P, S velocity and V_P/V_S ratio beneath the Toba caldera complex (Northern Sumatra) from local earthquake tomography

Ivan Koulakov,[1,2] Tedi Yudistira,[3] Birger-G. Luehr[2] and Wandono [4]

[1]*Institute of Petroleum Geology and Geophysics, SB RAS, Prospekt Akademika Koptuga,* 3, *Novosibirsk,* 630090, *Russia. E-mail: KoulakovIY@ipgg.nsc.ru*
[2]*GeoForschungsZentrum Potsdam, Telegrafenberg,* 14473 *Potsdam, Germany*
[3]*Department of Geophysics, Institut Teknologi Bandung, Indonesia*
[4]*Meteorological and Geophysical Agency (BMG), Jakarta, Jl. Angkasa* I/2, *Jakarta* 10720, *Indonesia*

SUMMARY
In this paper, we investigate the crustal and uppermost mantle structure beneath Toba caldera, which is known as the location of one of the largest Cenozoic eruptions on Earth. The most recent event occurred 74 000 yr BP, and had a significant global impact on climate and the biosphere. In this study, we revise data on local seismicity in the Toba area recorded by a temporary PASSCAL network in 1995. We applied the newest version of the LOTOS-07 algorithm, which includes absolute source location, optimization of the starting 1-D velocity model, and iterative tomographic inversion for 3-D seismic *P, S* (or the V_P/V_S ratio) and source parameters. Special attention is paid to verification of the obtained results. Beneath the Toba caldera and other volcanoes of the arc, we observe relatively moderate (for volcanic areas) negative *P*- and *S*-velocity anomalies that reach 18 per cent in the uppermost layer, 10–12 per cent in the lower crust and about 7 per cent in the uppermost mantle. Much stronger contrasts are observed for the V_P/V_S ratio that is a possible indicator of dominant effect of melting in origin of seismic anomalies. At a depth of 5 km beneath active volcanoes, we observe small patterns (7–15 km size) with a high V_P/V_S ratio that might be an image of actual magmatic chambers filled with partially molten material feeding the volcanoes. In the mantle wedge, we observe a vertical anomaly with low *P* and *S* velocities and a high V_P/V_S ratio that link the cluster of events at 120–140 km depth with Toba caldera. This may be an image of ascending fluids and melts released from the subducted slab due to phase transitions. However, taking into account poor vertical resolution, these results should be interpreted with prudence. Although the results show clear signatures that are quite typical for volcanic areas (low velocity and high V_P/V_S ratio beneath volcanoes), we do not observe any specific features in seismic structure that could characterize Toba as a super volcano.

Key words: Seismic tomography; Volcano seismology; Volcanic arc processes; Magma chamber processes; Calderas; Indian Ocean.

1 INTRODUCTION

The Toba volcanic complex, located in northern Sumatra, Indonesia, is part of a 5000 km long volcanic chain along the Sunda arc (Fig. 1). Toba volcano produced the largest known volcanic eruption on Earth during the past 2 Myr (Smith & Bailey 1968). About 74 000 yr ago, around 2800 km³ of magma were erupted. The eruption led to the final formation of one of the largest calderas, the 35 × 100 km wide Toba caldera. Super scale eruptions at Toba have occurred several times (at least four eruptions of more than VEI 7 over the last 2 Myr). In this sense, Toba seems to be a singularity in a chain of more than 100 other volcanoes with explosive potential along the Sunda arc. The reasons for this unique behaviour of the Toba volcanic activity are not yet clearly understood. We believe that understanding the mechanisms of the origin of this supervolcano will only be possible by investigating the entire subduction complex beneath Toba and the surrounding areas. In addition, comparison with subduction complexes from other areas of normal and supervolcanism can reveal specific features that distinguish Toba from other volcanoes.

Supervolcano eruption with a Volcanic Explosivity Index of 8 (VEI-8) is a colossal event that throws out at least 1000 km³ Dense Rock Equivalent (DRE). Besides Toba, there are at least two other known eruptions of VEI-8 that occurred during the last 1 000 000 yr: recent Yellowstone eruptions in the USA and the Oruanui eruption in Taupo volcanic zone in New Zealand. Within the Yellowstone area during the past 3 Myr, there were at least two VEI-8 eruptions: ∼2.2 Ma (2500 km³ of erupted material) and ∼640 000 yr ago (1000 km³). Since the last cataclysmic eruption at least 30 dominantly rhyolitic and basaltic flows, as young as 70 000 yr old, have been erupted, covering much of Yellowstone (Christiansen 2001).
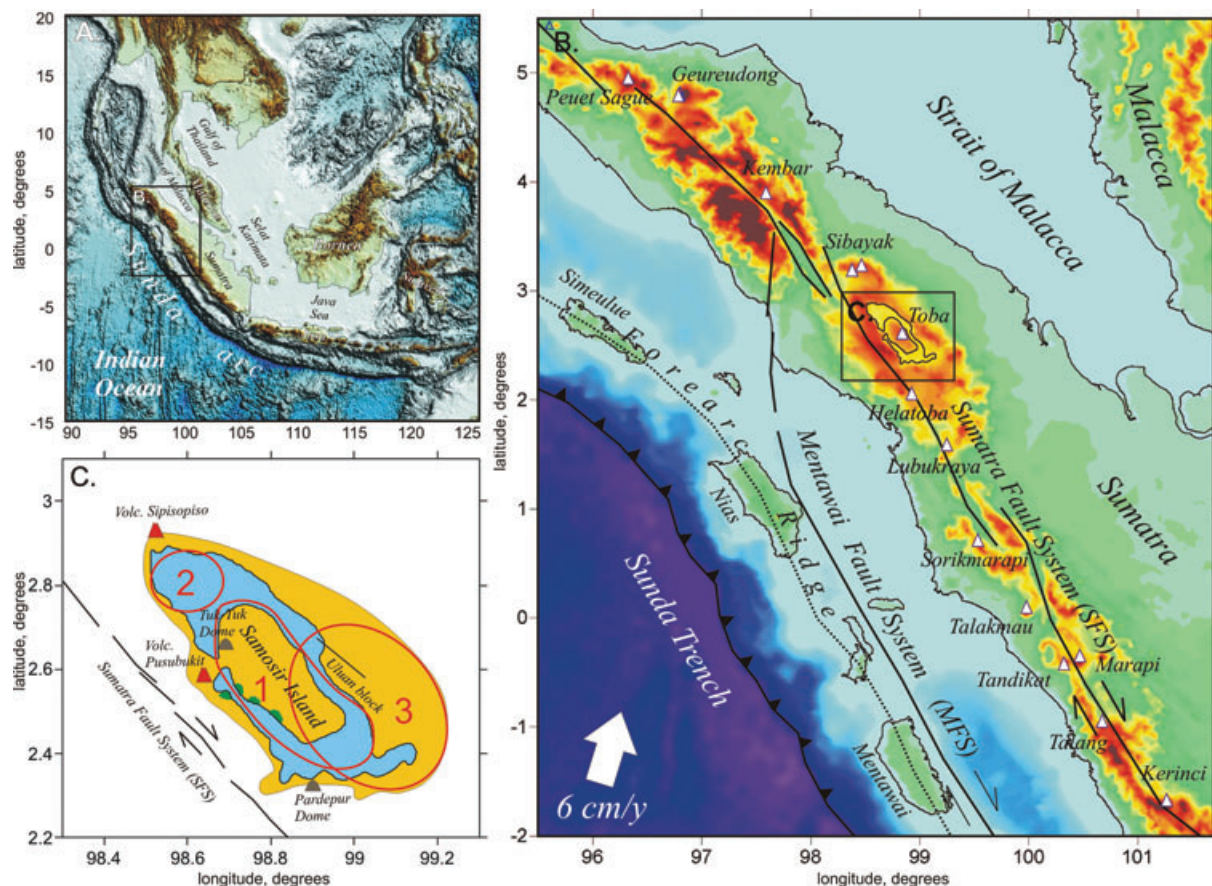
Figure 1. Study region in the context of the Sunda Arc. (a) Bathymetry and topography of the Sunda Arc and surrounding areas. (b) The main geological units of Northern Sumatra. The volcanoes shown by triangles are from (Simkin & Siebert 1994). (c) The main structural elements of Toba caldera. Yellow areas are the present-day topographic depression; green indicates uplifted areas. Red ellipses mark different caldera units. (1) Sibandung caldera: created 74 000 yr ago by the Toba YTT event (Young Toba Ash); (2) Haranggaol caldera: formed 500 000 yr ago by the Toba MTT event (Middle Toba Ash); (3) Sibandung caldera: formed 800 000 yr ago by the Toba OTT event (Old Toba Ash).

This area is one of the key American geological polygons that attracts attention of hundreds of specialists from different domains of geology and geophysics. Extensive multidisciplinary studies have provided much information about the deep state of the Yellowstone area. In particular, the seismic structure beneath the Yellowstone caldera has been investigated using various seismic schemes, such as teleseismic (e.g. Hadley *et al.* 1976; Humphreys *et al.* 2000) and local earthquake tomography (LET) (e.g. Miller & Smith 1999; Husen *et al.* 2004; Waite *et al.* 2006). All these studies reveal in the crust a large body with low velocities and a high V_P/V_S ratio. It is interpreted as an area of crystallizing magma with high content of gas and liquid fluids. However, the origin of Yellowstone super volcanism seems to be different of that of Toba, because it is located relatively far from the subduction zones. In many papers the origin of the Yellowstone volcanism is associated with the mantle plume (e.g. Waite *et al.* 2006), although this concept is actively debated by different authors (e.g. Humphreys *et al.* 2000).

The Oruanui super eruption in New Zealand is more similar to the Toba event. It took place in the Taupo volcanic zone about 26500 yr ago and resulted in ejection of ∼1170 km^3 of volcanic materials (e.g. Harrison & White 2004, Wilson *et al.* 2006). Although this amount is only slightly less than erupted in Toba, the Oruanui event did not have such a global effect, probably due to longer and less intensive regime of eruption. This area was intensively investigated during the last decades using different geophysical methods

such as gravity and magnetic modelling (Rogan 1982; Soengkono 1994), receiver function (Bannister *et al.* 2004) and LET (Sherburn *et al.* 2003; Reyners *et al.* 2006). Similar to the Yellowstone area, the tomographic results demonstrate low velocities and a high V_P/V_S ratio beneath the caldera, which can be an indicator for high content of fluids and melts.

It is quite obvious that the expected features beneath the Toba caldera should be the same as in the Yellowstone and Taupo volcanic zone: low velocity and a high V_P/V_S ratio. Furthermore, it is easy to predict that the large caldera filled with a thick layer of relatively recent sediments should be expressed as low-velocity body in the uppermost part of the tomogram. At the same time, the quantitative information about size, shape and intensity of the anomalies in the crust and uppermost mantle beneath the Toba caldera is necessary to single out the signatures of the previous event and assess the perspective of new eruptions. As suggested by Bachmann & Bergantz (2008), the explosive magma in super eruptions are produced by extracting interstitial liquid from long-lived 'crystal mushes' (magmatic sponges containing >50 vol per cent of crystals) and collecting it in unstable liquid-dominated lenses. If there is a potential of a new super eruption, such signatures should be clearly seen in tomograms.

In contrast to two other areas of super volcanism mentioned above, the geophysical investigations of the Toba area are much less intensive. The distribution of local earthquakes beneath Toba was

previously studied in Fauzi *et al.* (1996) based on local seismicity data recorded by a temporarily passive PASSCAL seismic experiment in 1995. The same data were used by Masturyono *et al.* (2001) to perform tomographic inversion resulted at *P*-velocity structure in the crust and the uppermost mantle beneath the Toba caldera. The undoubted advantage of this study was the combined analysis of seismic and gravity data. However, we suppose that this model is not sufficiently trustworthy, and it must be revised. First of all, Masturyono *et al.* (2001) used only *P* arrivals. Without using the *S* data, the stability of source locations is significantly lower, which reduces the reliability of the velocity model. Second, their solution is strongly affected by parameterization grid. The cell spacing used is too large, and, as a result, the solution oscillates from a cell to cell. For example, a strong negative anomaly in one cell may coexist with a positive anomaly in a neighbouring cell, which in its turn is conterminous with another negative anomaly in the next cell. It is clear that in this situation when one uses a grid with other parameters (e.g. shifted to a half step, or having slightly different spacing, or rotated grid), he or she would produce completely different images. Third, it is unclear how they selected damping parameters. The reported 37 per cent of *P*-velocity anomaly seems to be unrealistic and probably due to instability of the underdamped inversion.

In this study, we consider the data of the same PASSCAL experiment in the Toba area that were supplemented with *S* picks. Here we use another tomographic algorithm, LOTOS-07 (Koulakov 2009a), which has some important features compared to the code used for obtaining the previous results, such as quasi-continuous parameterization, 1-D velocity optimization and a more effective algorithm for source location in a 3-D model. It offers effective and unbiased ways for verification of the obtained results based on synthetic modelling and other tests. In this study, we pay special attention to studying the reliability of the results and the quantitative evaluation of the derived parameters. In particular, in this study, we use a new technique for estimating the amplitudes of anomalies. Based on this technique, we show that in the crust beneath Toba, the *P*-velocity anomaly does not exceed 16 per cent. This value is much lower than that predicted by Masturyono *et al.* (2001).

2 GEOLOGICAL OVERVIEW

The super large eruptions of Toba should be considered in the context of the entire subduction complex beneath Sumatra. Sumatra Island is a northwest trending physiographic expression located on the western edge of Sundaland, a southern extension of the Eurasian Continental Plate (Figs 1a and b). Sumatra Island has an area of about 435 000 km^2, and a width of about 100–200 km in the northern part and about 350 km in the southern part. The main geographical trend lines of the island are rather simple. Its backbone is formed by the Barisan Range, which runs along the western side and constitutes the active volcanic arc and divides the west and the east coasts. The distance between the Sunda trench and the Sumatran coast is about 200 km (Fig. 1b). In the offshore, a chain of islands form the Forearc ridge, a local high between the trench and the Sumatra coast. The island of Sumatra is interpreted to have been constructed by collision and suturing of discrete micro continents in the late Pre-Tertiary times (Pulunggono & Cameron 1984). At present, the Indo-Australian Ocean Plate is being subducted beneath the Eurasian Continental Plate in a N20°E direction at a rate of about 5.5 cm yr^{-1} (e.g. DeMets *et al.* 1990; Fig. 1b). This zone is marked by the active Sunda Arc–Trench system (Fig. 1a) that extends for more than 5000 km, from Burma in the north to where the Australian

Plate is colliding with Eastern Indonesia in the south. The input of sediments mainly comes from the Ganges River, India, leading to the development of the largest accretionary wedge at a subduction zone on Earth (e.g. Steckler *et al.* 2008). The oblique subduction of the Indian plate is responsible for strike-slip fault systems along Sumatra [e.g. Mentawai Fault System (MFS), Sumatra Fault System (SFS)].

In 1949, van Bemmelen (1949) reported that Lake Toba was surrounded by a vast layer of ignimbrite rocks. Later researchers found rhyolite ash in Malaysia similar to the ignimbrites around Toba, as well as 3000 km away in India (Aldiss & Ghazali 1984), and on the seafloor of the eastern Indian Ocean and the Bay of Bengal. These observations show that the Toba eruption, dated at 74 000 yr ago, was the most recent truly large eruption on Earth during the last several millions years. All information on the extent of the erupted Toba volcanic material has been compiled in Rose & Chesner (1987), Chesner & Rose (1991) and Chesner *et al.* (1991). According to estimates in these studies, the total amount of erupted material was about 2800 km^3. About 800 km^3 was ignimbrite that travelled swiftly over the ground away from the volcano, and the remaining 2000 km^3 fell as ash, with the wind blowing most of it to the west. Such a huge eruption probably lasted nearly 2 weeks. Ninkovich *et al.* (1978) estimated the height of the eruption column to have been 50–80 km. Rose & Chesner (1987), after a study of the shapes of the ash shards, concluded this estimate was too high by a factor of 5 or more. This event, called the eruption of the Youngest Toba Tuff (YTT), was responsible for the collapse structure of the Toba caldera visible today (Van Bemmelen 1949). Besides, the most recent large eruption of 73 000 yr ago, during the past 1.2 Myr, there have been at least three other ash flow tuff eruptions from the caldera complex (Chesner & Rose 1991; Chesner *et al.* 1991). The older Toba units are the Middle Toba Tuff (MTT; age 0.50 Myr, Chesner *et al.* 1991), the Oldest Toba Tuff (OTT; age 0.84 Myr, Diehl *et al.* 1987) and the Haranggoal Dacite Tuff (HDT; age 1.2 Myr, Nishimura *et al.* 1977). These were erupted alternately from northern and southern vent areas in the present caldera (Chesner & Rose 1991).

The eruption of such a huge amount of volcanic rock caused a large collapse, resulting in a huge caldera. This caldera filled with water and created Lake Toba (Fig. 1c). The Toba caldera, the largest known Cenozoic caldera in the world, has a size of 30 × 100 km and a total relief of 1700 m. After the YTT eruption, resurgent doming formed the massive Samosir Island and Uluan Peninsula structural blocks. Lake sediments on Samosir indicate an uplift of at least 450 m. Additionally, post-YTT eruptions include a series of lava domes, the growth of the solfatarically active Pusubukit volcano on the southwestern margin of the caldera, and the formation of Sipisopiso volcano at the NW-most rim of the caldera. Lack of vegetation suggests that this volcano may be only a few hundred years old. There have been no eruptions documented for Toba in historical time, but the area has been seismically active.

The recent enormous Toba super eruption should probably leave some markers, such as low-velocity anomalies and high values of the V_P/V_S ratio in the crust and uppermost mantle. The main purpose of this paper is to detect them using the seismic tomography approach.

3 DATA DESCRIPTION

The seismic network around Toba was operated for about 4 months (January–May, 1995) by Indonesian teams in cooperation with IRIS
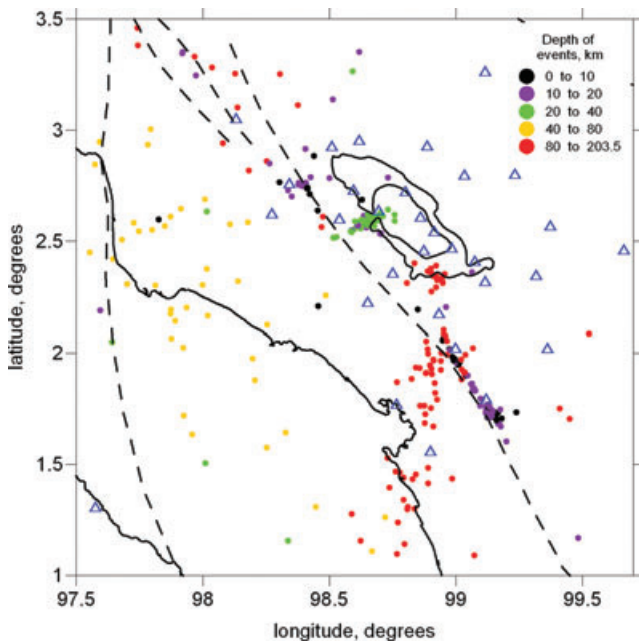
Figure 2. Distribution of stations (blue triangles) and events (coloured dots) used in this study. The colours of dots indicate the depth ranges of the events after the final locations. Dotted lines indicate the locations of the Mentawai and Sumatra fault zones.

and PASSCAL. The network comprised 30 short-period stations (three-component Mark Product L22C-3D) and 10 broad-band instruments (Guralp CNG-3ESP) covering an area of about 250 × 250 km. To obtain better control on the locations of the events in the offshore, one station was installed on Nias Island, west of North Sumatra. The network recorded ~1500 local earthquakes; however, for this study only the 390 most reliable events were used. The distribution of the stations and earthquakes used in this study is presented in Fig. 2.

   We selected events with a number of recorded phases at more than nine stations, which seems to be the optimum value for this data set. We considered cases of other values of this number and tested them with synthetic models. If fewer stations were chosen (e.g. seven), the total amount of data amount increased, but the trade-off between velocity and source parameters became more important. As a result of the joint inversion for seismic anomalies and velocity parameters, some artefacts appeared. On the other hand, for larger numbers of stations (e.g. 12), the total amount of data became much smaller, which lead to a significant loss of resolution. In total, we selected 390 events and the corresponding 3377 *P* and 2462 *S* rays.

   The LOTOS code, which is used here for inversion, does not require that sources be located inside the network of stations (having a GAP <180°), as in many tomographic studies. Koulakov (2009b) has shown that this requirement does not reflect the real importance of an event for tomographic inversion. Based on synthetic testing, he has demonstrated great importance of out-of-network events for the LET scheme. In LOTOS, we set the requirements of event selection less strictly than the GAP criterion. For example, an event is rejected if the lateral distance to the nearest station is more than 200 km.

   It should be mentioned that additional data for this region have been collected by Meteorological and Geophysical Agency (BMG), using Indonesian permanent seismic stations. However, the catalogue provided by the Agency contained only *P* arrivals, and number of picks for each single event was quite low. For most of events, the number of picks was less than 7, which is not sufficient for solving

the coupled inversion problem for velocity and source parameters. Therefore, these data were not used in this study.

4 ALGORITHM

4.1 General information

For tomographic inversion, we used the LOTOS-07 code, which is described in detail in Koulakov *et al.* (2007) and Koulakov (2009a), and is freely available on the web site www.ivan-art.com/science/LOTOS_07. Here we present a short description of the main steps of this code. The calculations start with two data files: coordinates of the stations and arrival times of *P*- and *S* seismic rays from local earthquakes to these stations. Also, additional information such as starting velocity model, parameters of grid and inversion and others is defined in a separate file. It is possible to use preliminary locations and origin times provided by picking tools or/and catalogues, but this information is not strictly required. In the case of the absence of any information about sources, *LOTOS-07* starts searching for the source hypocentre either from the centre of the network or from the station with minimal arrival times. The algorithm contains the following general steps:

   (1) simultaneous optimization for the best 1-D velocity model and preliminary location of sources;
   (2) location of sources in the 3-D velocity model and
   (3) simultaneous inversion for the source parameters and velocity model using several parameterization grids.

   Steps 2 and 3 are repeated in turn, one after another, in several iterations. Now let us describe some features of these steps.

### 4.2 Algorithm for one-dimensional velocity optimization

Preliminary location of sources and searching for the best 1-D model include the following steps:

   (1) **Calculation of a traveltime table in a current 1-D model.** The traveltimes for all possible combinations of source depths epicentral distances are computed in a 1-D model using analytical formulae (Nolet 1981).
   (2) **Source location in the 1-D model.** The calculations of traveltimes in this step are based on bilinear interpolation of values from the reference table and therefore very fast. For finding source location, we use the grid search method that provides very stable solution (e.g. Koulakov *et al.* 2007).
   (3) **Matrix calculation and inversion.** We compute the matrix that reflects the effect of velocity variation in each depth level upon traveltime of each ray. Matrix inversion is performed for the 1-D *P* - and *S*-velocity anomalies and for correction of source parameters (dx, dy, dz and dt). Damping is controlled by a special smoothing block. The inversion is performed using the LSQR method (Paige & Saunders 1982; Van der Sluis & Van der Vorst 1987).

   The 1-D velocity model is then updated according to the velocity anomalies obtained in step 3 and is used as a reference model for the next iteration that contains steps 1–3. Optimum values for free parameters (number of iterations, smoothing coefficients and weights for the source parameters) are evaluated on the basis of synthetic modelling.

   Results of the 1-D model optimization based on the real data set are shown in Fig. 3. Here we use different starting models to investigate the stability of the optimization. For example, in the cases
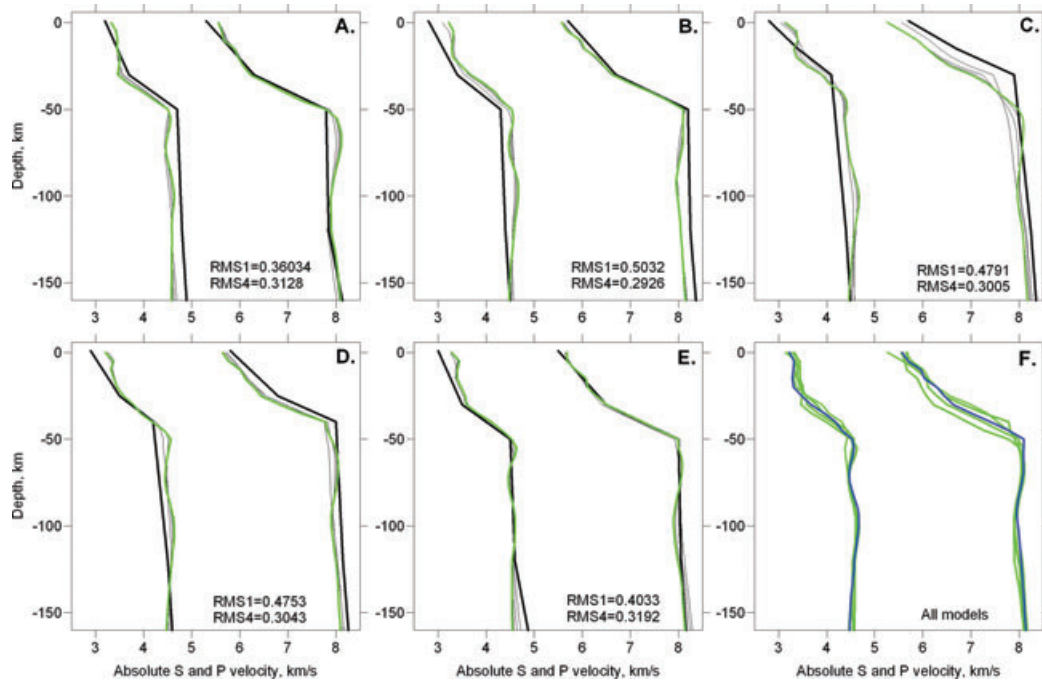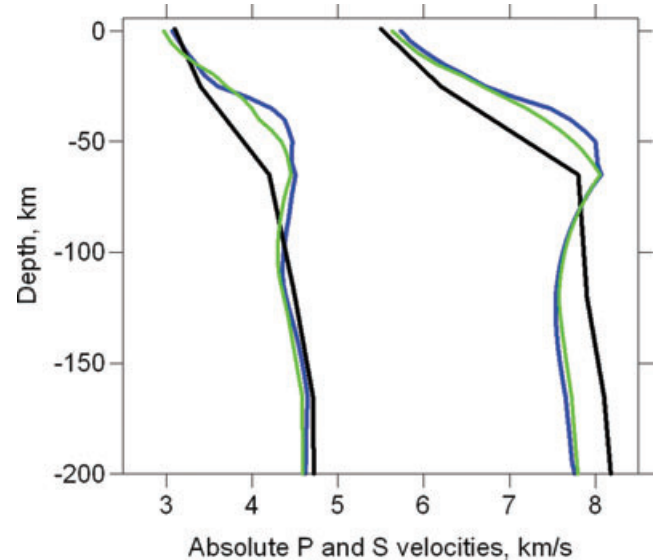
Figure 3. Optimization results for 1-D *P*- and *S*-velocity models based on real data with the use of different starting models (plots a–e) after performing four iterations. In plot f, all the results from a–e are combined in one plot. Bold black line is a starting model; thin black lines are results of inversion after 1–3 iterations; bold green line is a final optimization result after four iterations. In plot f, the best model corresponds to case b, which provides the best data fit indicated by blue. The values of rms after tracing in the starting and final models are shown with numbers.

plotted in Figs 3(a) and (b), the values of V_P/V_S in the starting 1-D velocity distributions are significantly different. Nevertheless, the optimization results in these cases are similar. In the case of shallower depth of the uppermost low-velocity layer (crust) in the starting model (Fig. 3c), the optimized model tends to deepen this layer. Combination of all the resulting velocity distributions in Fig. 3(e) shows that the most coherent results are obtained for depths below 50 km. This seems paradoxical because, as will be shown later, the vertical resolution below 50 km is rather poor. In the depth interval of 0–50 km, absolute velocities vary in the range of 10 per cent. A model with the best fit (Fig. 3b and blue line in Fig. 3f) is used as a reference distribution for further 3-D inversions. At the same time, it should be noted that the computed 3-D velocity anomalies based on various 1-D starting velocity distributions are quite similar, and all the models shown in Fig. 3 appear to be equally reliable.

To check the reliability of the optimization results and to estimate the optimum values of the free parameters, we performed an estimation of a 1-D velocity model in a synthetic test (Fig. 4). The synthetic model is represented by checkerboard anomalies superimposed with a 1-D absolute velocity distribution. The result of a 3-D reconstruction of this model is described in Section 6 and shown in Fig. 11. In this case, optimization of a 1-D model started with a model (black line in Fig. 4) that differs strongly from the 'true' synthetic 1-D basic model (blue line). The derived model (green line) appears to be fairly close to the synthetic 'true' model. The optimum free parameters that provided the best result were used for the case of real data processing.

4.3 Iterative algorithm for 3-D tomographic inversion

The tomographic inversion starts with the 1-D velocity model obtained at a preliminary step (previous section). The algorithm pro-

Figure 4. Result of optimization for 1-D *P*- and *S*-velocity distributions in the case of a synthetic checkerboard test. Blue line is the 'true' velocity distribution in the model; black line is a starting model; green line is the result of optimization.

vides two alternative options: inversion for V_P and V_S using *P* and *S* residuals (d_{tp} and d_{ts}) and inversion for V_P and V_P/V_S ratio using d_{tp} and differential residuals, $d_{ts} - d_{tp}$. Further calculations are based on performing several iterations; each iteration contains the following successive steps:

(1) **Source location in the 3-D velocity model** (in the first iteration, the 1-D velocity model is used). For the ray tracing, we used our own version of the bending algorithm that is based on

successive deformation of the ray path to achieve the minimum of the traveltime that is described in detail in Koulakov (2009a). The general idea of bending was proposed in previous studies (e.g. Um & Thurber 1987), but our practical realization is significantly different. In contrast to source location in a 1-D model that is based on the grid searching method, the location in a 3-D model uses the gradient method of searching the extreme value of goal function (Koulakov *et al.* 2007).

(2) **Parameterization.** A parameterization grid is constructed only in the first iteration. The velocity anomalies are computed in nodes distributed within the study volume according to the ray density, using the algorithm described in Koulakov *et al.* (2006). Between the nodes, the velocity distribution is approximated linearly. The minimal spacing between the nodes is significantly smaller than a size of the expected anomalies (e.g. 5 km spacing for vertical and horizontal directions in this study). To reduce an effect of node distributions on the results, we perform the inversion by using several grids with different basic orientations (e.g. $0°$, $22°$, $45°$ and $67°$) and then stack the results. This provides quasi-continuous param-

eterization in which the results are not affected by distributions of individual nodes. It is important to note that the total number of nodes (in our case ~4800 for the P and ~4600 for the S model) can be larger than the ray number (3377 P and 2462 S rays). This does not cause any obstacles for performing the inversion, because in our case, the unknown parameters associated with the parameterization nodes are not independent. They are linked with each other through a smoothing block in inversion. Examples of grids for orientations of $0°$ and $22°$ in depth interval of 5–15 and 15–25 km are shown in Fig. 5.

(3) **Matrix calculation and inversion.** Calculation of the first derivative matrix is performed using the ray paths computed after the source locations in the 3-D model. Details of matrix calculation for the cases of V_P–V_S and V_P–V_P/V_S inversions are described in detail in Koulakov *et al.* (2007). Unlike some other algorithms for V_P/V_S inversion, our code takes into account difference of the P- and S-ray paths. In addition to the V_P and V_S (or V_P/V_S) parameters, the matrix contains the elements responsible for the source corrections (dx, dy, dz and dt) and station corrections. The amplitude and smoothness
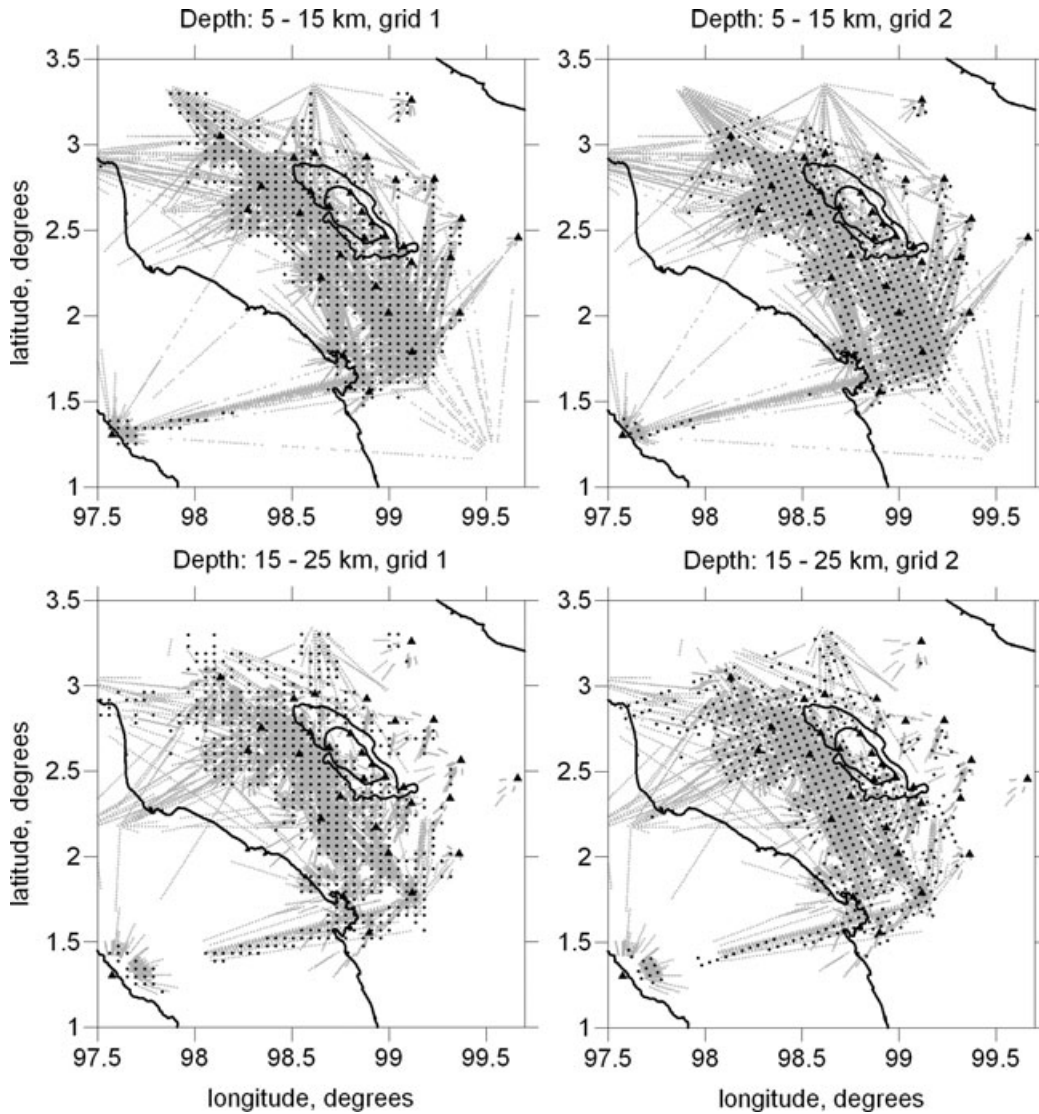
Figure 5. Example of grid construction (black dots) according to the ray distribution (grey lines). Upper and lower rows show the ray paths and nodes of the grid for the depth intervals of 5–15 km and 15–25 km, respectively. Left and right columns present two grids with orientations of $0°$ and $22°$, respectively. Black triangles depict the seismic stations.

of the solution are controlled by two additional blocks as described in Koulakov (2009a). Increasing the weights of these blocks causes either reduction in the solution amplitude or smoothness of the computed velocity fields. Inversion of the entire sparse matrix is performed using an iterative LSQR code (Paige & Saunders 1982; Van der Sluis & van der Vorst 1987).

(4) **Update of the 3-D velocity distribution.** The results obtained for several parameterization grids are averaged in one 3-D model that is used for the next iteration. This mode is defined in a regular grid that makes the ray tracing significantly faster compared to the case when the model is defined in an irregular grid.

The first iteration contains all the steps, whereas for the next iterations, only steps 1, 3 and 4 are executed. After the end of calculations, the results are visualized in arbitrary horizontal and vertical sections. The values of the resulting anomalies are shown if the distance to a nearest parameterization node is less than predefined value (e.g. 5 km in our case).

To define optimal values of free parameters for inversion (smoothing and amplitude coefficients, weights for source and station corrections, number of iterations, etc.), we perform a series of synthetic reconstruction with the conditions that represent the real situation as adequate as possible. In our opinion, the most important test for this purpose is creating a model that reproduces the realistic patterns. This modelling is discussed in Section 6.

In many studies, the damping coefficients are defined by analysing the relationship between the rms of residuals, amplitude of the solution and values of damping; these are called the trade-off curves (TOC; e.g. Eberhart-Phillips 1986). Koulakov (2009a) provides several arguments based on synthetic modelling why this approach is not valid in most cases. First, in most of the studies that use this approach, the TOC is computed for the first iteration of the inversion procedure. At the same time, the main results of these studies are obtained after performing several iterations (e.g. 3, 5, 10). The damping parameters obtained as corner points of the 'L-shaped' TOC do not necessarily coincide in the first and final iterations. It would be more reasonable to use the value that corresponds to the TOC in the final iteration, but in this case, the problem becomes too time-consuming. Second, even if we use a damping value corresponding to the TOC in the final iteration, it is not clear that this value is really optimal. Koulakov (2009a) has tested this question with several synthetic models, and he has observed that damping estimated in this way is too strong.

5 REAL DATA INVERSION

We performed a number of inversions for real and synthetic data with different parameters, data subsets and configurations of synthetic models. Here we present only a few of them that seem most important for showing the reliability of our main results, and which allow quantifying values of obtained velocity anomalies. All the real and synthetic models presented here are obtained after realization of five iteration steps. This number is taken as a compromise between solution quality and calculation speed, and it was checked with synthetic modelling.

For each solution, we compute the amplitude of the velocity variation as the difference between values of velocity deviations in prominent negative and positive anomalies. These values are computed in two small square areas ($0.12° \times 0.12°$ each) that are located inside the anomalies where the solution seems to be robust. The average values of the velocity anomalies are computed inside each square and then subtracted from each other. The amplitude

Table 1. Values of velocity contrasts at different depths observed in real data and synthetic modelling. All the values are given in per cent. For all models we consider three variants of smoothing parameters (sm1, sm2 sm3). In total, we present the results of real data inversion and reconstruction of two synthetic models with the same configuration of anomalies, but different amplitudes.

Model		Sm	5 km	15 km	25 km	35 km	45 km
REAL_sm1	P	0.7	18.3	11.6	7.3	5.3	4.1
	S	1.5	17.9	10.2	9.9	10.4	5.8
REAL_sm2	P	1.5	13.2	10.0	7.4	6.0	5.0
	S	2.5	12.8	9.1	8.3	8.2	5.1
REAL_sm3	P	0.9	17.3	12.9	8.5	6.1	4.8
	S	0.9	21.2	16.3	13.2	10.7	8.4
Synthetic model 1							
Initial anomalies	P		18	13	9	9	8
	S		18	13	9	9	8
SYN1_sm1	P	0.7	21.3	15.0	8.9	5.1	4.6
	S	1.5	16.6	10.5	8.6	5.7	4.3
SYN1_sm2	P	1.5	15.7	12.7	9.2	6.6	5.3
	S	2.5	12.9	8.6	6.8	4.5	3.0
SYN1_sm3	P	0.9	22.0	14.4	9.1	6.5	4.6
	S	0.9	17.9	11.8	7.4	3.8	5.3
Synthetic model 2							
Initial anomalies	P		16	10	7	7	6
	S		18	14	10	10	8
SYN2_sm1	P	0.7	18.7	12.5	7.3	3.9	3.7
	S	1.5	15.0	13.6	8.0	4.4	4.2
SYN2_sm2	P	1.5	12.9	9.9	7.1	4.9	4.0
	S	2.5	11.2	10.1	6.3	3.8	3.0
SYN2_sm3	P	0.9	19.4	12.8	8.1	5.6	3.8
	S	0.9	20.5	11.8	11.6	11.5	7.4

values computed in this way for some real and synthetic models are presented in Table 1. The rms values for *P* and *S* residuals and variance reduction for the real and synthetic models are presented in Table 2. The values correspond to the moment just after location at a corresponding step. This means that in the first iteration, these values reflect the rms of residuals after the location of sources in a 1-D velocity model.

We performed the inversion using two different schemes: for V_P–V_S and for V_P –V_P/V_S. In both the cases, the starting 1-D model was the same that was obtained after optimization at a preliminary location step described in Section 4.2. In this model, starting 1-D distribution of the V_P/V_S ratio varies with depth. Independent inversion according to these two schemes provides additional constraints to the damping. When performing both the schemes, the smoothing values were adjusted to achieve similar amplitudes and shapes of *P*, *S* and V_P/V_S distributions in these two cases.

The main results of the inversion for V_P, V_S and V_P/V_S ratio are presented in horizontal and vertical sections (Figs 6–9). It is clearly seen that the *P*- and *S*-velocity anomalies correlate well with each other. At the same time, we observe significant V_P/V_S variations in all depth levels that seem to be more localized in space than V_P and V_S anomalies.

At a depth of 5 km, the strongest V_P and V_S anomalies are observed beneath Toba. It may reflect volcanic deposits related to the supervolcano eruption 74 000 yr ago and sediments accumulated in the caldera. The background V_P/V_S value in the uppermost section is rather low (1.65). However, few local high V_P/V_S patterns are

Table 2. Values of data fits for the real and synthetic data results, the same as presented in Table 1. The values of rms of residuals just after location in the 1-D model (rms, 1 it) and after tracing in a final model (rms, 5 it), as well as variance reduction (var. red., per cent) are presented.

		Sm	rms, 1 it	rms, 5 it	var. red., per cent
REAL_sm1	P	0.7	0.2807	0.1942	30.8
	S	1.5	0.4335	0.2246	48.2
REAL_sm2	P	1.5	0.2807	0.2116	24.6
	S	2.5	0.4335	0.2448	43.5
REAL_sm3	P	0.9	0.2807	0.1959	30.2
	S	0.9	0.4335	0.2124	51.0
Synthetic models					
SYN1_sm1	P	0.7	0.2928	0.1779	39.23
	S	1.5	0.3540	0.1995	43.63
SYN1_sm2	P	1.5	0.2928	0.1965	32.88
	S	2.5	0.3540	0.2202	37.78
SYN1_sm3	P	0.9	0.2928	0.1738	40.28
	S	0.9	0.3540	0.1770	44.74
SYN2_sm1	P	0.7	0.2684	0.1748	34.87
	S	1.5	0.3659	0.1973	46.07
SYN2_sm2	P	1.5	0.2684	0.1896	29.35
	S	2.5	0.3659	0.2198	39.92
SYN2_sm3	P	0.9	0.2684	0.1840	33.75
	S	0.9	0.3659	0.2031	44.85

observed along the Sumatra fault and beneath the recent active volcanoes. These patterns are probably the signatures of present magma chambers beneath the volcanoes filled with partially molten material. The size of these patterns is 7–15 km, and it is on the lower resolution limit of the tomographic method. The real size of these chambers may be smaller than we observe in the V_P/V_S plot.

For the deeper sections, the most prominent feature is a strong low P- and S-velocity anomaly and a high V_P/V_S pattern that is observed beneath the Toba caldera and adjacent segments of the volcanic arc. At 15–35 km depth, we observe two approximately equal anomalies beneath Toba and Helatoba volcanoes; they probably reflect the distribution of magmatic chambers and diapirs. At greater depths, this anomaly becomes weaker but remains at the same position. In vertical sections (Fig. 9), this anomaly seems to be oriented vertically. This anomaly links the cluster of earthquakes at 120–140 km depth with Toba and Helatoba volcanoes, which may support a model of upward migration of fluids from the slab that feeds the volcanoes in the arc. However, it should be taken into account that the vertical resolution for depths below 40 km is fairly low, as will be demonstrated by synthetic tests in Section 6.

The resulting low-velocity pattern in the crust can also be associated with the location of the Great Sumatran fault indicated in Figs 6–8 with dotted lines. There is clear correlation in shape of the V_P, V_S anomalies and the high V_P/V_S ratio with the fault zone. The faulting zone represents a weakened part of the crust where penetration of deep fluids and melts is most probable. In this case, the Sumatra fault may have influenced the rise of magma through the crust that could facilitate the origin of volcanoes in the fault zone.

5.1 Distribution of local seismicity

An important result of our simultaneous inversion is the imaging of the distribution of local seismicity in the Toba region, which can be observed in horizontal and vertical sections (Figs 6, 7 and 9). It
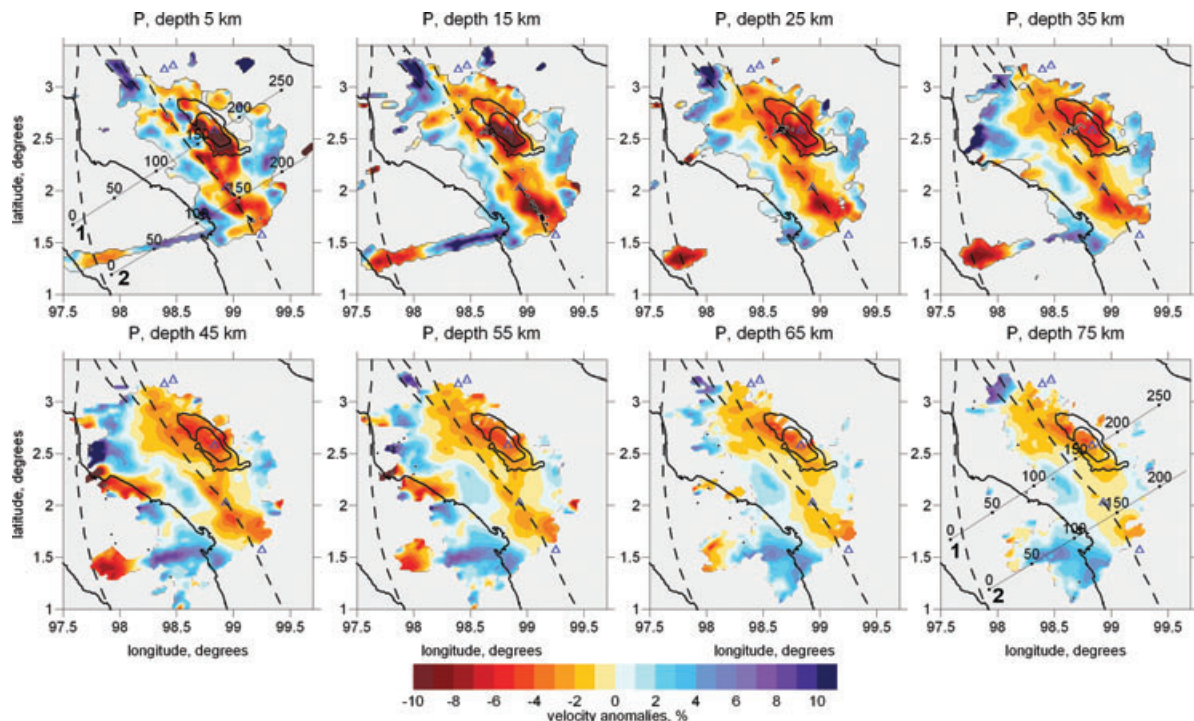

Figure 6. *P*-velocity anomalies in horizontal sections obtained as a result of real data inversion. The anomalies are given in per cent with respect to the 1-D model obtained as a result of optimization. Black contour lines within the 'red' area indicate the levels of 10 and 20 per cent anomalies. Black dots indicate the final locations of sources in a corresponding depth interval. Blue triangles show positions of active volcanoes. Dotted lines indicate the locations of the Mentawai and Sumatra fault zones. Positions of the profiles for two cross sections in Fig. 9 are given in the plots for 5 and 75 km depth.

© 2009 The Authors, *GJI*
Journal compilation © 2009 RAS
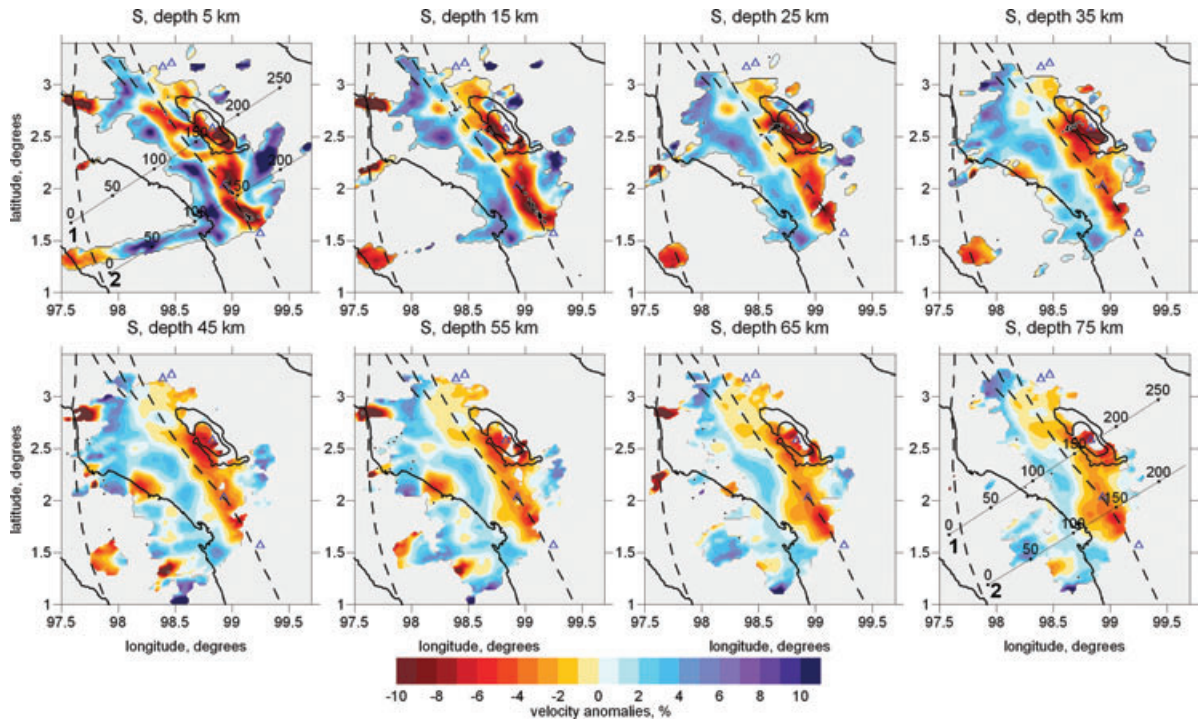
Figure 7. Same as Fig. 6, but for *S*-velocity anomalies.
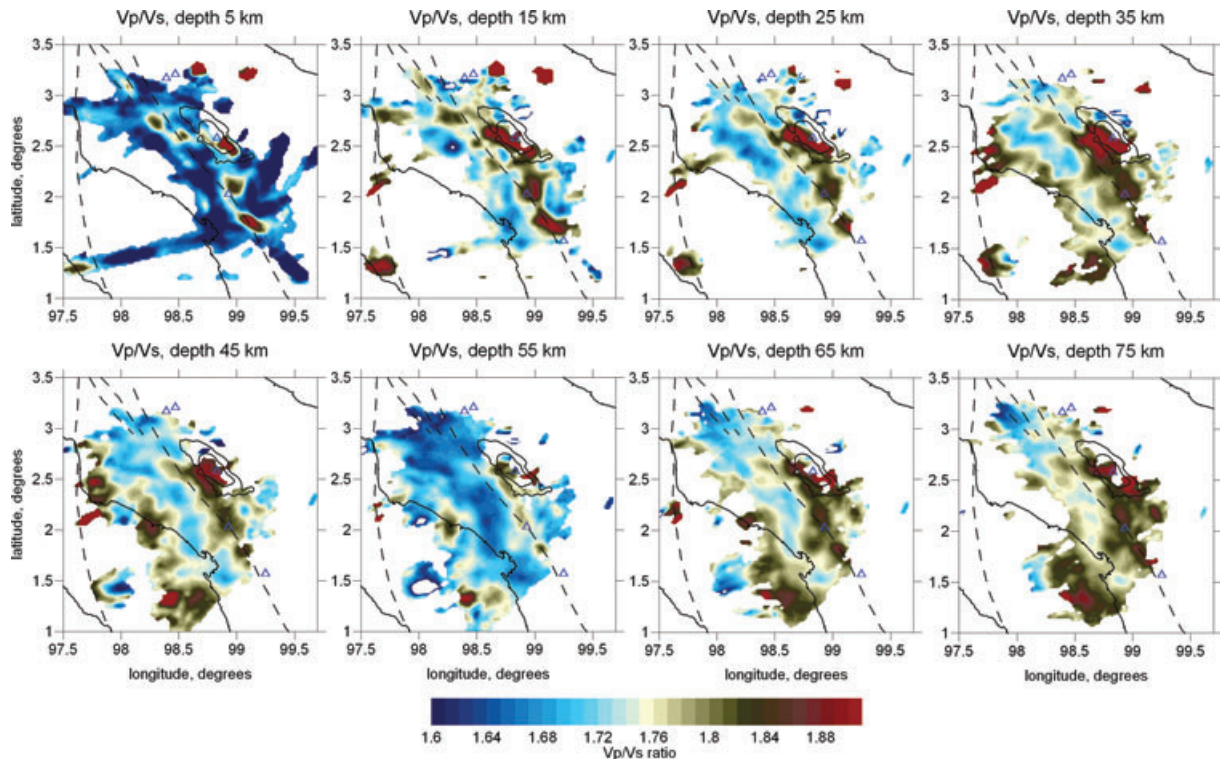
Figure 8. Same as Figure 6, but for the *V$_P$/V$_S$* ratio.

is possible to single out several groups of seismic events that may have different origin. One cluster represents the relatively shallow events (down to 20 km depth) located in the SE part of the study area and aligned along the Sumatra Fault System, between Helatoba and Lubikraya volcanoes. Another cluster is concentrated quite densely beneath a middle part of the SW border of the Toba caldera in the depth interval between 15 and 30 km. Lying directly beneath the volcanic complexes, these two clusters may possibly have volcanic or volcano-tectonic origin. They may represent the processes in the magma chambers (especially one beneath the Toba caldera). On the other hand, penetration of fluids and magma products beneath the volcanoes may cause activation of brittle fractures and stronger seismicity in a segment of the fault that coincides with the volcanic complexes. However, we cannot discuss these features at a more
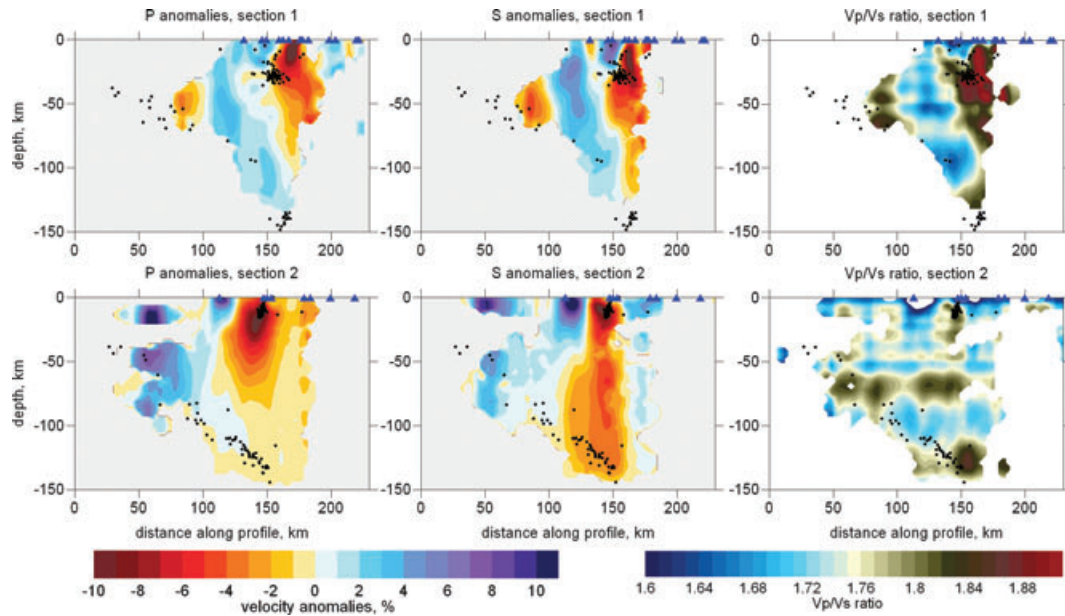
Figure 9. *P*- and *S*-velocity anomalies obtained as a result of real data inversion presented in two vertical sections. Positions of sections are shown in Figs 6 and 7. The anomalies are given in per cent with respect to the 1-D model obtained as a result of optimization. Blue triangles indicate projections of stations located inside a 40 km range to the cross-sections. Black dots indicate the final locations of sources after executing five iterations located at distances less than 40 km from the cross-sections.

quantitative level without studying the focal mechanisms, which are not available to us.

The second seismicity group includes the intermediate depth events in the Benioff zone, which can be seen in the vertical sections (Fig. 9). In the interval from 40 to 70 km depth, we observe moderate seismicity in the coupling zone of the subducted slab and overlying crust. Between 70 and 100 km depth, the seismic activity becomes weaker. The major seismic activity is observed in a cluster around 120–140 km depth. This cluster is located just beneath the Toba caldera and neighbouring volcanoes of the study area. We do not observe any evidence for a double seismic zone, as is found in some other subduction zones, such as beneath Central Java (Koulakov *et al*. 2007) and Japan (Nakajima *et al*. 2001).

5.2 Odd/even data test

To estimate the effects of random factors, we performed a test with the inversion of independent data subsets (odd/even test). In this test, the entire data set is randomly subdivided into two similar groups (for example, with events having odd or even numbers). Then the inversion is performed using the same conditions as for the entire data set. This test is very important for allocating artefacts and checking the robustness of the obtained anomalies. In our experience, this test works quite well for data with sufficient quality, but sometimes this test provides non-coherent solutions due to high noise levels. It is obvious that these results had to be rejected. Traditional tests, such as a checkerboard test, could provide satisfactory quality of solutions for these cases but did not reveal the problem of poor data quality. Therefore, we believe that the odd/even test provides very important supplementary information, and we suggest applying this test in any tomographic study.

The results of the odd/even test for this study are presented in horizontal sections in Fig. 10. It should be noted that due to the halving the data set, the solution becomes smoother than in the main results. Indeed, in the inversion step, reducing the size of

the main matrix block by two times increases the importance of supplementary smoothing blocks that leads to smoother solutions. Comparison of the inversion results for different subsets with each other, as well as with the entire data inversion, demonstrates a good correlation for both the *P* and *S* models and demonstrates that the result is rather robust and is almost unaffected by random factors.

6 SYNTHETIC MODELLING

6.1 Conditions of synthetic modelling

The most important and difficult task in tomographic inversion is not showing the results, but demonstrating that these results have any relation to the structure in the real Earth. Synthetic modelling is a key step in any tomographic study that helps in solving this problem. In particular, the synthetic testing has the following main purposes:

(1) evaluating the spatial resolution of the obtained velocity model;
(2) determining the optimal values for the free inversion parameters (smoothing coefficients, weights for source parameters and station corrections, number of iterations, etc.);
(3) estimating the real amplitudes of the anomalies and
(4) constructing a model that provides the best agreement with the real observations.

The synthetic tests are performed in a way that reproduces the procedure of real data processing as far as possible. The traveltimes are computed for sources and receivers corresponding to the real observation system using 3-D ray tracing by the bending algorithm. The traveltimes are then perturbed with noise. Our algorithm provides two options for defining the noise. According to the first option, we add the remnant residuals from the real data set after obtaining the final results. This scheme seems to us most adequate, since it defines the noise with the same magnitude and distribution
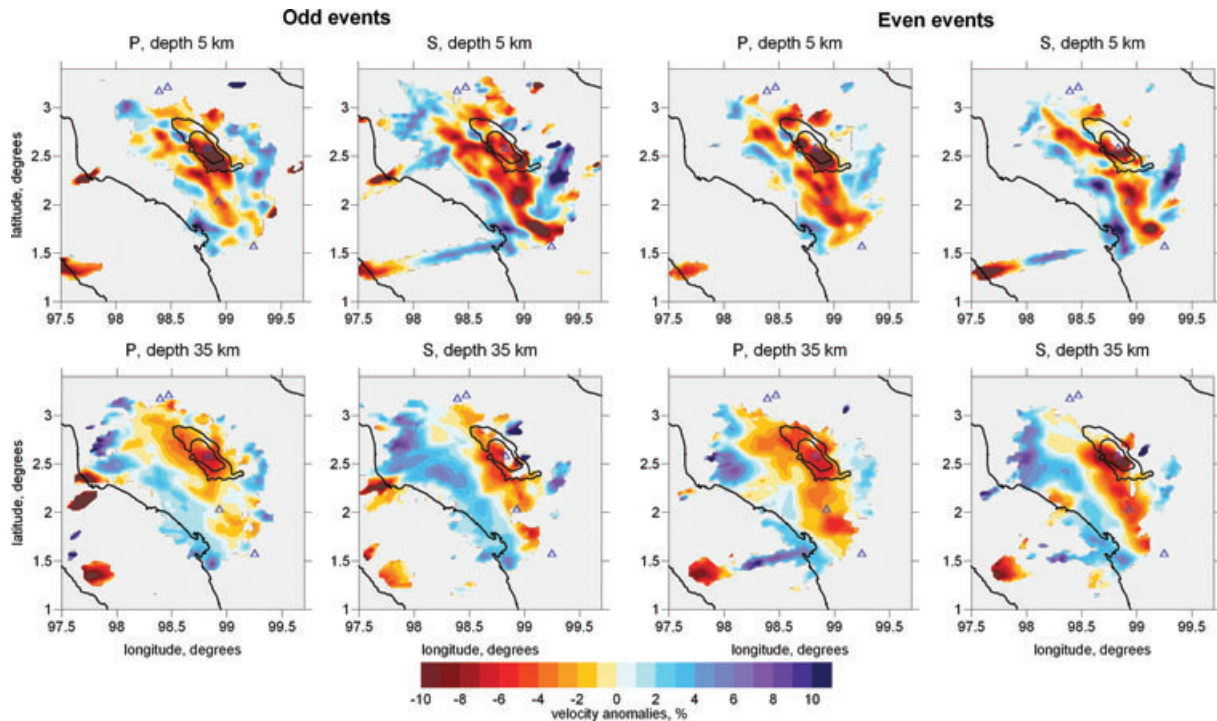
Figure 10. Inversion results for *P* and *S* models based on two independent data subsets (odd/even test) presented in two horizontal sections. Two left and two right columns represent the results for subsets with odd and even events, respectively. Comparison of these results shows the robustness of the results towards random factors (picking error, mainly).

as in the real data. This is used in most of the tests presented below. The second option produces the noise with a predefined rms for *P* and *S* data and a histogram shape by using a special generator of random numbers. We use a histogram derived from analysis of residuals from several data sets. We compared the results of reconstructions using the data with noise produced by both options. For the second option, we defined the rms of 0.15 s for *P* and 0.25 s for *S* data. The results of these two cases were very similar.

After computing the synthetic traveltimes, we 'forget' information about the coordinates and origin times of sources and anything about the velocity distribution. Thus, we simulate the situation as in processing the real data, when we had only arrival times and co-ordinates of stations. The reconstruction of the synthetic model is performed in the same way as in the real data processing, including the optimization of the 1-D model and the absolute source location. We use the same values of free parameters as in the real case. If, after performing the test, we realized that these parameters are not optimal, we find improved values and then execute the real data inversion again, using the new values. Thus, in our work, synthetic and real data inversions are always performed in a reciprocal link.

6.2 Horizontal checkerboard test

Fig. 11 presents the results of a traditional checkerboard test for investigating the horizontal resolution for V_P, V_P anomalies and the V_P/V_S ratio. In this test, the synthetic model is represented by unlimited vertical columns with a lateral size of 30 km. The amplitudes of *P* and *S* anomalies are ±4 and ±7 per cent, respectively, which produce the anomalies of V_P/V_S distribution, which are presented in the lower row of Fig. 11. Here we present the result after inversion according to the V_P–V_P/V_S scheme. In general, 30 km patterns are robustly reconstructed in all plots. Poorer quality of *P*-velocity

reconstruction is due to lower amplitude of synthetic anomalies and lower signal/noise ratio. We observe moderate diagonal smearing in the NW–SE direction, which is probably related to dominant ray orientations. This effect deforms slightly the anomalies, but the general result of this test seems to be satisfactory.

6.3 Vertical checkerboard test

The next test presented in Fig. 12 is specially aimed at checking the vertical resolution and the horizontal resolution in the NW–SE direction. Here we present only *P* anomalies. The reconstruction results for the *S* anomalies appear to be similar (Fig. S1 of the Supporting Information). The model consists of two layers of alternated blocks with positive and negative anomalies of ±7 per cent amplitude and a sign change interface at 30 km depth. In map view, they alternate along the NW–SE diagonal where the previous test showed poorer resolution. The reconstruction results show fairly good horizontal resolution for the first layer of anomalies (0–30 km) and poorer resolution for the deeper layer (30–60 km). In the vertical section, the interface between the layers is reconstructed correctly only in the central part of the profile (110–180 km along the profile). The structure below 60 km depth is strongly smeared and cannot be considered as reliable. Therefore, the results for these depths should be considered carefully.

6.4 Synthetic model with realistic patterns

One of the most important tasks of synthetic modelling consists in the creation of a model for which performing a forward and inverse problem would produce the same results as observed in a real case. Tomography is like a photograph that is taken by a camera with blurred and defected lenses that deforms the shape and colours
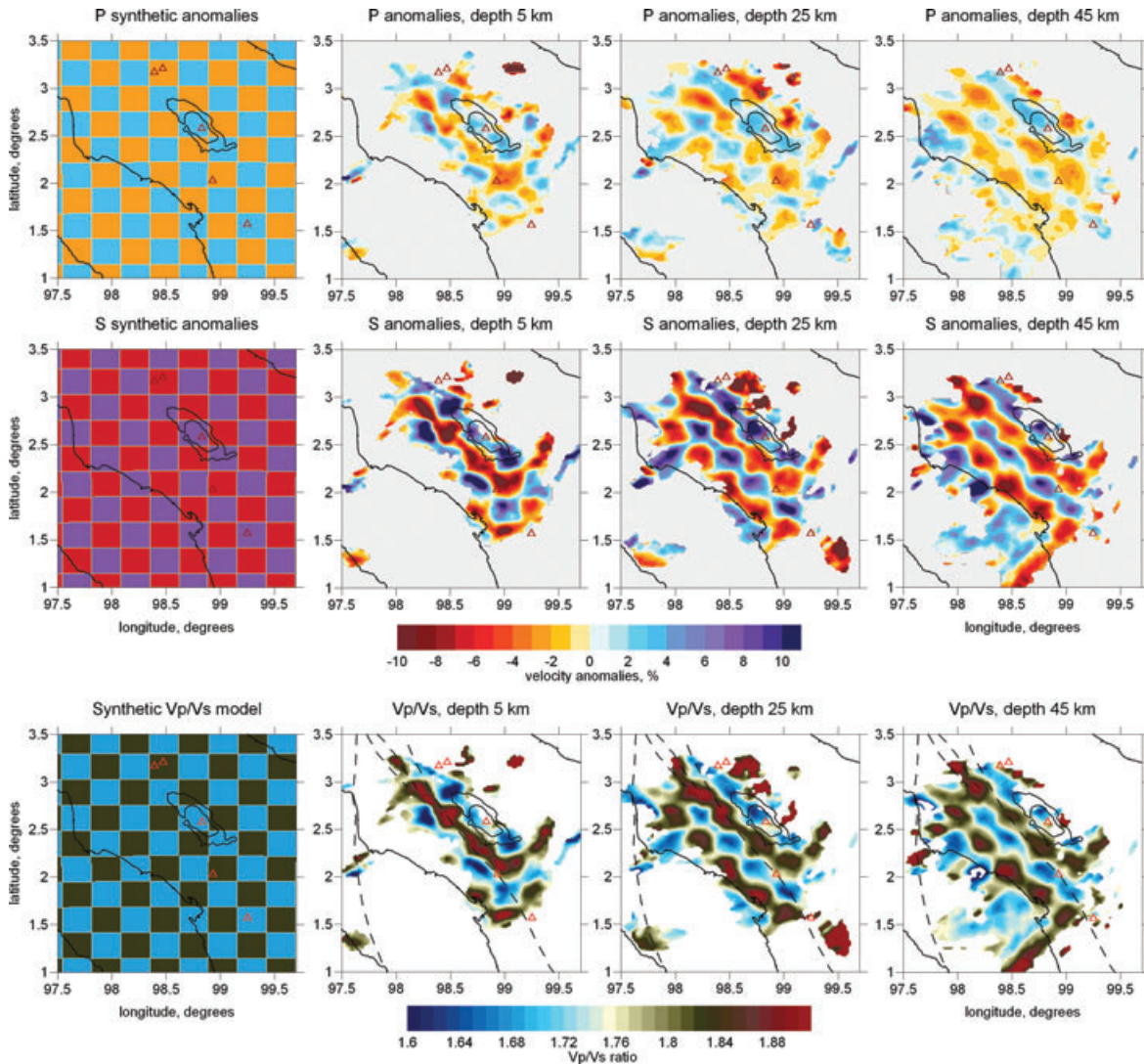
Figure 11. Reconstruction of the checkerboard synthetic model by performing the inversion for V_p and V_p/V_s. Synthetic anomalies are represented by vertically non-limited alternated columns with a lateral size of 30 km and amplitudes of ± 4 and 7 per cent for *P* and *S* anomalies, respectively. The modelling conditions are described in the text. The results for *P*, *S* anomalies and V_P/V_S ratio are presented in three horizontal sections.

of objects. The results of tomographic inversion are often quite different of the true real distribution. Just reporting the parameters of the retrieved model (e.g. shapes and amplitudes of anomalies) observed after real data inversion, as done in many studies, seems not correct to us.

Our approach is an attempt to assess more realistic distribution than images in tomograms based on synthetic modelling. To describe more clearly the idea of modelling, we present Fig. 13. A real unknown distribution in the Earth is indicated as *A1*. Real rays pass through this velocity structure and produce the real data (arrival times). Then these data are processed with the inversion algorithm that produces model *A2*, which is probably not the same as *A1*. Then we create a synthetic model *B1* (which can be similar to *A2*) as a first trial of modelling. For this model, we compute synthetic traveltimes based on tracing algorithm, which is presumed adequately representing the process of real ray propagation. Then these data are inverted using the same tomographic scheme with absolutely identical parameters as in the real case and obtain model *B2*. Now we can see how our 'photo camera' deforms the model (in the presented scheme *B2* shows smoother solution with lower amplitude than the

'true' *B1* model). Model *B2* is not similar to real inversion result *A2* (red dotted line), and it means that model *B1* is not similar to real distribution *A1*. Now we can play with the shape and amplitudes of the synthetic model to achieve the maximal similarity with *A2* (for example, model *C1* and reconstruction *C2*). If $A2 \sim C2$, we can assume that model *C1* is similar to real velocity distribution in the Earth, *A1*.

Correspondence of a synthetic model for the real observations can be evaluated by comparing several formal parameters. The two most important ones are the amplitude contrast and data fit provided by the retrieved model. These two parameters obtained for several real and synthetic models are presented in Tables 1 and 2. The data fit is computed as the rms of residuals after tracing of the rays in a current velocity model. We tried to define the synthetic model and the noise to achieve similar values of rms in all iterations in the real and synthetic cases. The value of rms in the first iteration (just after the source location in the 1-D model) is a first hint for estimating the amplitudes of the real anomalies. In fact, if the rms in this stage is lower than in the real case, it means that the amplitude in our synthetic model is underestimated.
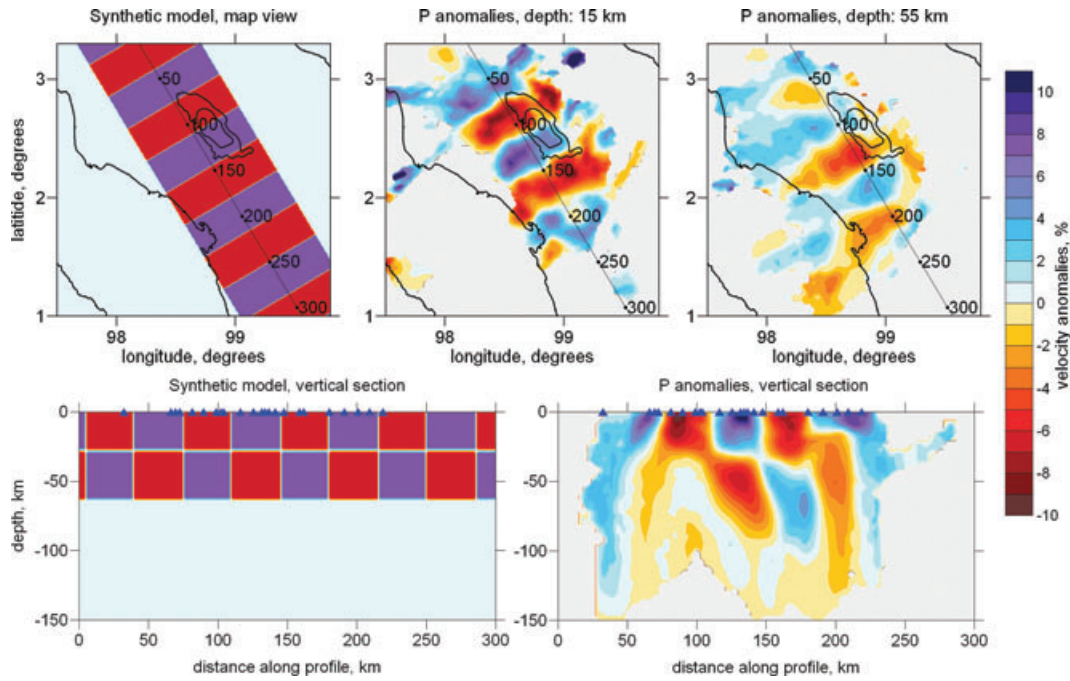
Figure 12. Test for checking the vertical and horizontal resolution in the NW–SE direction. The synthetic model is shown in the left plots both in map view and in a cross-section. Reconstruction results for *P*-velocity anomalies are shown in horizontal sections at 15 and 55 km depth (upper row, central and right plots) and in a vertical section (lower-right plot). Location of the profile is shown in all maps. Blue triangles depict the stations projected to the profile. The images for the resulting *S* anomalies are shown in Fig. S1 of the Supporting Information.

The amplitude contrast presented in Table 1 was computed as the difference between amplitude values in two areas: inside the Toba caldera (negative) and at the SW border of the resolved area (positive). Consideration of these parameters for the real and synthetic models obtained with the use of the same free parameters allows evaluating of the true amplitudes of anomalies in the real Earth. A similar approach was used (Koulakov *et al*. 2007) to show evidence of extremely strong anomalies beneath Central Java, which reached 30 and 35 per cent for the *P* and *S* models, respectively.

Let us assume that an unknown value of a velocity anomaly in the Earth, A^{real}, is reconstructed after the inversion as an anomaly with the amplitude of B^{real}. At the same time, a synthetic anomaly with known amplitude, A^{syn}, after performing the same inversion procedure, is reproduced as an anomaly with amplitude B^{syn}. According to the scheme illustrated in Fig. 13, if $B^{syn} \sim B^{real}$ in all depths, we can conclude that the true model, A^{real}, is similar to the synthetic model, A^{syn}, created by ourselves. It is important that if we found such a synthetic model, the similarity of $B^{syn} \sim B^{real}$ should take place for any values of free parameters (e.g. smoothing).

We have tested this scheme with the Toba data. In Table 1, we present the amplitudes of anomalies at different depths for the real data and two synthetic models. In each case, we performed the inversion with three different sets of smoothing parameters. Analysis of these values allows us to evaluate the amplitudes of unknown real anomalies. For example, model 'REAL_sm2' shows 10 per cent of *P* anomaly at 15 km depth. The synthetic model 'SYN1_sm2' with 13 per cent of synthetic anomaly shows 12.7 per cent of the resulting anomaly, which is slightly higher than in 'REAL_sm2'. A much more similar value (9.9 per cent) was obtained for model 'SYN2_sm2' with 10 per cent of synthetic anomaly. The similar relationships are obtained for model 'SYN2_sm2' at all depths and with all smoothing coefficients. It means that model 'SYN2_sm2'

is more adequate than 'SYN1_sm2'. The amplitudes of anomalies corresponding to this model (Table 3) are probably close to the real velocity distribution in the Earth.

The synthetic model SYN2_sm1 presented in Fig. 14 is a result of work performed by the trial and error method. We explored a dozen different synthetic configurations to find one that provides the best data fit and correlation with the results of the real inversion. The anomalies in this model are defined as vertical prisms in a depth interval corresponding to the depth level used for visualization of the results (e.g. section at 15 km depth corresponds to a prism in a depth interval of 10–20 km). It can be seen that both the amplitudes and shapes of the reconstructed anomalies are generally consistent with the results of the real data inversion (Figs 6 and 7), which is natural since the correspondence was the main criteria for the construction of this synthetic model. After performing this test, we conclude that if the real Earth distribution of *P* and *S* velocities is the same as in the model presented in Fig. 14, the results of inversion would be similar to the observed ones.

We admit that an effect of non-uniqueness may take place in this modelling. Theoretically, two different synthetic models can reproduce similar velocity patterns. However, in a practice, when we generate the synthetic models for this test, we do not have much freedom in defining amplitudes and shapes anomalies. For example, we would never be able to construct a synthetic model with an anomaly of more than 30 per cent amplitude, as reported by Masturyono *et al*. (2001), to achieve the similar reconstructions as observed in the real data tomograms. Based on this test, we conclude that such values are absolutely unrealistic.

One of the manifestations of the non-uniqueness problem is a model uncertainty that is related to poor vertical resolution. In Fig. 15, we present two vertical sections of the initial and reconstructed anomalies for the model SYN2_sm1 described above.
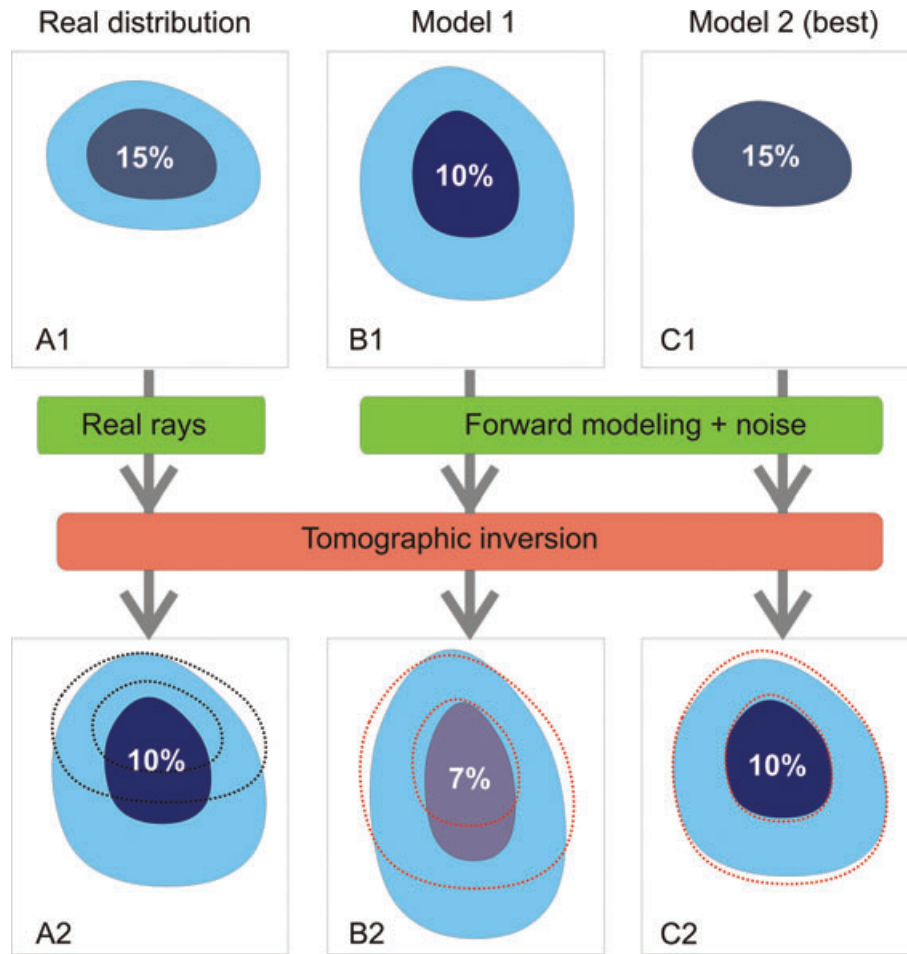
Figure 13. Scheme for the ground of finding the realistic model based of forward and inverse modelling. *A1* is the real distribution. *A2* is result if real data processing. Dotted line in *A2* indicates true velocity distribution from *A1*. *B1* and *C1* are the first and final trials of synthetic models; *B2* and *C2* are the reconstruction results corresponding to models *B1* and *C1*, respectively. Red dotted lines in *B2* and *C2* indicate the shape of the anomaly in *A2*.

Table 3. Amplitudes of *P* and *S* anomalies at different depth in the real Earth estimated from synthetic modelling.

| Depth (km)	5	15	25	35	45
dV_P (per cent)	16	10	7	7	6
dV_S (per cent)	18	14	10	10	8

In general, these images correlate quite well with the results of real data inversion (Fig. 9). However, another test with a model SYN3_sm1, presented in Fig. 16, shows that we must be careful with interpreting these results. Above the depth of 40 km, this model coincides with the model SYN2_sm1 presented in Figs 14 and 15. Below 40 km, the model is homogeneous. Reconstruction results presented in vertical sections (Fig. 16) demonstrate that the crustal anomalies are strongly smeared downwards with a deep trace into the mantle. The results for SYN2_sm1 and SYN3_sm1 appear to be very similar. Thus, it is difficult to select one of them as more probable for explaining the observed patterns.

7 DISCUSSION AND INTERPRETATION

The seismic structure beneath Toba allows quite clear interpretation which is illustrated in Fig. 17. We observe a vertical low-velocity pattern that links the seismicity cluster at 120–150 km depth with the Toba caldera. We should, however, keep in mind that the results of synthetic modelling reveal poor vertical resolution, so we cannot exclude the possibility that the observed vertical anomaly is partly due to vertical smearing of shallower anomalies beneath Toba. Thus, interpretation of this anomaly should be performed carefully. On the other hand, even without considering the seismic structure, it is quite clear that the arc volcanism and seismicity cluster located just beneath the volcanoes are linked with each other in some way. The most probable explanation of this link is that the phase transitions in the slab cause active release of fluids. Ascension of these fluids leads to decreasing melting temperature above the slab and the origin of diapirs and magmatic chambers (Poli & Schmidt 1995). In this case, a prominent cluster of events at ~40 km depth beneath the western boundary of the caldera may delineate the border of an area with high content of molten material beneath Toba.

It is unexpected that beneath Toba, we observe moderate values of *P*- and *S*-velocity anomalies. At a depth of 5 km, the anomalies reach 16–18 per cent; however, they may be partly related to young sediments that filled the Toba caldera. Some anomalies can also be related to fracturing of rocks along the Great Sumatra Fault. When considering only seismic velocities, it is not easy to distinguish these factors from volcano related anomalies (e.g. magma chambers). At 15 km depth beneath Toba and Helatoba volcanoes, we observe anomalies that are slightly stronger than 10 per cent. This is likely
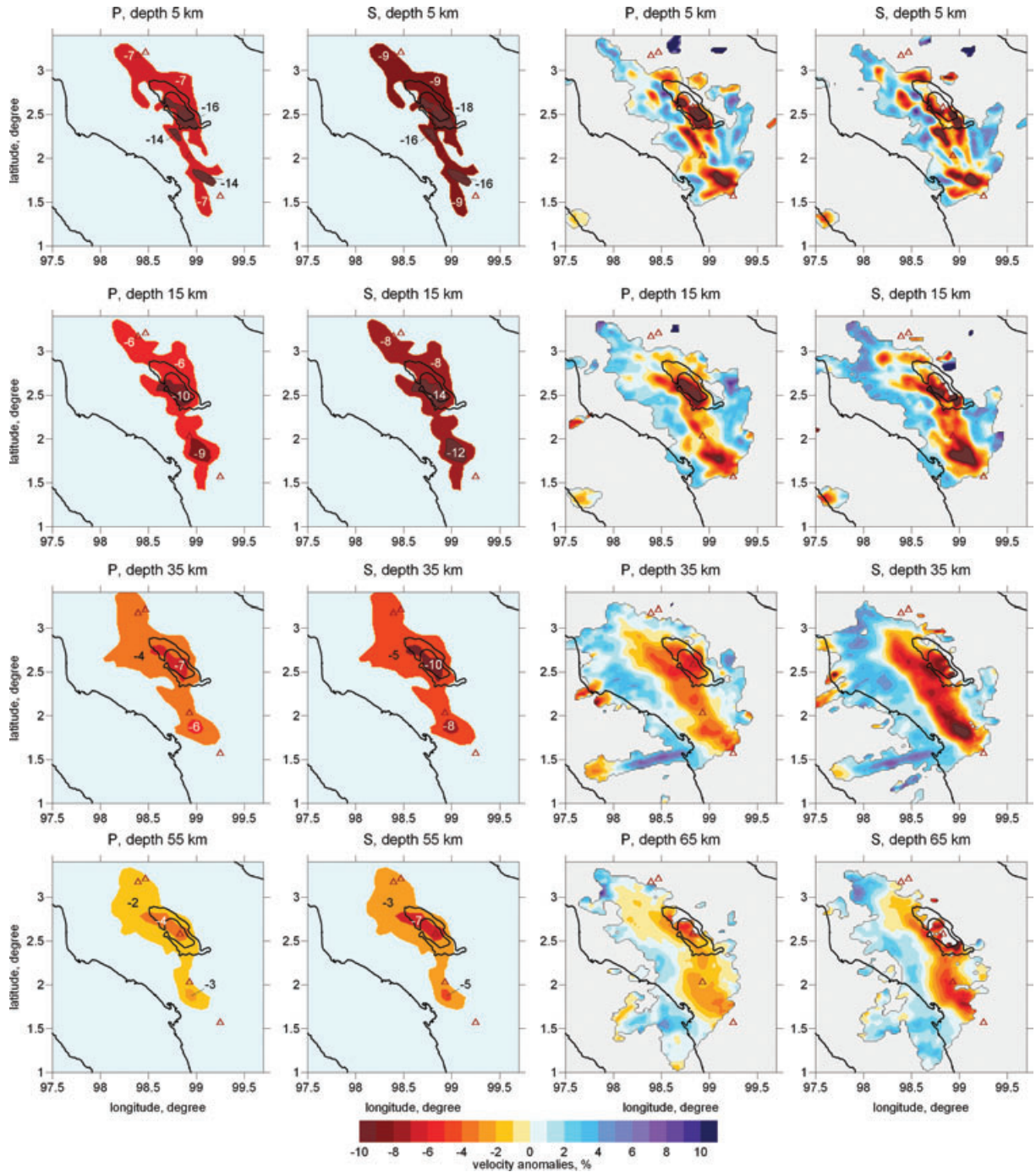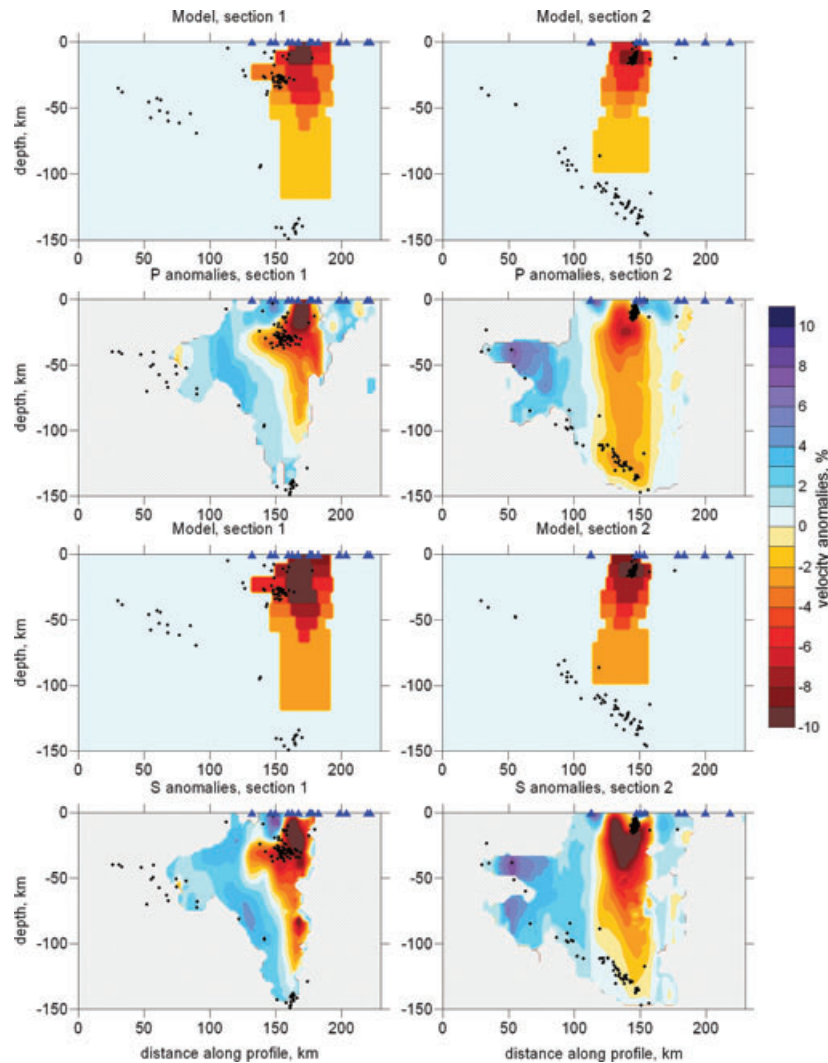
Figure 14. Synthetic model SYN2_sm1, which reproduces the images and values of formal parameters as in the real case. The configurations of the *P*- and *S*-velocity anomalies in the model are presented in horizontal sections in the left two columns. Numbers indicate the values of velocity anomalies in the model. The right two columns represent the reconstruction results after executing forward and inverse problems for the presented model. Conditions for performing the modelling and the strategy for defining the synthetic model are described in the text.

a signature of volcanic chambers and magma paths; however, this value is not higher than observed in normal volcanic areas like Central Java (Koulakov *et al*. 2007, 2009) and Central Chile (Koulakov *et al*. 2006). In the mantle, the amplitude of anomalies is less than 7 per cent, which is relatively low for volcanic areas. In this sense, no significant signature of super volcanism is observed here.

Much more interesting features are observed at the distribution of the V_P/V_S ratio, which roughly reflects the content of melts and fluids. Variations of this parameter observed at different depths are

much more significant than those of *P*- and *S*-velocity anomalies. At 5 km depth, we obtain dominantly low values of the V_P/V_S ratio with the value of about 1.62. At the same time, just beneath the presently active volcanoes, we observe few local patterns with a very high V_P/V_S ratio of about 1.87. These patterns possibly indicate the magmatic chambers that feed the volcanoes. The concept of magma chamber is widely accepted in popular scientific literature (e.g. a typical cartoon in every school book where a volcano is connected with a large spherical body of magma chamber). At the same time, in

Figure 15. Model SYN2_sm1 and reconstruction results as shown in Fig. 13, but presented in two vertical sections, same as was used for presenting the results in Fig. 9. Rows 1 and 3 represent *P*- and *S*-synthetic models, respectively; rows 2 and 4 are the reconstruction results. Black dots are the positions of sources in the starting model and in the final results. Blue triangles show the stations projected to the cross-sections.

many fairly robust and high-resolution tomographic studies for other volcanoes, the chambers are not clearly detected. For example, in Central Java, Koulakov *et al.* (2007, 2009) used data with more than 100 stations operated simultaneously and obtained high-resolution images of the crust and upper mantle beneath Merapi volcano and surrounding areas. These results did not provide any feature just beneath Merapi or other volcanoes, which could be interpreted as a magma chamber. In contrast, beneath some dormant volcanoes, they observed local high-velocity patterns that possibly indicated frozen lava bodies and channels. Beneath Toba and surrounding volcanoes at shallow depths, we observe quite small bodies with an extremely high V_P/V_S ratio that are quite clear indicators to the magma chambers.

The high V_P/V_S ratio patterns are also observed beneath the volcanic arc in the middle and lower crust (15 and 25 km depth), where their sizes and amplitudes increase (up to 1.90). For the mantle wedge, the resolution of V_P/V_S becomes rather poor, mostly due to trade-off with source depth and origin times. However, high values of V_P/V_S are robustly resolved beneath the Toba caldera in all depth sections. The fact that V_P/V_S contrasts are much more important beneath Toba caldera than *P*- and *S*-velocity anomalies

is probably an indicator that the observed anomalies are mostly related to melting processes rather than to temperature, chemical or mechanical reasons.

Another important observation is that in the Toba area, the events in the slab at 130–150 km depth that are possibly responsible for magma origin are located directly beneath the volcanic arc. In many other subduction zones, they are laterally biased. However, it is unlikely that this observation is a distinguishing feature that might be connected with the origin of the Toba supervolcano.

Are there any features of a structure beneath Toba caldera that can be called as a signature of the previous super eruption that occurred 74 000 yr BP or indicate to a potential large-scale eruption in the future? The general seismic image is quite similar to other results obtained for normal volcanoes: low *P* and *S* velocities and a high V_P/V_S ratio are observed beneath the volcanic arc in most studies of volcanic areas. In contrast, comparing the results of the tomographic inversion in Toba and Java, one might conclude that the 30 per cent anomaly between Merapi and Lawu volcanoes in Central Java is more indicative of a super volcano than the moderate low-velocity anomalies of 12–18 per cent amplitude beneath Toba. From this study, we can conclude that most of signatures of previous
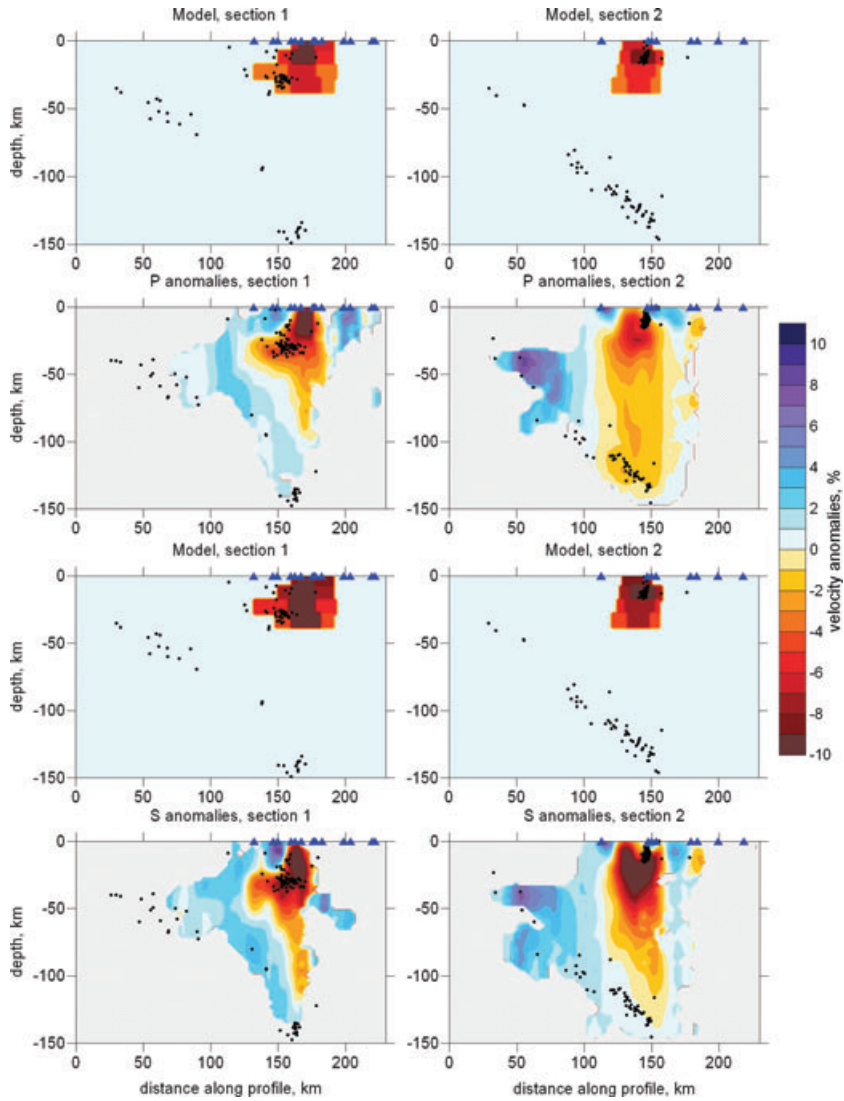
Figure 16. Reconstruction of model SYN3_sm1, which is the same as *SYN2_sm1* (Figs 14 and 15) in the depth interval of 0–40 km, but homogeneous in the deeper sections. The indications are the same as in Fig. 14. It can be seen that the crustal anomalies are strongly smeared downwards, and the reconstruction results are hardly distinguishable from those in Fig. 14.

super eruption beneath Toba have disappeared, and no new super scale features are observed there. The issue of why several eruptions of super scale have been localized in the area of Toba during the last 2 Ma remains open. In spite of the fairly robust and clear images obtained, the tomographic inversion does not provide an answer to this question.

8 CONCLUSIONS

Using a new version of the LOTOS-07 algorithm (Koulakov 2009a), we have revised a rather old data set that was collected in the Toba area in 1995. We obtained images of *P* and *S* anomalies, as well as the V_P/V_S ratio, beneath the Toba caldera that are higher in resolution and more reliable than the previous models obtained for the same region. The results are supported by various synthetic tests that show a rather good horizontal resolution. At the same time, we show the limits in vertical resolution, which appears to be rather poor. In this study, we estimate the values of amplitudes of *P* and *S*

anomalies using synthetic reconstructions of a model with realistic anomaly shapes.

We show that the negative anomalies of *P* and *S* seismic velocities beneath Toba do not exceed 15–18 per cent, which appears to be a moderate value for volcanic areas. At the same time, the lateral contrasts of the V_P/V_S ratio are much more significant. Beneath the Toba caldera, we observe a very high value of V_P/V_S that reaches 1.9. This could be an indicator of a dominant effect of melting in the origin of seismic anomalies in the crust and uppermost mantle beneath the Toba caldera.

At a depth of 5 km, the high V_P/V_S patterns are rather small in size (7–15 km) and coincide with active volcanoes. We suppose that they represent the magma chambers beneath the volcanoes.

In spite of quite clear and robust tomographic images, we do not see any patterns that could distinguish Toba as an area of super volcanism. The obtained tomographic structure beneath the caldera area is not considerably different of that observed in other 'normal' volcanic areas.

It is clear that multidisciplinary investigations of the Toba area should be continued. This area merits much more intensive
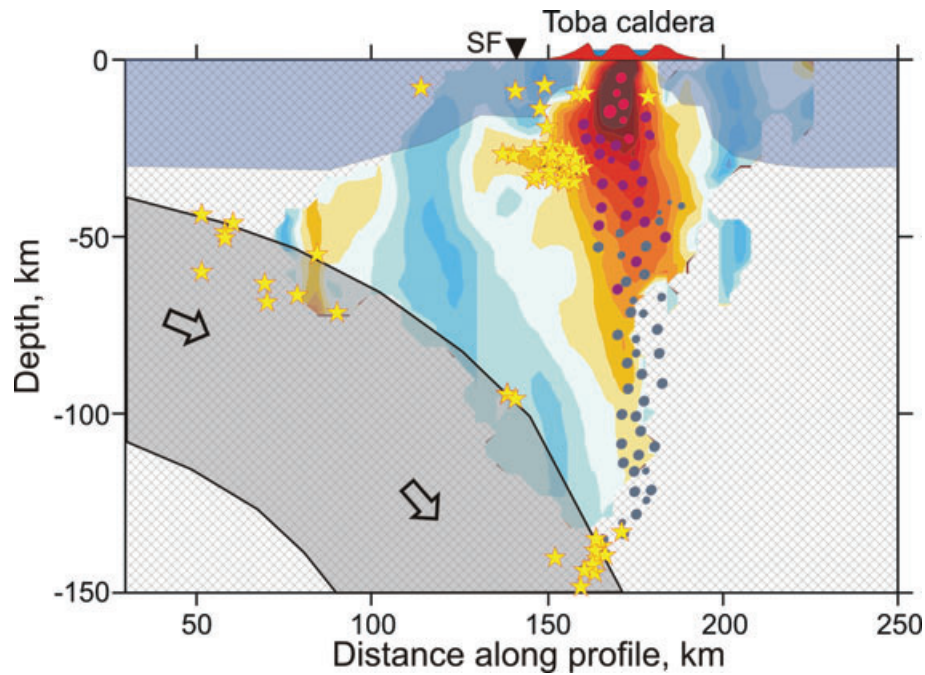
Figure 17. Interpretation of the results along profile 1, same as in Fig. 9. Background is the distribution of *P*-velocity anomalies. Yellow stars are the final location of events used in this study. Blue, violet and red drops mark schematically the path of ascending fluids and partially molten material. Black triangle and SF indicates the Sumatra fault.

geophysical investigations, at least at a level of those performed in two other super volcano areas: Yellowstone and Taupo. We believe that additional information about the local seismicity, as well as increasing the spatial resolution of the tomographic models, will reveal new features about the deeper structure beneath Toba. It will help to answer many open issues about the phenomena of the Toba supervolcano, which had a global effect in the very recent history of the Earth.

REFERENCES

Aldiss, D.T. & Ghazali, S.A., 1984. The regional geology and evolution of the Toba volcano-tectonic depression, Indonesia, *J. Geol. Soc. Lond.,* **141,** 487–500.

Bachmann,, O. & Bergantz,, G., 2008. The magma reservoirs that feed supereruptions, *Elements,* **4**(1), 17–21.

Bannister, S., Bryan, C.J. & Bibby, H.M., 2004. Shear wave velocity variation across the Taupo Volcanic Zone, New Zealand, from receiver function inversion, *Geophys. J. Int.,* **159,** 291–310.

Chesner, C.A. & Rose, W.I., 1991. Stratigraphy of the Toba tuffs and the evolution of the Toba caldera complex, Sumatra, Indonesia, *Bull. Volc.,* **53**(5), 343–356, doi:10.1007/BF00280226

Chesner, C.A., Rose, W.I., Deino, A. & Drake, R., 1991. Eruptive history of Earth's largest Quaternary caldera (Toba, Indonesia) clarified. *Geology,* **19,** 200–203.

Christiansen, R.L., 2001. The Quaternary and Pliocene Yellowstone plateau volcanic field of Wyoming, Idaho, and Montana, No. 729-G, U.S. Geological Survey.

DeMets, C., Gordon, R., Argus, D. & Stein, S., 1990. Current plate motions, *Geophys. J. Int.,* **101,** 425–478.

Diehl, J.F., Onstott, T.C., Chesner, C.A. & Knight, M.D., 1987. No short reversals of Brunhes age recorded in the Toba tuffs, north Sumatra, Indonesia, *Geophys. Res. Lett.,* **14**(7), 753–756.

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

Fauzi, McCaffrey, R., Wark, D., Sunaryo & Prih Haryadi, P.Y., 1996. Lateral variation in slab orientation beneath Toba caldera, *Geophys. Res. Lett.,* **23,** 443–446.

Hadley, D. M., Stewart, G.J. & Ebel, J.E., 1976. Yellowstone: seismic evidence for a chemical mantle plume, *Science,* **193,** 1237–1329.

Harrison, A.J. & White, R.S., 2004. Crustal structure of the Taupo volcanic zone, New Zealand: stretching and igneous intrusion, *Geophys. Res. Lett.,* **31,** L13615.

Humphreys, E.D., Dueker, K.G., Schutt, D.L. & Smith, R.B., 2000. Beneath Yellowstone: evaluating plume and nonplume models using teleseismic images of the upper mantle, *GSA Today,* **10,** 1–7.

Husen, S., Smith, R.B. & Waite, G.P., 2004. Evidence for gas and magmatic sources beneath the Yellowstone volcanic field from seismic tomographic imaging, *J. Volcanol.. Geotherm. Res.,* **131,** 397–410.

Koulakov, I. *et al.,* 2007. P and S velocity structure of the crust and the upper mantle beneath central Java from local tomography inversion, *J. geophys. Res.,* **112,** B08310, doi:10.1029/2006JB004712.

Koulakov, I., Sobolev, S.V. & Asch, G., 2006. *P-* and *S*-velocity images of the lithosphere-asthenosphere system in the Central Andes from local-source tomographic inversion, *Geophys. J. Int.,* **167,** 106–126.

Koulakov, I., 2009a. LOTOS code for local earthquake tomographic inversion. Benchmarks for testing tomographic algorithms, *Bull. seism. Soc. Am.,* **99**(1), 194–214, doi:10.1785/0120080013.

Koulakov, I., 2009b. Out-of-network events can be of great importance for improving results of local earthquake tomography, *Bull. seism. Soc. Am.,*, submitted, available at: www.ivan-art.com/science/PAPERS/gap.pdf.

Koulakov, I., Jakovlev, A. & Lühr, B., 2009. Anisotropic structure beneath Central Java from local earthquake tomography, *Geochem. Geophys. Geosyst.,* **10,** Q02011, doi:10.1029/2008GC002109.

Masturyono, McCaffrey, R., Wark, D.A., Roecker, S.W., Fauzi, Ibrahim, G. & Sukhyar, 2001. Distribution of magma beneath Toba caldera, North Sumatra, Indonesia, constrained by 3-dimensional P-wave velocities, seismicity, and gravity data, Geochemistry, *Geophys Geosyst.* **2.**

Miller, D.S. & Smith, R.B., 1999. P and S velocity structure of the Yellowstone volcanic field from local earthquake and controlled source tomography, *J. geophys. Res.,* **104,** 15105–15121.

Nakajima, J., Matsuzawa, T., Hasegawa, A. & Zhao, D., 2001. Three-dimensional structure of Vp, Vs and Vp/Vs beneath northeastern Japan: implication for arc magmatism and fluids, *J. geophys. Res.,* **106,** 21843–21857.

Ninkovich, D., Shackleton, N.J., Abdel-Monem, A.A., Obradovich, J.D. & Izett, G., 1978. K-Ar age of the late Pleistocene eruption of Toba, north Sumatra. *Nature,* **276,** 574–577.

Nishimura, S., Abe, E., Yokoyama, T., Wirasantosa, S. & Dharma, A., 1977. Danau Toba-the outline of Lake Toba, North Sumatra, Indonesia. *Paleolimnol. Lake Biwa Japanese Pleistocene,* **5,** 313–332.

Nolet, G., 1981. Linearized inversion of (teleseismic) data, in *The Solution of the Inverse Problem in Geophysical Interpretation,* pp. 9–37, ed. Cassinis, R. Plenum, New York, USA.

Paige, C.C. & Saunders, M.A., 1982. LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Soft.,* **8,** 43–71.

Poli, S. & Schmidt, M.W., 1995. H_2O transport and release in subduction zones: experimental constrains on basaltic and andesitic systems, *J. geophys. Res.,* **100,** 22 299–22 314.

Pulunggono, A. & Cameron, N. R., 1984. Sumatran microplates, their characteristics and their role in the evolution of the Central and South Sumatra basins, in *Proceedings of the Indonesian Petroleum Association,* Vol. 13, pp. 121–143.

Reyners, M., Eberhart-Phillips, D., Stuart, G. & Nishimura, Y., 2006. Imaging subduction from the trench to 300 km depth beneath the central North Island, New Zealand, with Vp and Vp/Vs, *Geophys. J. Int.,* **165,** 565–583, doi:10.1111/j.1365-246X.2006.02897.x.

Rogan, M., 1982. A geophysical study of the Taupo volcanic zone New Zealand, *J. geophys. Res.,* **87**(B5), 4073–4088.

Rose, W.I. & Chesner, C.A., 1987. Dispersal of ash in the great Toba eruption, 75 ka. *Geology,* **15,** 913–917.

Sherburn, S., Bannister, S. & Bibby, H., 2003. Seismic velocity structure of the central Taupo Volcanic Zone, New Zealand, from local earthquake tomography, *J. Volc. Geotherm. Res.,* **122,** 69–88.

Simkin, T. & Siebert, L., 1994. *Volcanoes of the World,* 2nd edn, Geoscience Press in association with the Smithsonian Institution Global Volcanism Program, Tucson AZ, 368 p.

Smith, R.L. & Bailey, R.A., 1968. Resurgent cauldrons, in Studies in Volcanology, (A memoir in honor of Howel Williams), *Geol. Soc. America Mem. V.,* **116,** 613–662

Soengkono, S., 1994. A magnetic model for deep plutonic bodies beneath the central Taupo Volcanic Zone, North Island, *NZ J. Volc. Geotherm. Res.,* **68,** 193–207.

Steckler, M.S., Akhter, S.H. & Seeber, L., 2008. Collision of the Ganges–Brahmaputra Delta with the Burma Arc: implications for earthquake hazard, *Earth planet Sci. Lett.,* **273,** 367–378.

Um, J. & Thurber, C., 1987. A fast algorithm for two-point seismic ray tracing, *Bull. seism. Soc. Am.,* **77,** 972–986.

Van Bemmelen, R.W., 1949. The geology of Indonesia, in *General Geology of Indonesia and Adjacent Archipelagos,* Vol. 1A. Government Printing Office, The Hague, The Netherlands, 732 pp.

Van der Sluis, A. & van der Vorst, H.A., 1987. Numerical solution of large, sparse linear algebraic systems arising from tomographic problems, in *Seismic Tomography,* pp. 49–83, ed. Nolet, G., Reidel, Dordrecht.

Waite, G.P., Smith, R.B. & Allen, R.M., 2006. VP and VS structure of the Yellowstone hot spot: evidence for an upper mantle plume, *J. geophys. Res.,* **111,** B04303, doi:10.1029/2005JB003867.

Wilson, C.J.N., Blake, S., Charlier, B.L.A. & Sutton, A.N., 2006. The 26.5 ka Oruanui eruption, Taupo volcano, New Zealand: development, characteristics and evacuation of a large rhyolite magma body, *J. Petrol.,* **47,** 35–69.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Test for checking the vertical and horizontal resolution in the NW–SE direction for the *S* velocity anomalies.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.