Bulletin of the Seismological Society of America, Vol. 99, No. 4, pp. 2556–2563, August 2009, doi: 10.1785/0120080365

Ⓔ

Short Note

Out-of-Network Events Can Be of Great Importance for Improving Results of Local Earthquake Tomography

by Ivan Koulakov

Abstract In most local earthquake tomography (LET) studies the data are selected according to the azimuthal-gap (GAP) criterion, which means rejecting all the sources outside the station network perimeter. In this article I show that in some cases this criterion is inappropriate and can be a reason for significant decimation of the relevant data, in turn leading to lower quality inversion results. This study presents several qualitative and quantitative arguments for why the GAP criterion is not adequate. The fact of great importance to out-of-network events for improving the results of tomographic inversion is supported by synthetic testing using realistic distribution of events in the area of Central Java and station locations according to the Merapi Amphibious Experiments project. I consider three models with different criteria of event selection: (1) dataset with GAP $< 180°$, (2) dataset with GAP $< 280°$, and (3) dataset with all events within a radius of 5°. The synthetic modeling reproduces the real situation when neither coordinates of sources nor starting 1D models are initially known. The reconstruction results show that the best resolution is obtained for model 3 with all data available, while the worst solution is observed in model 1. This study demonstrates that the commonly used GAP criterion that rejects the out-of-network events is injurious for LET tomography. In future experiments and when reconsidering the old datasets, I encourage the use of data from all events to improve the results of tomographic inversion, though at large distances from networks (at least, up to 400–500 km).

Online Material: Results of synthetic tests with events selected according to different criteria.

Introduction

Here, I consider a problem of data selection when performing local earthquake tomography (LET). There is a presumption widely held in the tomographic community that sources located outside the perimeter of the station network are useless for performing tomographic inversion. In most LET algorithms (see examples in Thurber *et al.*, 1995; Husen *et al.*, 2000; Paul *et al.*, 2001; Vlahovic and Powel, 2001; Chiarabba and Amato, 2003; Husen *et al.*, 2003; Barberi *et al.*, 2004; Husen *et al.*, 2004; Husen and Smith, 2004; Reyners *et al.*, 2006; Dias *et al.*, 2007; Daly *et al.*, 2008), all sources located outside the network are rejected *a priori*. To make this selection, a GAP criterion, the maximum empty azimuthal angle, is used. In these studies events having a GAP of more than 180° are rejected. Even if the GAP criterion is not mentioned directly in some studies (e.g., Eberhart–Phillips and Bannister, 2002; Kato *et al.*, 2007), the presented figures show that only events located inside the network perimeter are used in the majority of the LET studies. This criterion is grounded on an assumption that the accuracy of source locations inside the network is better than that outside. In most studies, using the GAP criterion significantly shrinks the dataset. Nevertheless, the authors of such works believe that this criterion increases the reliability of the results. I could not, however, find in the literature any appropriate argumentation that confirmed the statement that rejecting the out-of-network data according to the GAP criterion really improves the results of tomographic inversion. Taking into account the costs of a network deployment and maintaining the stations during an experiment, throwing out a considerable part of the data just because of this ungrounded criterion seems to be unacceptable. Therefore, it is a matter of great importance to confirm or disprove the efficiency of the GAP criterion based on quantitative estimates. The main purpose of this article is to explore the effect of adding/rejecting the out-of-network sources upon the results of tomographic inversion using synthetic modeling.

Qualitative Arguments against Using the GAP Criterion in LET

The statement that the accuracy of out-of-network source locations is lower, and the error increases with distance from the network, seems to be intuitively reasonable, and it is correct in most cases. However, there are some obvious exceptions. For example, an event located slightly outside the network and having 100 clear picks will probably be much better determined than another event located inside the network and having only 10 noisy picks. However, in most LET studies that use the GAP criterion, the former event is rejected, and the later is used in processing.

Second, introducing the GAP criterion in most algorithms presumes computing the angles in horizontal projection. However, it is clear that the accuracy of source location depends on the GAP value in three dimensions. When sources are located below a certain depth for surface station arrays of finite extent, the 3D GAP is always greater than 180°, that is, no down going rays are observed at the source. The deeper a source is, the greater will be its 3D GAP, and the lower the accuracy of the source location. In some current LET schemes, using the 2D GAP criterion creates the dubious situation illustrated in Figure 1. A shallow source located just a few kilometers outside the network is rejected (circles in clusters 1 and 2) while an earthquake several hundreds of kilometers deep (crosses in cluster 3) is taken into consideration. It is likely that the former events are better located than the deep events.

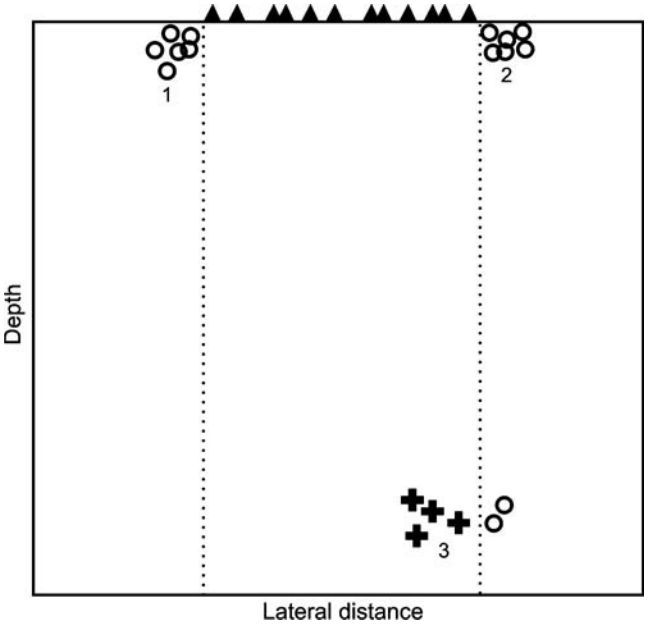

Figure 1. Sketch showing inadequacy of the GAP criterion. Crosses and circles indicate events in three clusters in a vertical profile; triangles on surface indicate station projections. It is possible that events in clusters 1 and 2 (circles) are better located than crosses in the deep cluster 3. At the same time, the GAP criterion selects only events in cluster 3 as indicated by crosses.

However, even if one accepts for argument's sake that location accuracy outside the network is low, it does not follow *a priori* that these events are useless for deriving the velocity structure. For example, in tomographic schemes based on using the teleseismic rays from remote events (see examples in Aki *et al.*, 1977; Evans and Achauer, 1993; Lippitsch *et al.*, 2003; Sandoval *et al.*, 2004), which represent an extreme case of large GAP, the source parameters are taken from international catalogs, and errors in their location are often rather large. Nevertheless, this constraint does not cause practical trouble when performing tomographic inversions. In LET algorithms where the GAPs for the events outside the network are much more reasonable, the problem of source uncertainty is less dramatic. I claim that adding sources located outside the network perimeter does not cause a problem in LET inversions and, in most cases, improves the resolution of the resulting velocity model. In this article, I present examples of synthetic modeling that disprove the assertion that the GAP criterion is appropriate in LET schemes.

Synthetic Modeling

The synthetic modeling has been performed using the LOTOS algorithm for LET inversion (e.g., Koulakov *et al.*, 2007; Koulakov, 2009; see the Data and Resources section). Set up of the synthetic model and data processing represent, as close as possible, the realistic situation. The distribution of stations (Fig. 2a) corresponds to a real experiment in Central Java performed in 2005 in the framework of the MARAMEX project (Koulakov *et al.*, 2007). Part of the events is taken from the real catalog of seismicity recorded during the experiment. Unfortunately, when the processing of the initial data was performed, the operators who picked the phases were under the pressure of the GAP stereotype, and they did not pay much attention to collecting out-of-network events. Only few of them were detected in the final stage of picking. Ideally, these data should be reprocessed, but unfortunately, I do not have either human or financial resources to repeat this routine work. In the modeling presented in this study the Merapi Amphibious Experiments (MERAMEX) dataset was supplemented with 500 events from the International Seismological Center (ISC, 2001) catalog located within a radius of 5° with respect to the center of the network (110° E and 8° S). For the MERAMEX events I used the same source-receiver (S-R) pairs as in the observed dataset. For the ISC events the S-R pairs were generated artificially. For such events the number of *P* picks varied randomly from 0 to 50. The stations for each S-R pair were selected randomly. For 20%–60% of the randomly selected S-R pairs having *P* picks, I created *S* picks. After generating the dataset, the events with less than 20 picks were removed from consideration.

In this study I consider three different datasets. In the first dataset I use only events having the GAP $< 180°$. A total of 130 events within the area M1 in Figure 2 (including 27 ISC events) with 4879 *P* and 2668 *S* picks were selected. In

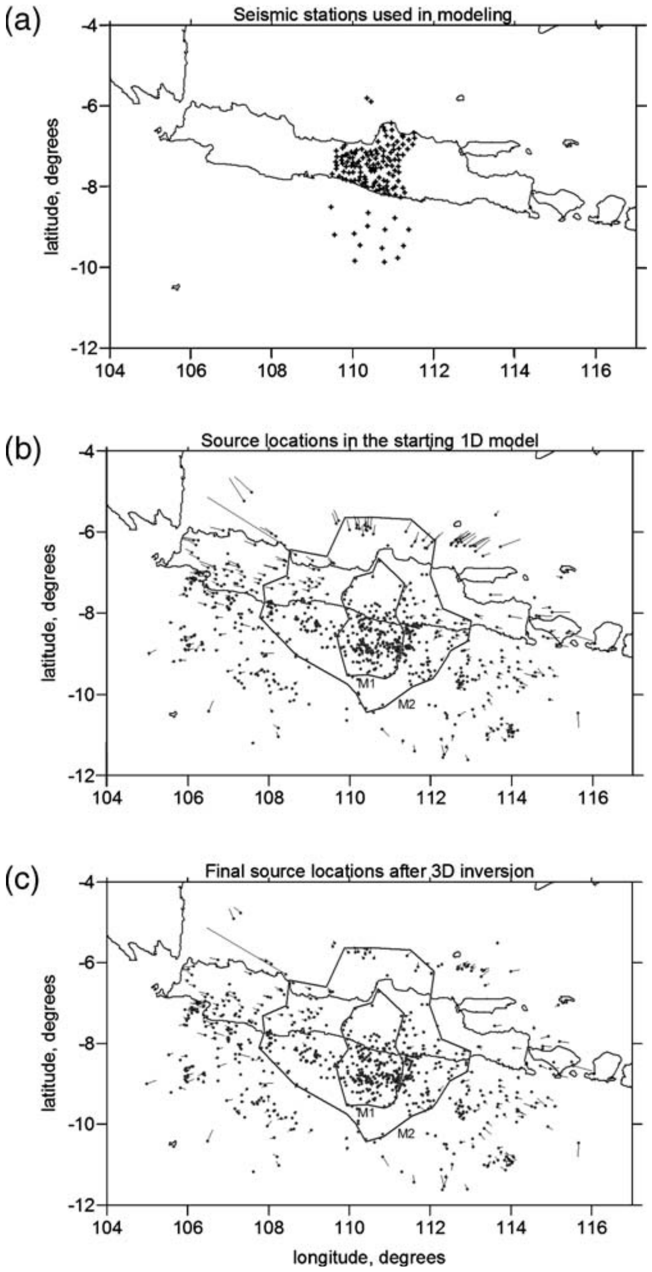

Figure 2. Stations and relocated sources used for modeling. (a) Stations used for modeling indicated by crosses. (b, c) Locations of events in the starting 1D velocity model (plot B) and after final iteration of 3D tomographic inversion (plot C) are depicted with gray dots. Bars show errors with respect to true locations. Areas marked with M1 and M2 indicate the areas of event selection for the cases of GAP < 180° and GAP < 280°, respectively.

the second dataset, the allowed GAP was increased up to 280° (area M2 in Fig. 2). In this case the number of events was 343 (including 109 ISC events) with a corresponding 13,060 P picks and 6649 S picks. In the third dataset I used all the events within a radius of 5° with respect to the center of the network (a total 838 events including 424 ISC events, 31,954 P picks, and 15,797 S picks). The distribution of events for the latter case is presented in Figure 2.

The average summary number of P and S picks per event is similar in all cases (58, 57, and 57 for models 1, 2, and 3, respectively). Adequacy of the modeling setup can be grounded by simple estimates. In the area of about 200×300 km corresponding to model 1 (GAP > 180°) I have 130 events. In model 3, the area for event selection is 500 km in radius, and its space is about 20 times larger than in the case of model 1. Statistically, in model 3 one can expect 20 times more events of the same class than in model 1 (~2600 events). In fact, for the specific case of the Central Java experiment, the number of out-of-network events can be even larger because the network is deployed in the area of a relative seismic gap that is clearly observed in the ISC data. In the modeling I use ~700 out-of-network events (total 838 events minus 130 events located inside the network). This is almost four times less data amount than the statistically possible value and represents the selection of events with sufficiently large magnitudes. Such out-of-network events are recorded by a similar average number of stations as the weaker local events located inside the network area.

A synthetic model is defined as a 3D checkerboard pattern ($30 \times 30 \times 25$ km in size along the X, Y, and Z directions, respectively). Anomalies of ±2% amplitude are given with respect to the true 1D reference velocity distribution, V_{true} (bold black lines in plots d in Figs. 3–5). For sources located far outside the network, a large part of the ray paths travels outside the resolved area, which roughly coincides with the network location. If such rays pass through outside velocity anomalies, they accumulate additional residuals and may bias the computed structure in the resolved area. In order to investigate this effect upon the velocity structure beneath the network, I defined the periodical checkerboard structure far outside the resolved area (in a square of 1600×1600 km size). In the vertical direction the checkerboard is defined down to 150 km depth. In Figures 3–5 the configuration of the checkerboard in vertical and horizontal sections is indicated with grid lines.

Synthetic travel times are computed by 3D ray tracing in the checkerboard model using our own version of the bending algorithm (Koulakov, 2009). I add the random noise with realistic rms (0.1 sec for P and 0.15 sec for S data) and distribution histogram. Furthermore, the origin times of each event are perturbed with a random bias. After computing these synthetic data, I forget the coordinates and origin times of sources and everything about the velocity model. I then have only the coordinates of stations and randomly perturbed arrival times of P and S rays. I use a starting 1D model, V_{start} (dotted line in plot G in Figs. 3–5), which is presumably different from true values of V_{true}. Initial source coordinates are the same for all sources, the center of the network at the depth of 10 km. Sources located in the starting model and mislocation errors with respect to true values are shown in Figure 2B. It can be seen that sources in the northern clusters that are located at ~600 km depth are strongly biased, that is, probably related to wrong V_P/V_S ratio in the starting model (1.74) with respect to the true one (1.7).

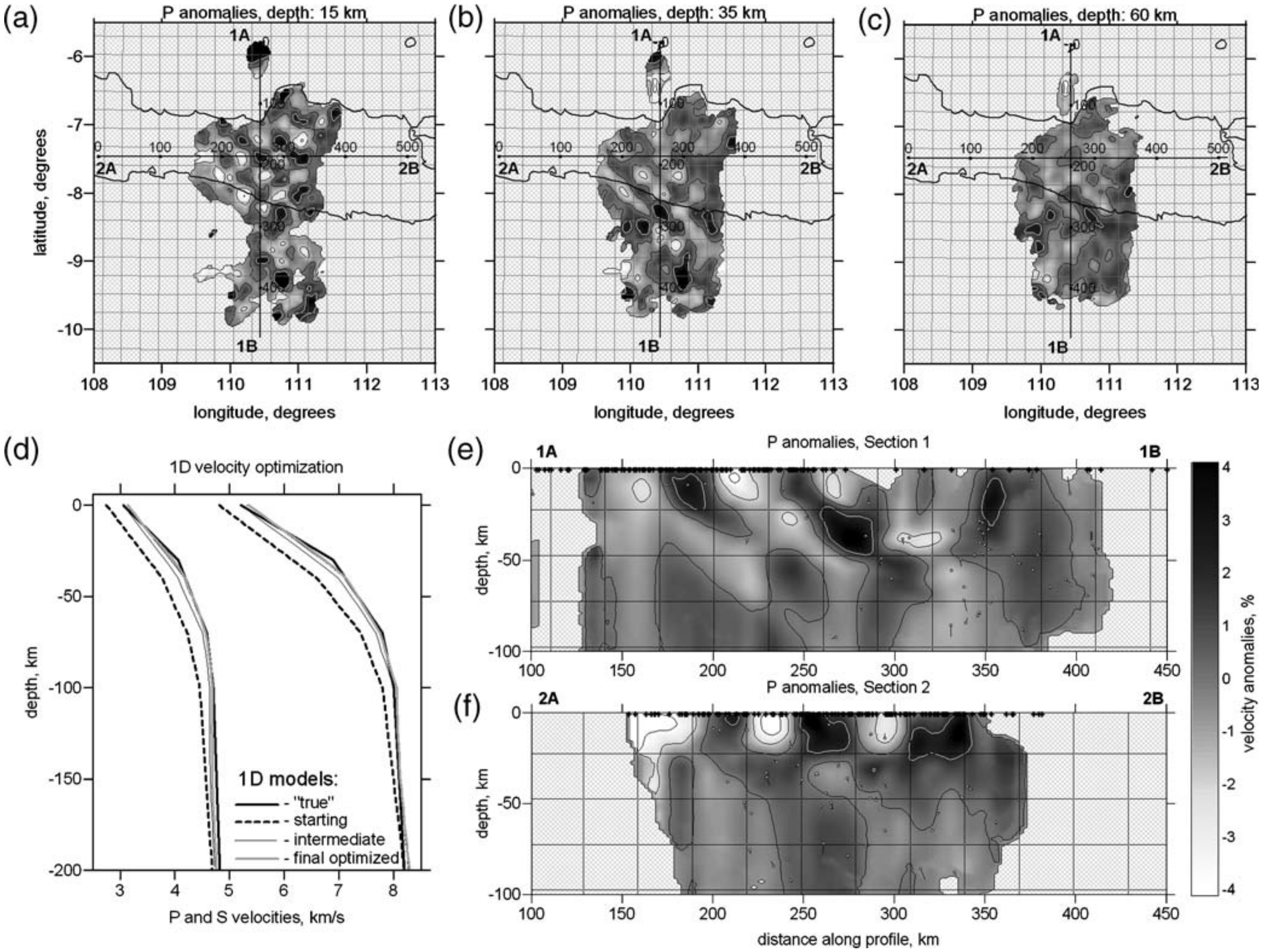

Figure 3. Inversion results using the sources selected according to GAP < 180°. (a, b, c) Resulting *P* velocity anomalies in horizontal sections (contour lines are shown at levels of −2%, 0 and 2%). Gray grid lines indicate the locations of the checkerboard anomalies in the synthetic model. (d) Result of the 1D model optimization. Black line is the true model; dotted black line is the starting model; bold gray line is the retrieved model. (e, f) Resulting velocity anomalies of *P* velocities in two vertical sections (positions of the sections are indicated in a–c). Gray dots are relocated sources at distances of up to 200 km from the profiles; bars show errors with respect to true coordinates. Ⓔ Color version is available in the electronic edition of *BSSA*.

Processing begins with the preliminary location of the sources and finding an optimal 1D model, V_{res}. The algorithm for the preliminary location of sources is based on a 1D velocity model and uses tabulated travel times. This allows performing grid searching for the hypocenter coordinates that is rather fast and stable, even when a true event is located far from the initial point of search (e.g., at distances of 400–500 km). 1D velocity optimization is performed iteratively in parallel with the grid search location of sources. The resulting distributions of the 1D *P* and *S* velocities for three considered models are shown with gray lines in plots D in Figures 3–5. In all cases the retrieved velocity distribution fit the true model (black bold lines) rather well. However, for model 3, which contains many long ray paths from the out-of-network events, the optimized velocity seems to be higher than the true one. This can be explained by the fact of dominating negative residuals in the synthetic dataset caused by preferential traveling of the bending rays through positive velocity blocks.

The derived 1D distribution is used as a starting velocity model for an iterative tomographic inversion and further source relocations based on the LOTOS code (Koulakov, 2009). It is important that the velocity model is computed in several parameterization grids with different predefined orientations (e.g., 0, 22, 45, and 67 deg) and then averaged into one model. The parameterization nodes are installed according to the ray distributions. The minimal node spacing is 5 km, and it is much smaller that the size of the synthetic patterns (30 km). In this case the solution is practically independent of the grid configurations, and the parameterization can be called quasicontinuous. For model 3 the numbers of parameters were ~23000 for the *P* model and ~21000 for the *S* model.

The resulting locations of events after five iterations of simultaneous inversion for velocity distribution and source

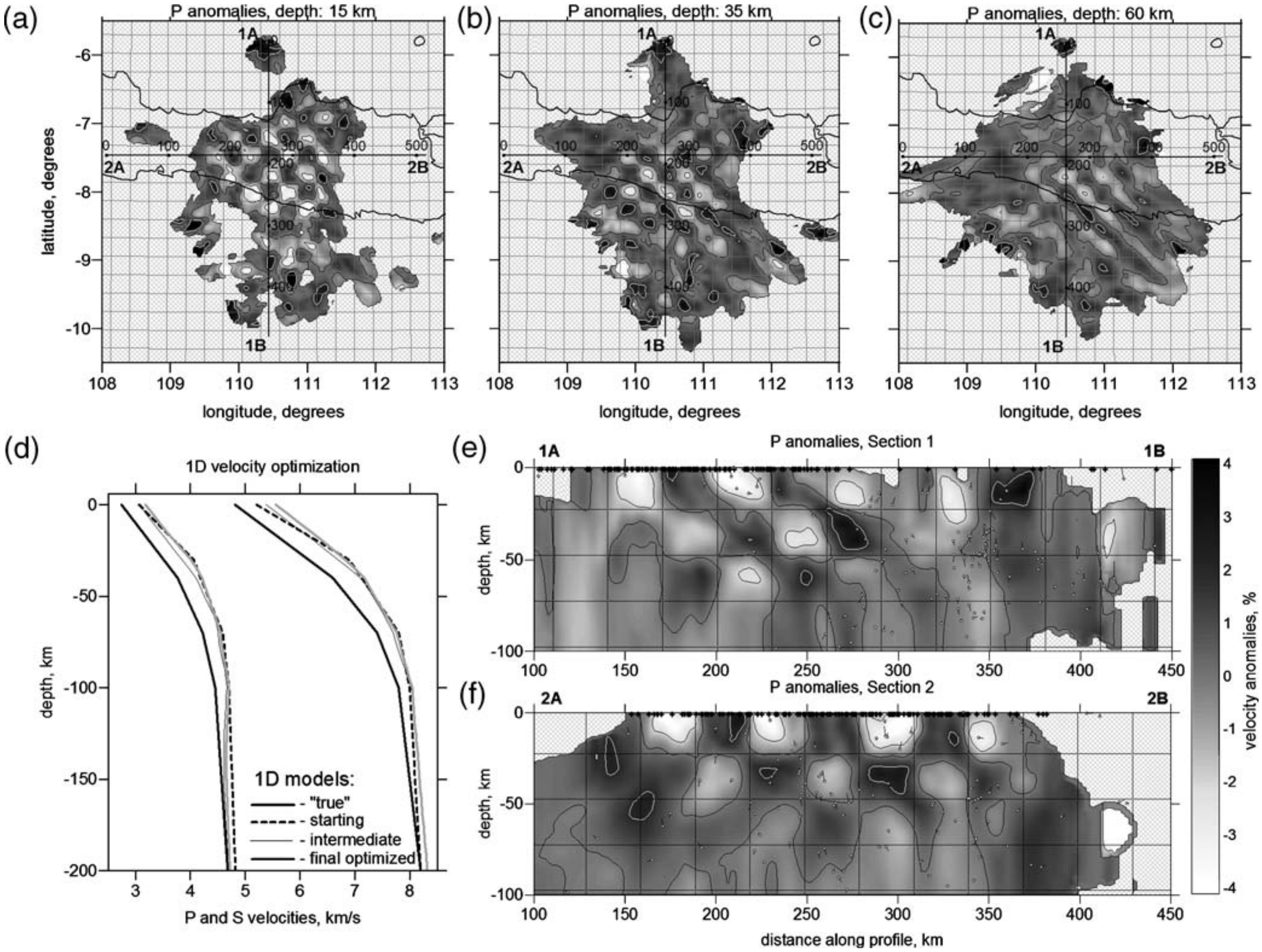

Figure 4. Inversion results using the sources selected according to GAP < 280°. Meaning of plots and symbols are the same as in Figure 3. ⒺColor version is available in the electronic edition of *BSSA*.

parameters are shown in Figure 2c. It can be seen that the 3D inversion leads to a considerable decrease of mislocation of the events. However, for events located far outside the network the error remains important. For some events it reaches 20–40 km. Nevertheless, as will be shown later, adding such events although located with considerable error, does not prevent but helps to improve the quality of tomographic inversion.

The results of *P* velocity reconstructions for the three considered models are presented in Figures 3–5 in three horizontal and two vertical sections. The horizontal sections at depths of 15, 35, and 60 km correspond to three upper layers of the different checkerboard polarities. The results for *S* velocity are practically identical and are not shown here. Model 1 (Fig. 3), for which only events with GAP < 180° were selected, provides rather poor reconstruction quality, especially in the lower section. The satisfactory reconstruction quality is only observed for the uppermost layer (15 km depth). In the second layer, only qualitative correspondence with the checkerboard structure is observed in the central part of the study area. For the third layer the reconstructed patterns seem to be chaotic. In the vertical sections the anomalies are strongly smeared diagonally. For model 2 (Fig. 4) with GAP limitation of 280°, the reconstruction quality is much higher. The good resolution area is larger in both horizontal and vertical sections. The uppermost two layers of the checkerboard are reconstructed robustly and in the third layer the general periodicity of anomalies is visible. However, the best results are obtained for the case when all events within the radius of 5° are used (model 3, Fig. 5). Indeed, the checkerboard structure is robustly reconstructed in the three uppermost layers of the checkerboard. The higher quality of reconstruction in model 3 is related to two advantageous factors of using out-of-network events: a significant increase in the amount of data and much better spatial coverage of the ray paths. Their positive contribution appears to be quite important, and it completely compensates for any negative effect related to the poorer locations of out-of-network events.

Similar tests were performed for different realistic and artificial observation schemes. In particular, my colleagues

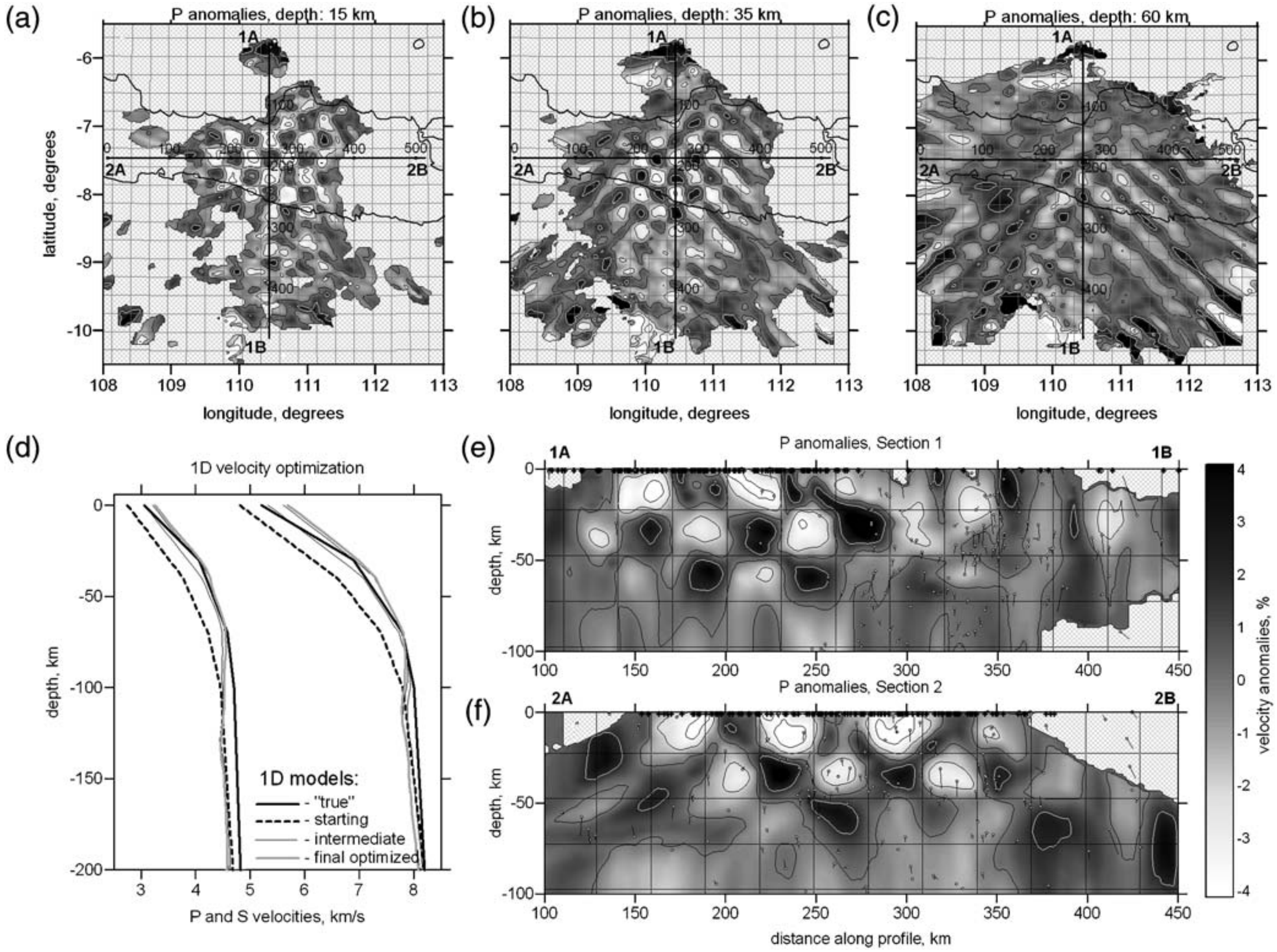

Figure 5. Inversion results using all sources within the area of 5° with respect to the center of the network. Meaning of plots and symbols are the same as in Figure 3. ⒺColor version is available in the electronic edition of *BSSA*.

and I have found a positive effect from using the out-of-network events for real networks in the area of North Anatolian fault and Toba Caldera (Indonesia). However, due to the limited space of this short note, we cannot present these results here.

Conclusions

The results of three tests presented in this article show that rejecting the sources outside the network is contraindicated or at least inadequate in LET modeling. It was shown here that for the considered observation schemes, increasing the area for the event selection improves the quality of the tomographic inversion. The same results have been obtained for several other realistic and artificial experiment configurations. I did not check larger than 5° areas for the source selection, but I suggest that adding events located at larger distances may cause further improvement of the results. It will probably lead to increase in the variability of incidence angles and azimuths of rays in the target volume, which is obviously favorable for tomographic inversion. In particular, I suggest that it would be very fruitful to combine local and teleseismic data. For example, in some subduction zones deep seismic clusters in the slab are focused in narrow zones. The rays from these clusters are quasiparallel and often cause strong smearing of the resulting anomalies. If these rays are combined with teleseismic rays coming from other directions it would obviously improve the spatial resolution and would increase depth penetration of tomographic models.

At the same time I admit that I cannot consider all the variety of possible source/receiver configurations that could exist in LET schemes. Thus, I do not pretend that the statement about improving the results when out-of-network sources are used is universal in all situations. For example, I expect that adding out-of-network events that are clustered in one localized area would possibly not provide improvement of the results and might cause smearing of the resulting anomalies. On the other hand, it should be mentioned that similar effect of smearing caused by clustering may equally appear for events located inside the network, so it is not a specific problem of using out-of-network events.

The possibility of the existence of opposite cases means that the data selection criteria should be checked individually for each source/receiver configuration. The most clear and effective way is to perform synthetic modeling, such as a checkerboard test as was done in this study, for various datasets. In my opinion, it is impossible to propose a universal scheme for data selection that would be suitable for any possible experiment setup.

In any case, the rays from events at any distances provide some information about the deep structure if they travel through the target area. The purpose of tomography is to decipher it. Most LET algorithms are not working *a priori* with data from out-of-network events, and their users usually adapt the datasets to the existing codes and reject a lot of relevant information. It seems to me that it would be much more constructive to adapt the algorithms to the existing data and to use as much of available information as possible including events from all ranges of epicentral distances.

The strong positive effect of adding out-of-network events demonstrated in this study shows that the strategy of LET experiment performed in the commonly accepted practice should be revised. It is clear now that executing costly work on performing experiments and then rejecting most of relevant information just because of the GAP stereotype is very inefficient. This causes strong limiting of the resolution and decreasing the scientific importance of the results. In my opinion, all the published results on local earthquake tomography that used only events with GAP < 180 could be significantly improved if the datasets were revised to include a broader distribution of sources.

Data and Resources

In the article I present synthetic data based on real configuration of sources and receivers corresponding to a real experiment in Central Java (Koulakov *et al.*, 2007). A part of the events used in modeling is taken from the International Seismological Center (ISC, 2001) catalog. The LOTOS code for tomographic inversion that is used in this study is freely available online at www.ivan-art.com/science/LOTOS_09 (last accessed March 2009). This Internet site provides an executable version of the code, its detailed description, manuals, examples of observed and synthetic datasets, and other necessary information. The three synthetic datasets presented in this article are available online in the file: www.ivan-art.com/science/LOTOS_09/data_for_lotos_9.zip in folders, GAP__180, GAP__280, and GAP__360. All the pictures presented in this article can be easily reproduced by any person by running the executable version of the LOTOS-09 code for each of these datasets.

Acknowledgments

This paper was inspired by criticism of several anonymous reviewers who rejected my previous papers just because of using a more liberal than GAP < 180° selection criteria. Heated discussions with Edi Kissling were very important in constructing the argumentation scheme of this article. I am very grateful to Andrew J. Michael for constructive suggestions on preparing this study. I thank the associate editor Diane I. Doser for important comments and help in editing the text of the manuscript. This study is supported by the Russian Foundation for Basic Researches (RFBR; 08-05-00276-a), Heimholtz Society, and RFBR joint research project 09-05-91321-SIG_a, multidisciplinary projects SB RAS 44, 21 and SB-UrO-DVO RAS 96, and project ONZ RAS 7.4.

References

Aki, K., A. Christofferson, and E. S. Husebye (1977). Determination of the three dimensional structure of the lithosphere, *J. Geophys. Res.* **82,** 277–296.

Barberi, G., M. T. Cosentino, A. Gervasi, I. Guerra, G. Neri, and B. Orecchio (2004). Crustal seismic tomography in the Calabrian Arc region, south Italy, *Phys. Earth Planet. Inter.* **147,** 297–314.

Chiarabba, C., and A. Amato (2003). V_p and V_p/V_s images in the M_w 6.0 Colfiorito fault region (central Italy): A contribution to the understanding of seismotectonic and seismogenic processes, *J. Geophys. Res.* **108,** no. B5, 2248, doi 10.1029/2001JB001665.

Daly, E., D. Keir, C. J. Ebinger, G. W. Stuart, I. D. Bastow, and A. Ayele (2008). Crustal tomographic imaging of a transitional continental rift: The Ethiopian rift, *Geophys. J. Int.* **172,** 1033–1048.

Dias, N. A., L. Matias, N. Lourenço, J. Madeira, F. Carrilho, and J. L. Gaspar (2007). Crustal seismic velocity structure near Faial and Pico Islands (AZORES), from local earthquake tomography, *Tectonophysics* **445,** 301–317.

Eberhart-Phillips, D., and S. Bannister (2002). Three-dimensional crustal structure in the Southern Alps region of New Zealand from inversion of local earthquake and active source data, *J. Geophys. Res.* **107,** no. B10, 2262, doi 10.1029/2001JB000567.

Evans, J., and U. Achauer (1993). Teleseismic velocity tomography using the ACH method: theory and application to continental scale studies, in *Seismic tomography: Theory and Practice*, H. Iyer and K. Hirahara (Editors), Chapman and Hall, London, 319–360.

Husen, S., and R. B. Smith (2004). Probabilistic earthquake relocation in three-dimensional velocity models for the Yellowstone National Park region, Wyoming, *Bull. Seismol. Soc. Am.* **94,** 880–896.

Husen, S., E. Kissling, and E. R. Flueh (2000). Local earthquake tomography of shallow subduction in north Chile: A combined onshore and offshore study, *J. Geophys. Res.* **105,** 28,183–28,198.

Husen, S., R. Quintero, E. Kissling, and B. Hacker (2003). Subduction-zone structure and magmatic processes beneath Costa Rica constrained by local earthquake tomography and petrological modeling, *Geophys. J. Int.* **155,** 11–32.

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

International Seismological Centre (2001). Bulletin Disks 1-9 [CD-ROM], *Internatl. Seis. Cent.*, Thatcham, United Kingdom.

Kato, A., T. Iidaka, E. Kurashimo, S. Nakagawa, N. Hirata, and T. Iwasaki (2007). Delineation of probable asperities on the Atotsugawa fault, central Japan, using a dense temporary seismic network, *Geophys. Res. Lett.* **34,** L09318, doi 10.1029/2007GL029604.

Koulakov, I. (2009). LOTOS code for local earthquake tomographic inversion. Benchmarks for testing tomographic algorithms, *Bull. Seismol. Soc. Am.* **99,** no. 1, doi 10.1785/0120080013.

Koulakov, I., M. Bohm, G. Asch, B.-G. Lühr, A. Manzanares, K. S. Brotopuspito, P. Fauzi, M. A. Purbawinata, N. T. Puspito, A. Ratdomopurbo, H. Kopp, W. Rabbel, and E. Shevkunova (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.

Lippitsch, R., E. Kissling, and J. Ansorge (2003). Upper-mantle structure beneath the Alpine orogen from high-resolution teleseismic tomography, *J. Geophys. Res.* **108,** no. B8, 2376–2390.

Paul, A., M. Cattaneo, F. Thouvenot, D. Spallarossa, N. Bethoux, and J. Frechet (2001). A three-dimensional crustal velocity model of the south-western Alps from local earthquake tomography, *J. Geophys. Res.* **106,** no. B9 19367–19389.

Reyners, M., D. Eberhart-Phillips, G. Stuart, and Y. Nishimura (2006). Imaging subduction from the trench to 300 km depth beneath the central North Island, New Zealand, with V_p and V_p/V_s, *J. Geophys. Res.* **87,** no. B5, 4073–4088.

Sandoval, S., E. Kissling, and J. AnsorgeThe SVEKALAPKO STWG (2004). High-resolution body wave tomography beneath the SVEKALAPKO array: II. Anomalous upper mantle structure beneath the central Baltic Shield, *Geophys. J. Int.* **157,** 200–214.

Thurber, C. H., S. R. Atre, and D. Eberhart–Phillips (1995). Three-dimensional V_p and V_p/V_s structure at Loma Prieta, California, from local earthquake tomography, *Geophys. Res. Lett.* **22,** no. 22, 3079–3082.

Vlahovic, G., and C. A. Powell (2001). Three-dimensional *S*-wave velocity structure and V_pV_s ratios in the New Madrid seismic zone, *J. Geophys. Res.* **106,** no. B7, 13,501–13,514.

Institute of Petroleum Geology and Geophysics
SB RAS, Prospekt Akademika Koptuga, 3
Novosibirsk 630090, Russia
KoulakovIY@ipgg.nsc.ru

Manuscript received 15 December 2008