviation.

As seen in Fig. 3, the synthetic hypocentres are precisely recovered (synthetic and recovered positions almost plot on top of each other, with bars indicating the uncertainty (1σ) almost invisible at this map scale). A closer look at the differences between the synthetic and the recovered values (histograms in Fig. 5) indicates that this is in particularly true for the epicentral coordinates and the station corrections (the differences between synthetic and recovered parameters have almost zero mean and standard deviations of ∼250 m and ∼0.04 s, respectively). The depths and origin times are somewhat less well recovered (also almost zero mean but with larger standard deviations of ∼600 m and ∼0.2 s, respectively) hence indicating a trade-off of the latter parameters.

The modified averages of all hypocentres (after completion of the burn-in phase; only every 2000th analysed) seem to be a good measure for the 'final' (reference) hypocentre positions since they coincide to a very large extent with the synthetic (input) locations. Fig. 8 (left-hand panel) shows the convergence of the hypocentre positions towards the synthetic locations for 10 arbitrarily selected chains. The heat map in Fig. 8 (right-hand panel) shows the peaked distribution of all hypocentre positions after the burn-in phase, with the average hypocentre location hitting the synthetic location while the depths are more scattered hence indicating a larger uncertainty of the depth coordinate. The uncertainty of the depth estimates is obviously depth-dependent and in the range between 700 and 1500 m (Fig. 6). The epicentral uncertainty (1σ) for our synthetic data set is better than 500 m.

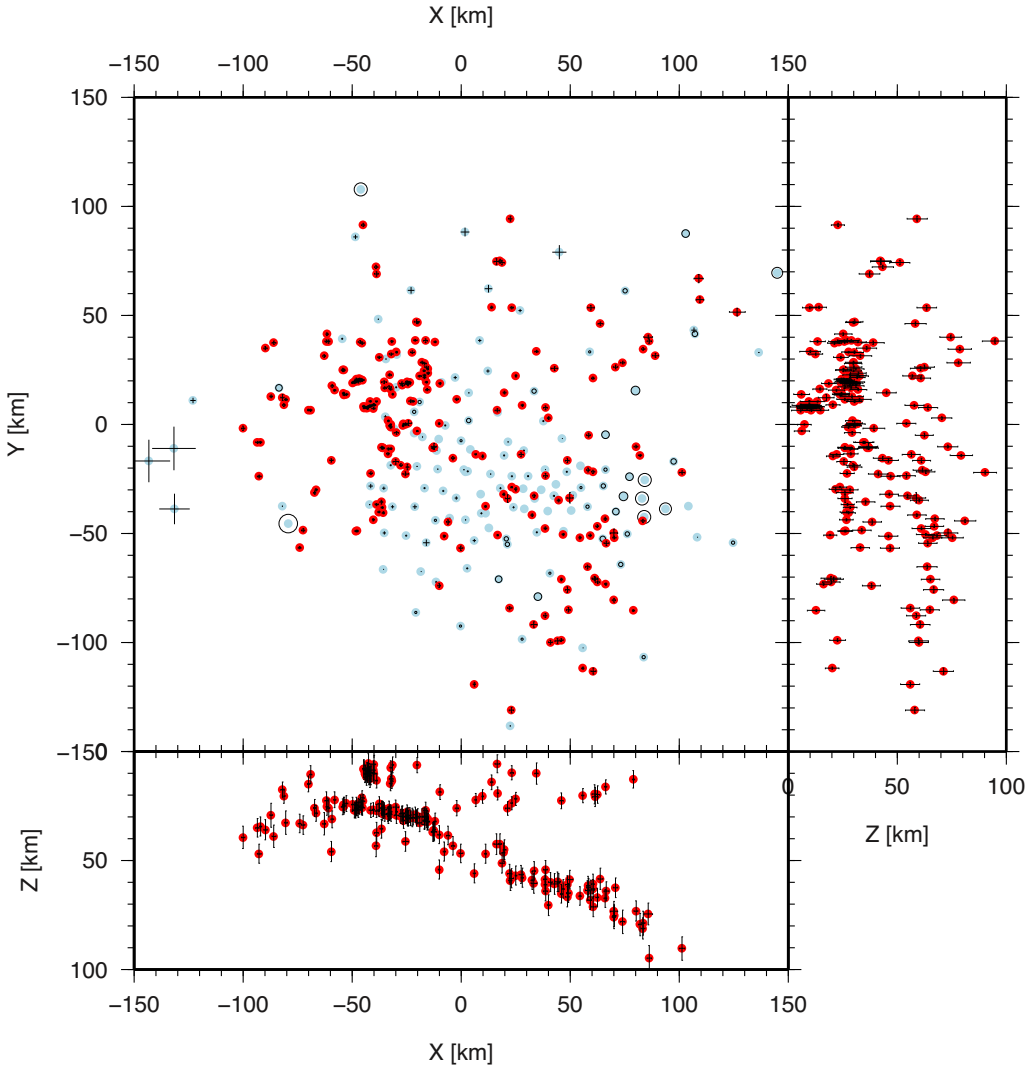

Figure 9. Distribution of earthquakes (red circles with uncertainties) and stations (light blue circles) derived from the inversion of the real data. Station corrections are indicated by open circles (negative corrections) and crosses (positive corrections). Note that the depth uncertainty of earthquakes is significantly larger than for the epicentral locations.

Also the synthetic velocities are well resembled by the averages of all inferred velocity models (post burn in; only every 2000th model considered), and the standard deviations of the velocities of all models after the burn-in phase are in the order of 0.2 km s^{-1} down to a depth of around 60 km (Fig. 4). Very large standard deviations below 60 km indicate the depth limit of resolution.

The average velocities (red line in Fig. 4) show a smooth increasing trend nicely resembling the synthetic model (black line in Fig. 4) while the heat-map plot of Fig. 4 indicates rather a step-like behaviour. Fig. 7 shows examples where the posterior distribution of *Vp* can be multi-modal and is typically non-Gaussian. As discussed above, instead of showing the conventional average velocity values, a modified averaging procedure was used.

4 APPLICATION TO REAL DATA

In a second example we apply our method to real data from a temporary, amphibious, local seismic network deployed 2004/2005 in South-Central Chile. The network, which ran for ∼1 yr, consisted of up to 130 seismic stations, 10 of them OBS/OBH stations. These stations were deployed in a region of ∼200 by 300 km at the western South American margin south of the City of Concepción. Average station spacings were in the order of 7 km in the central part of the network, decreasing to around 40 km at the periphery. In the current study we used a high-quality subset of 220 local earthquakes with a total of 8988 P and 5766 S picks which were manually determined. We only used events with GAP $\leq 180^\circ$ and at least 10 P and 5 S observations. The maximum hypocentre depth was ∼120 km and multiple events at the same time were removed. The observed earthquakes were located at the plate interface between the oceanic Nazca Plate and the South American continent, within the subducting plate (intermediate depth seismicity) as well as within the continental crust, mainly along continental transverse faults. More details of the network and the data set can be found in Haberland *et al.* (2006) and Haberland *et al.* (2009).

The McMC search of the real data was conducted in the same way as for the synthetic data (see above). The detailed results are shown in Figs 9–11. The uncertainties of the epicentres are small ($\sigma \sim 750$ m) while, similar to the synthetic example, in depth they are much larger and reach 3.5 km for the real data set (Fig. 9).

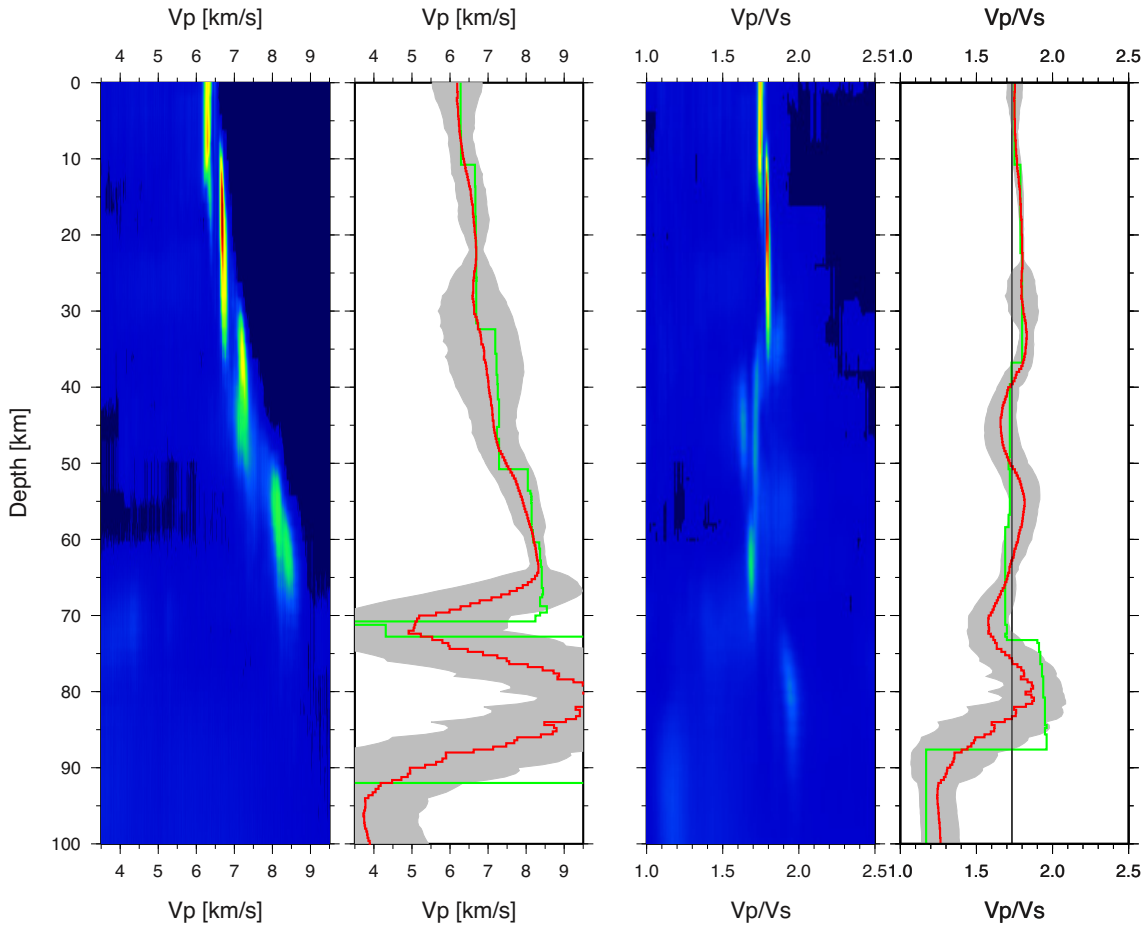

Figure 10. Recovered *Vp* and *Vp*/*Vs* models (real data example), similar to Fig. 4. Note that the velocity models are well recovered down to ∼65 km depth (indicated by small standard deviations).

For comparison, Haberland *et al.* (2006) gave location uncertainty estimates based on relocations of shots with known locations and origin times (see also below) of 1 km horizontally and 500 m vertically. The derived pick noise is ∼0.27 s for P-picks and ∼0.34 s for S-picks. The heat map and the maximum probability velocity models show a step-like behaviour with roughly 4 layers for *Vp* with increasing velocity from 6.3 to 8.2 km s^{-1}. Not unexpectedly, the average *Vp* values have a much smoother course with σ in the range between 0.25 and 0.75 km s^{-1} down to a depth of 65 km. Much larger σ values at depths greater than 65 km depict the resolution limit. *Vp*/*Vs* ratios are around 1.75 at depths shallower than 40 km, slightly decreasing toward greater depth. On average, the algorithm tends to need ∼12 layers to appropriately fit the real data. In Haberland *et al.* (2006) the parametrization of the (final) model was set to 18 layers.

Overall, the results agree very well with the previous results from Haberland *et al.* 2006 derived with the *velest* program (Kissling *et al.* 1994; see Fig. 11). The epicentral differences between the McMC locations and *velest* locations have almost zero mean and a σ of 2.0 km in the *x*-direction and 1.0 km in the *y*-direction. Depths are systematically shifted by roughly 3 km toward greater depth with a σ of ∼ 3.7 km. Station corrections are perfectly matched, and also the origin times correspond very nicely.

As an additional test we inverted arrival time data of seismic borehole blasts from a concurrent controlled source experiment (see Haberland *et al.* 2006 for details) with known locations and origin times. The shots were recorded by the seismological network. We kept the previously determined velocity model and station corrections determined with the McMC method fixed and inverted the arrival times of the seismic waves (mainly *P*-arrival times) for the source locations only. Fig. 12 shows the differences between the recovered and known shot coordinates. As in a similar relocation test using *velest* (Haberland *et al.* 2006), the shot locations could be well recovered with a few hundred meters accuracy and with a slight tendency to larger source depths. The latter could be related to our relatively simple approach to account for the effect of station elevations on the travel time calculation in our McMC routine in the current version (see above), which is probably not ideal for very shallow sources.

5 DISCUSSION AND CONCLUSIONS

With our method we closely follow the general approach to simultaneously derive hypocentre coordinates, origin times, station corrections and velocity structure as it has been proposed for example by Thurber (1983) for the 3-D case and Kissling *et al.* (1994) for the 1-D case. The simultaneous inversion for structure and hypocentres is required because the source location and the velocity structure are inherently linked to each other (hypocentre – velocity coupling, e.g. Kissling 1988, Thurber 1992; Kissling *et al.* 1994). Assuming a limited number of travel time observations, a limited aperture of the network, finite inter-station distances and noise (e.g. travel time

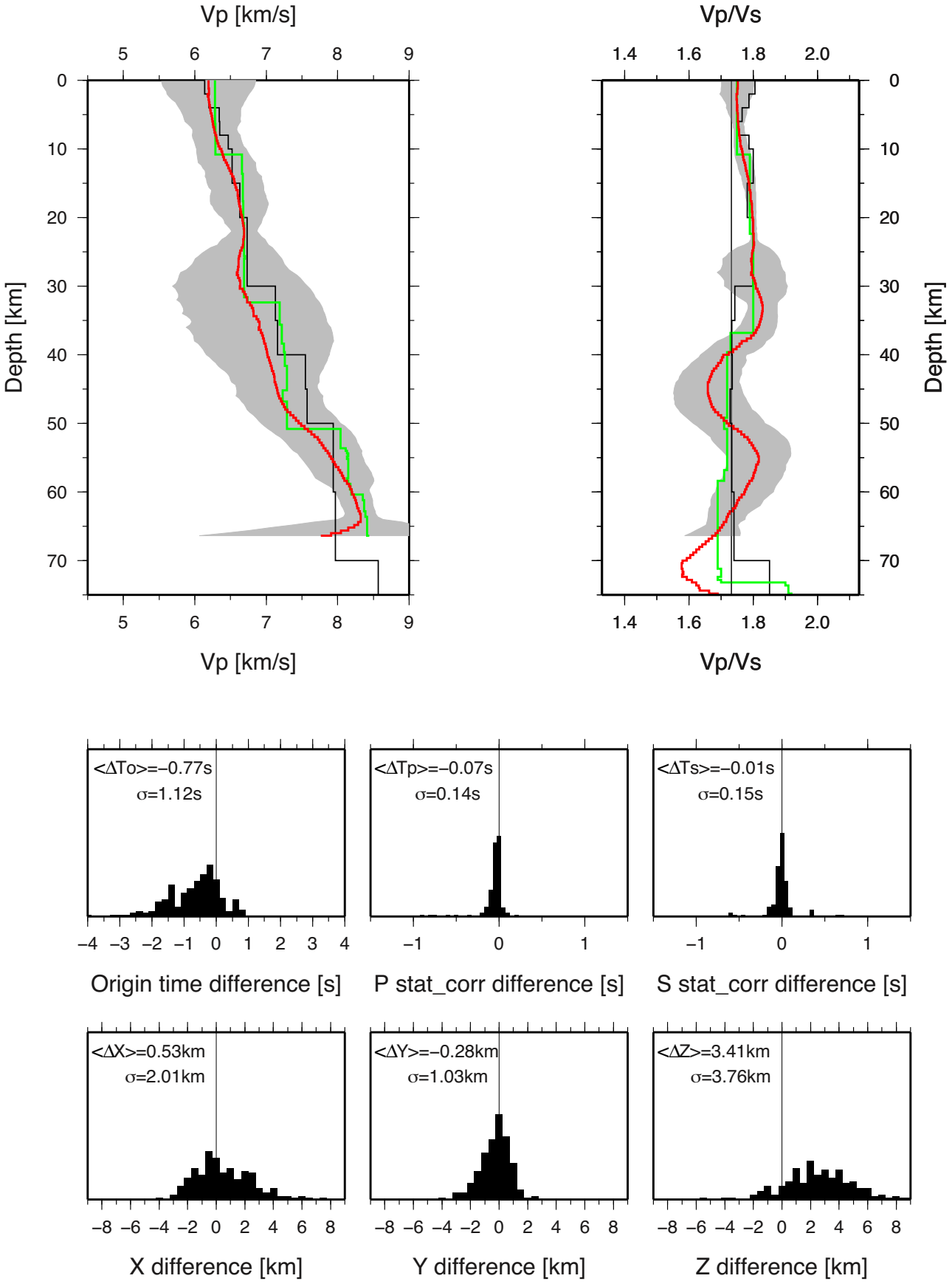

Figure 11. Comparison of McMC recovered velocity models with those derived by *velest* (top). The black line shows the *Vp* (*Vp*/*Vs*) model derived by the *velest* (Kissling *et al.* 1994) method. The red line represents the average *Vp* (*Vp*/*Vs*), the grey region shows the one σ uncertainty and the green line is the model with the maximum probability. The bottom panels show the distribution of the differences of the quake epicentral coordinates and depth. The middle diagrams show the differences of the quake origin times and station corrections.

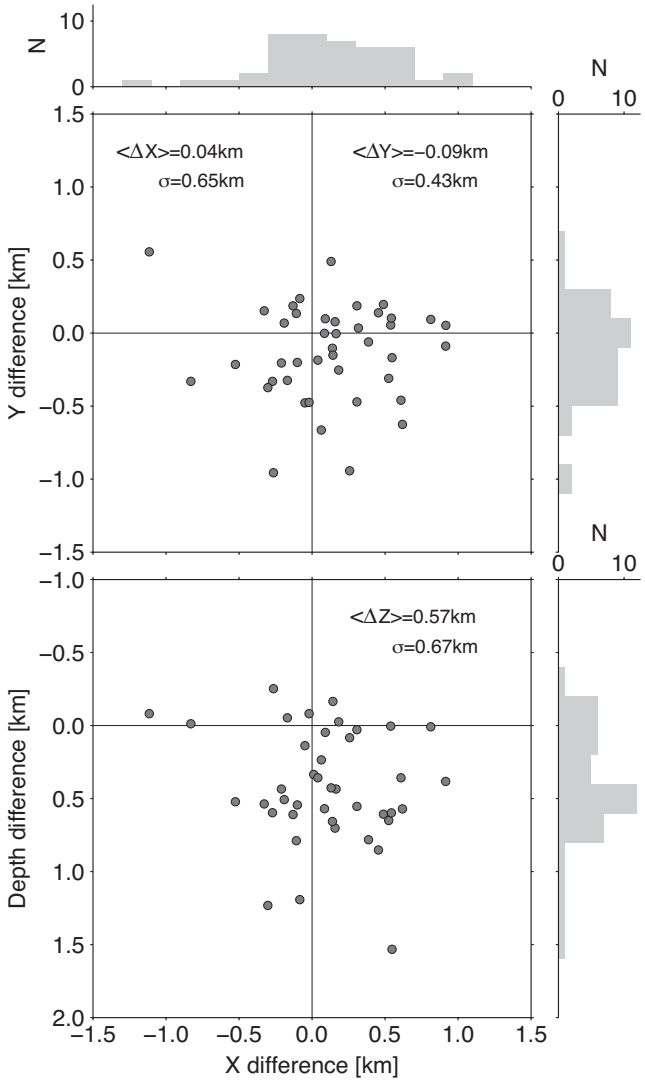

Figure 12. Misfits and misfit distributions (top and right) of recovered coordinates of 45 blasts with known locations and origin times. The velocity models and station corrections previously determined in Figs 9 and 10 were used to determine the event locations. While the misfits of epicentral coordinates are small the depth recovery is somewhat poorer and appears to be slightly shifted.

pick errors), earthquakes could be a bit deeper (or more distant or happen earlier) and velocities higher, and vice versa (within limits). In the 1-D case, the station corrections form an important ingredient since they account for all deviations of the 'real world' data from the 1-D simplification.

In contrast to the traditional inversions utilizing linearised and regularized formal inversions such as damped least squares we apply a hierarchical, transdimensional Monte Carlo Markov Chain algorithm. As shown in the synthetic example the algorithm performs very well and recovers—of course depending on the amount of synthetic noise added—the synthetic input model (hypocentre coordinates, origin times, velocity model and station corrections) precisely without making any assumption about starting velocity models and initial earthquake location. The entire McMC inversion procedure needs only very minor prior information: a very (unrealistic) wide range of velocity boundaries, upper and lower limits of station corrections and data noise levels. We carefully checked if these 'technical' boundaries might have an impact on our final models, but found no influence whatsoever.

The results of the McMC inversion of the real data are very similar to the results derived by a conventional (i.e. *velest*) inversion (Haberland *et al.* 2006). This is particularly true for the epicentral positions, whereas we notice a systematic shift of the focal depths towards slightly greater depths (∼2 km). The derived velocity models, in particular the reference model based on the maximum likely model (mode), correspond very well with the *velest* model in the upper part, whereas in the lower part (of the crust) the velocities derived by the McMC search are slightly smaller. This could be one reason for the larger earthquake depths derived by the McMC method. Moreover, the McMC velocity models reach mantle velocities at depths larger than 50–55 km (most visible in the reference model based on the maximum likely model (mode); Fig. 11). In the *velest* model these velocities were reached at greater depth (not well resolved). Overall, the results of the McMC inversion nicely confirm the reliability and quality of hypocentres determined by *velest*. We would like to point out that the hypocentre uncertainty derived in our McMC inversion takes into account the uncertainties of the velocity model.

It has to be mentioned that the *velest* locations had been carefully calculated (Haberland *et al.* 2006) including several location runs with a wide range of initial models (so-called 'minimum 1-D model'; see e.g. Kissling *et al.* 1984, 1994). Furthermore, the TIPTEQ data set was carefully quality-checked and selected, only incorporating earthquakes located within the network and with a relatively large number of high-quality observations per event (outliers removed). Studying the performance of our McMC method for data sets with very noisy data (which we expect for example for automatic picks), sparse data (sparse networks or very small events) or events outside the network will be important topics of future investigations. Furthermore, in contrast to the previous *velest* inversion we have not used the pick quality (weights) assigned to each individual pick. This could explain possible differences between the results from the two methods. Instead, we obtained the data noise from the McMC search. However, incorporating individual noise inversion for each quality class (or a similar consideration of the quality/weights in the McMC search) would be straight forward and will be targeted in the future.

We would like to emphasize that we strictly followed a minimum prior knowledge approach. Accordingly, we start our inversion runs with random initial values for almost all parameters (i.e. hypocentres, velocity model, station corrections, initial number of model points) and use wide bounds on the velocities (priors). We think that this approach makes full use of the McMC potential and yields the globally best fitting model(s). Of course, all this comes at a certain price: A typical McMC search for the data sets presented in this paper (∼200 earthquakes, ∼120 stations) takes about 8 hours (1000 chains on 1000 CPUs) when using a forward grid spacing of 2 km. Eventually, the successful relocation of the shots with known locations and origin times confirms the usefulness of the derived 1-D velocity model (and station corrections) and the proposed method.

Our future efforts will concentrate on extending our McMC method to 3-D velocity models thus yielding a full simultaneous inversion for hypocentres and 3-D velocity structure (commonly termed local earthquake tomography; e.g. the simul2000 software; Thurber 1983, 1993; Eberhart-Phillips 1993) using the hierarchical, transdimensional McMC approach (similar to Agostinetti *et al.* 2015). In the meantime, the results of the 1-D McMC search, both with respect to velocity models and hypocentral parameters, will

form well suited initial data (starting locations and initial 1-D model) for conventional tomographic inversions.

ACKNOWLEDGEMENTS

All computations were performed on the computer cluster of the GFZ Potsdam. The real data were collected within the TIPTEQ project financed by the German BMBF and DFG through the R&D program GEOTECHNOLOGIEN, grant 03G0594C. Data are available from the GEOFON archive, network code ZW (2004-2005). The GIPP (GFZ Potsdam) provided the seismic instruments (grant number GIPP200404). We would like to thank Eric Bergman, Jim Mechie and an anonymous reviewer for their thorough suggestions and comments which helped to improve the manuscript.

References

Agostinetti, N.P., Giacomuzzi, G. & Malinverno, A., 2015. Local three-dimensional earthquake tomography by transdimensional Monte Carlo sampling, *Geophys. J. Int.,* **201,** 1598–1617.

Aki, K. & Lee, W., 1976. Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes 1. A homogeneous initial model, *J. geophys. Res.,* **81**(23), 4381–4399.

Bai, C., Zhao, R. &Greenhalgh, 2009. Fast and accurate earthquake location within complex medium using a hybrid global-local inversion approach, *Earthq Sci.,* **22**(5), 435 .

Billings, S.D., Kennett, B.L.N. & Sambridge, M.S., 1994. Hypocentre location: genetic algorithms incorporating problem-specific information, *Geophys. J. Int.,* **118**(3), 693–706.

Bodin, T., Sambridge, M., Rawlinson, N. & Arroucau, P., 2012a. Transdimensional tomography with unknown data noise, *Geophys. J. Int.,* **189**(3), 1536–56.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K. & Rawlinson, N., 2012b. Transdimensional inversion of receiver functions and surface wave dispersion, *J. geophys. Res.,* **117,** B02301, doi:10.1029/2011JB008560.

Douglas, A., 1967. Joint hypocenter determination, *Nature,* **215,** 47–48.

Eaton, J.P., 1970. HYPOLAYR, a computer program for determining hypocenters of local earthquakes in an earth consisting of uniform flat layers over a half space, U.S. Geological Survey, Open-File Report 69–85, pp. 155.

Eberhart-Phillips, D., 1986a. Three-dimensional structure in northern California coast ranges from inversion of local earthquake arrival times, *Bull. seism. Soc. Am.,* **76**(4), 1025–1052.

Eberhart-Phillips, D., 1993. Local earthquake tomography: earthquake source regions, in *Seismic Tomography: Theory and Practice,* eds Iyer, H.M. & Hirahara, K., Springer.

Eberhart-Phillips, D., Christensen, D.H., Brocher, T.M., Hansen, R., Ruppert, N.A., Haeussler, P.J. & Abers, G.A., 2006. Imaging the transition from Aleutian subduction to Yakutat collision in central Alaska, with local earthquakes and active source data, *J. geophys. Res.,* **111,** B11303, doi:10.1029/2005JB004240.

Eberhart-Phillips, D. & Michael, A.J., 1998. Seismotectonics of the Loma Prieta, California, region determined from three-dimensional Vp, Vp/Vs, and seismicity, *J. geophys. Res.,* **103**(B9), 21 099–21 120.

Eberhart-Phillips, D., Reyners, M., Chadwick, M. & Chiu, J.-M., 2005. Crustal heterogeneity and subduction processes: 3-D Vp, Vp/Vs and Q in the southern North Island, New Zealand, *Geophys. J. Int.,* **162,** 270–288.

Evans, J., Eberhart-Phillips, D. & Thurber, C.H., 1994. User's manual for simulps12 for imaging Vp and Vp/Vs: a derivative of the "Thurber" tomographic inversion simul3 for local earthquakes and explosions, U.S. Geol. Surv. Open File Rep., 94–431.

Geiger, L., 1910. Herdbestimmung bei Erdbeben aus den Ankunftszeiten. Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, *Mathematisch-Physikalische Klasse,* **1910,** 331–349.

Geiger, L., 1912. Probability method for the determination of earthquake epicenters from the arrival time only (translated by Peebles, F. W. L., and Corey, A. H. from Geiger's 1910 German article), *Bull. St. Louis Univ.,* **8**(1), 56–71.

Gesret, A., Desassis, N., Noble, M., Romary, T. & Maisons, C., 2015. Propagation of the velocity model uncertainties to the seismic event location, *Geophys. J. Int.,* **200,** 52–66.

Haberland, C., Rietbrock, A., Lange, D., Bataille, K. & Dahm, T., 2009. Structure of the seismogenic zone of the southcentral Chilean margin revealed by local earthquake travel time tomography, *J. geophys. Res.,* **114,** B01317, doi:10.1029/2008JB005802.

Haberland, C., Rietbrock, A., Lange, D., Bataille, K. & Hofmann, S., 2006. Interaction between forearc and oceanic plate at the southcentral Chilean margin as seen in local seismic data, *Geophys. Res. Lett.,* **33,** L23302, doi:10.1029/2006GL028189.

Husen, S., Kissling , E., , Flueh , E. & Asch , G., 1999. Accurate hypocentre determination in the seismogenic zone of the subducting Nazca Plate in northern Chile using a combined on/offshore network, *Geophys. J. Int.,* **138,** 687–701.

Husen, S. & Kissling, E., 2001. Postseismic fluid flow after the large subduction earthquake of Antofagasta, Chile, *Geology,* **29**(9), 847–850.

Husen, S., Kissling, E., Deichmann, N., Wiemer, S., Giardini, D. & Baer, M., 2003. Probabilistic earthquake location in complex three-dimensional velocity models: application to Switzerland, *J. geophys. Res.,* **108**(B2), 2077.

Kissling, E., 1988. Geotomography with local earthquake data, *Rev. Geophys.,* **26,** 659–698.

Kissling, E., Ellsworth, W.L. & Cockerham, R., 1984. Three-dimensional structure of the Long Valley Caldera, California, region by geotomography, U.S. Geol. Surv. Open File Rep. 84–939, 188–220.

Kissling, E., Ellsworth, W.L., Eberhart-Philips, D. & Kradolfer, U., 1994. Initial reference models in local and regional earthquake tomography, *J. geophys. Res.,* **99,** 19 635–19 646.

Lahr, J.C., 1989. HYPOELLIPSE/Version 2.0: a computer program for determining local earthquakes hypocentral parameters, magnitude, and first motion pattern, U.S. Geological Survey Open-File Report 89–116, 92.

Lee, W.H.K. & Lahr, J.C., 1975. HYPO71 (revised): a computer program for determining hypocenter, magnitude and first motion pattern of local earthquakes, U.S. Geological Survey Open-File Report 75–311, pp. 116.

Lienert, B.R.E., Berg, E. & Frazer, L.N., 1988. HYPOCENTER: an earthquake location method using centered, scaled, and adaptively least squares, *Bull. seism. Soc. Am.,* **76,** 771–783.

Lienert, B.R.E. & Havskov, J., 1995. A computer program for locating earthquakes both locally and globally, *Seism. Res. Lett.,* **66,** 26–36.

Lomax, A., 2005. A reanalysis of the hypocentral location and related observations for the great 1906 California earthquake, *Bull. seism. Soc. Am.,* **95,** 861–877.

Lomax, A., Michelini, A. & Curtis, A., 2009. Earthquake location, direct, global search methods, in *Encyclopedia of Complexity and Systems Science: Assembles for the First Time the Concepts and Tools for Analyzing Complex Systems in a Wide Range of Fields,* 1 edn, pp. 2449–2473, Springer-Verlag New York, Inc.

Lomax, A., Virieux, J., Volant, P. & Berge, C., 2000. Probabilistic earthquake location in3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations, in *Advances in SeismicEvent Location,* pp. 101–134, eds Thurber, C.H. & Rabinowitz, N., Kluwer.

Metropolis, N., Rosenbluth, M.N., Rosenbluth, A.W., Teller, A.H. & Teller, E., 1953. Equation of state calculations by fast computing machines, *J. Chem. Phys.,* **21**(6), 1087–1092.

Mosegaard, K., 1998. Resolution analysis of general inverse problems through inverse Monte Carlo sampling, *Inverse Problems,* **14**(3), 405–426.

Myers, S.C., Johannesson, G. & Hanley, W., 2007. A Bayesian hierarchical method for multiple-event seismic location, *Geophys. J. Int.,* **171**(3), 1049–1063

Podvin, P. & Lecomte, I., 1991. Finite difference computation of travel times in very contrasted velocity models: a massively parallel approach and its associated tools, *Geophys. J. Int.,* **105**(1), 271–284.

Pujol, J., 2000. Joint event location – the JHD technique and applications to data from local networks, in *Advances in Seismic Event Location. Modern Approaches in Geophysics,* eds Thurber, C.H. & & Rabinowitz, N., **Vol. 18,** pp. 163–204, Springer.

Pujol, J., 2003. Software for joint hypocentral determination using local events, *Int. Geophys.,* **81,** 1621–1623.

Ružek, B. & Kvasnička, M., 2001. Differential evolution algorithm in earthquake hypocenter location, *Pure appl. Geophys.,* **158,** 667–693.

Ryberg, T. & Haberland, C., 2018. Bayesian inversion of refraction seismic travel-time data, *Geophys. J. Int.,* **212**(3), 1645–1656.

Sambridge, M.S. & Kennett, B.L.N., 1986. A novel method of hypocentre location, *Geophys. J. R. astr. Soc.,* **87,** 679–697.

Schott, J.-J., Roussignol, M., Menvielle, M. & Nomenjahanary, F.R., 1999. Bayesian inversion with Markov chains—II. The one-dimensional DC multilayer case, *Geophys. J. Int.,* **138**(3), 769–783.

Shen, W., Ritzwoller, M.H., Schulte-Pelkum, V. & Lin, F., 2013. Joint inversion of surface wave dispersion and receiver functions: a Bayesian Monte-Carlo approach, *Geophys. J. Int.,* **192**(2), 807–836.

Theunissen, T. *et al.*, 2018. Absolute earthquake locations using 3-D versus 1-D velocity models below a local seismic network: example from the Pyrenees, *Geophys. J. Int.,* **212**(3), 1806–1828.

Thurber, C., 1983. Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, *J. geophys. Res.,* **88**(B10), 8226–8236.

Thurber, C., 1992. Hypocenter-velocity structure coupling in local earthquake tomography, *Phys. Earth planet. Inter.,* **75**(1–3), 55–62.

Thurber, C., 1993. Local earthquake tomography: velocities and Vp/Vs-theory, in *Seismic Tomography: Theory and Practice,* pp. 563–583, eds Iyer, H. & Hirahara, K., CRC Press.

Tryggvason, A. & Bergman, B., 2006. A travel time reciprocity discrepancy in the Podvin and Lecomte time3d finite difference algorithm, *Geophys. J. Int.,* **165,** 432–435.