Resolution analysis of joint inversion of seismic receiver function and surface wave dispersion curves in the "13 BB Star" experiment

Kajetan Chrapkiewicz[1,2], Monika Wilde-Piórko[1,3], Marcin Polkowski[1], and Marek Grad[1]

[1]University of Warsaw, Faculty of Physics, Institute of Geophysics, Warsaw, Poland
[2]currently at Imperial College London, Department of Earth Science and Engineering, London, United Kingdom
[3]currently at the Institute of Geodesy and Cartography, Warsaw, Poland

Correspondence to: Kajetan Chrapkiewicz (k.chrapkiewicz17@imperial.ac.uk)

Abstract. ~~Joint inversion of Rayleigh wave phase velocity dispersion and P receiver function has been applied to study the structureof the upper mantle beneath the south-western margin of the East European Craton. The data were gathered in the passive seismic experiment "13 BB Star" (2013–2016) in the area of the crust recognized from previous borehole and refraction surveys~~A multistep workflow for a credible simultaneous inversion of surface wave dispersion and receiver function data has been proposed as a tool to illuminate a whole-lithosphere structure. Several fundamental issues inherent in the ~~linearised inversion were addressed in this work~~linearized inversion are addressed in it, including exploitation of a priori knowledge, choice of model's depth, trapping by local minima of the misfit function associated with non-uniqueness of the ~~misfit-function~~ optimization problem, a proper weighting of data sets characterized by different uncertainties, and credibility of the final models. The last was investigated with the aid of novel 1D checkerboard tests ~~juxtaposed with resolution matrixanalysis~~ – a more intuitive and feasible alternative to the resoluton matrix. We advocate the usefulness of ~~linearised~~ the linearized approach when handled with proper care, ~~and show that~~ including the resolution analysis ~~is~~ as an indispensable step when choosing the inversion parameters. ~~It allowed us to obtain reliable S-wave velocity models down to 200~~

We applied our workflow to study the south-western margin of the East European Craton. Rayleigh wave phase velocity dispersion and P-wave receiver function data were gathered in the passive seismic experiment "13 BB Star" (2013–2016) in the area of the crust recognized from previous borehole and refraction surveys. Final models of S-wave velocity structure down to 300 km depth beneath the ~~"13 BB Star" array , indicating the presence of a Paleozoic asthenosphere and the ceiling of the deeper, Precambrian,~~ array are characterized by proximity in the model space and good data fit. A low-velocity zone starting at the depth $180 - 200$ km and $v_S \approx 4.85$ km/s is likely to indicate the deep cratonic lithosphere-asthenosphere ~~transition zone.~~ boundary.

1 Introduction

~~Limited resolvable depth and sensitivity are the two major problems inherent in receiver~~ Receiver function (RF) and surface wave dispersion (SWD) analysis ~~. The first issue results mainly from the presence of noise and imperfection of the measurements. The second constitutes the intrinsic feature of the data. Combined, they make the credible models of subsurface difficult to obtain. Although it is the inversion of RF which is usually regarded as exceptionally non-unique (Ammon et al., 1990; Bodin et a~~

~~the SWD lacks uniqueness either, not being able~~ both suffer from their inherent limitations. In the former case it is a non-uniqueness of the solution caused by a trade-off between the depth of the discontinuity and the velocity of the overburden (Ammon et al., 1990) . The latter often fails to discriminate between ~~the fine structures , especially at greater depths~~ fine structures due to a substantial width of surface waves sensitivity kernels (Tsuboi and Saito, 1983; Romanowicz, 2002).

The simultaneous inversion of RF and SWD, having been applied to study deep lithosphere for 20 years (Özalaybey et al., 1997; Du and Foulger, 1999; Julià et al., 2000), mitigates these issues thanks to the complementary sensitivity of the data~~(Shen et al., 2013)~~ . Compared to inversion of each of these data types alone, it provides better vertical resolution than SWD, and, unlike RF, constrains absolute shear velocities (e.g. Shen et al., 2013) . Even in this case however, there ~~remains~~ may remain a certain ambiguity ~~, showed in the following sections,~~ that should be taken into careful consideration throughout inversion. A single model obtained by minimization of the data-misfit function cannot be regarded as a full solution of the inverse problem (Gubbins, 2004), which is often ~~a case of linearised~~ the case of linearized inversion (Julià et al., 2003; Horspool et al., 2006; Wang et al., 2014; Sosa et al., 2014; Bao et al., 2015; Li et al., 2016). One needs to probe the whole space of physically plausible solutions, by drawing inferences from an ensemble (Sambridge and Mosegaard, 2002). Bayesian Monte Carlo methods (Green and Hastie, 2009; Bodin et al., 2012; Shen et al., 2013; Deng et al., 2015; Fontaine et al., 2015) appear to be a natural approach to achieve this goal, not only performing importance sampling (Sambridge, 1999), but also properly exploiting a priori knowledge of the problem (Malinverno, 2002). They are not flawless though, but suffer from their own inherent nuisance, e.g. the lack of objective criterion of convergence (Roberts et al., 1996, 1997), not to mention the computational cost. They also require nontrivial tuning to prevent solutions from following the shape of prior velocity distribution other than homogeneous ~~(Minato et al., 2008; Wathelet, 2008)~~ (Minato et al., 2008) .

Here we propose the compromise between the simplicity and transparency of the method on the one hand, and a full solution of the inverse problem on the other, by performing ~~linearised inversion with~~ a linearized inversion with the ensemble of starting models covering the entire space of acceptable solutions. A workflow involving synthetic tests and resolution analysis was applied to tune the inversion parameters and determine expected resolving power of the data. We demonstrate this approach with teleseismic data collected in the area of well-recognized crust – the a priori knowledge which we introduce into ~~inversion. We show a remnant non-uniqueness of the joint RF and SWD inversion, which justifies the careful selection of the values of initial parameters. To this end, we perform the resolution analysis based on novel 1D checkerboard tests and resolution matrices before the actual inversion~~ the inversion in two steps.

2 Data

Passive seismic experiment “13 BB Star” was dedicated to study ~~the~~ deep structure of the Earth’s interior in the marginal zone of the East European Craton (EEC) in northern Poland (Grad et al., 2015)~~. The seismic network consisted of 13 broadband stations on the area of ca. 120 km in diameter~~ , close to the Trans-European Suture Zone (TESZ) (Fig. 1). The TESZ, a contact zone between the EEC and the Palaeozoic Platform (PP) located to SW, is referred to as the most fundamental lithospheric boundary in Europe (Pharaoh, 1999) . In large part it is located in Poland, hidden beneath the cover of sedimentary rocks at least

10 km thick (Fig. 1). ~~Location of "13 BB Star" broadband seismic stations on the background of tectonic map of north Poland and scheme of Europe. TESZ – Trans-European Suture Zone, Pre-Karelian granitoid massifs: Pm, Pomorze; Db, Dobrzy; Cn, Ciechanw; RG, rapakivi granites; Pre-Karelian metamorphic belts: Kb, Kaszuby; Karelian metamorphic-magmatic complexes: Kmr (Kampinos); (Ryka, 1982) . Epicentral distribution of earthquakes used for calculation of receiver functions and dispersion curves of surface waves plotted relatively to a position of A0 central station of network. The network was located~~ The TESZ is associated with a steep dip in the Moho depth, from 30 – 35 km in the PP to 42 – 52 km in the EEC over a short lateral distance of less than 200 km (Grad et al., 2016) . The "13 BB Star" network of ca. 120 km in diameter consisted of 13 broadband stations in the area of ~~well-known sedimentary cover and crustal structure~~ relatively flat Moho topography (about 40-45 km variation beneath the stations), favourable from the point of view of deep structure studies. The sedimentary and consolidated crust beneath the array has been well-recognized, mainly with refraction and borehole data (Grad et al., 2009; Polkowski and Grad, 2015; Grad et al., 2016)~~.~~ , the upper-mantle structure however remains poorly understood.

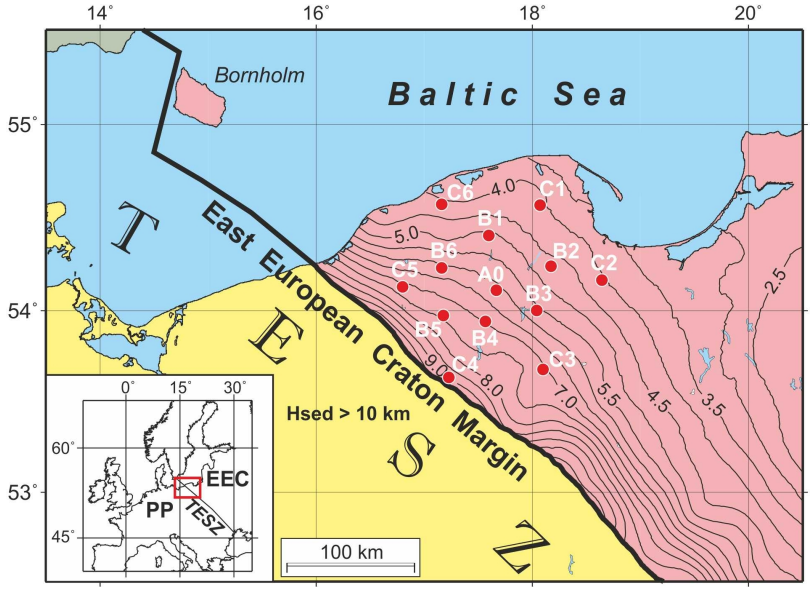

Figure 1. Location of the "13 BB Star" broadband seismic stations (red dots and white codes) on the background of the main tectonic features of the region; Key: pink color – East European Craton; TESZ – Trans-European Suture Zone; isolines: thickness (km) of the sedimentary cover of the EEC.

The "13 BB Star" ~~seismic~~ stations were equipped with the Reftek 151-120 Observer seismometers and Reftek 130 data ~~logger. The stations were~~ loggers and operated from June 2013 to October 2016. More technical details about the experiment can be found in (Grad et al., 2015) . The distance and azimuthal epicentral distribution of ~~analysed earthquakes~~ the subset of the earthquakes used in this study are shown in Fig. 2.

2.1 P receiver function

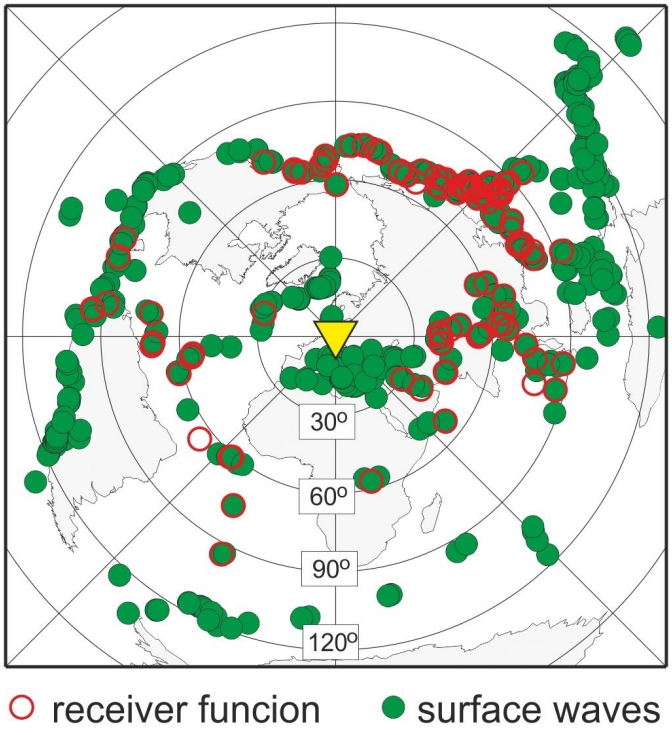

Figure 2. Epicentral distribution of the earthquakes used for the calculation of the receiver functions and the dispersion curves of surface waves plotted relative to the position of the A0 central station of the network.

~~The receiver functions techniques (Langston, 1977; Vinnik, 1977) have been used to investigate the structure of the lithosphere-asthenosphe system. Receiver functions~~

Receiver functions (Langston, 1977; Vinnik, 1977) were calculated by slightly modified method presented by Wilde-Piórko (2015); Wilde-Piórko et al. (2017)~~. Seismograms used in receiver function analysis has been selected manually~~ , from the seismograms selected manually corresponding to the events of the magnitude 5.7 and higher. Additionally, the second manual selection was done after the calculation of receiver functions to choose the traces with the highest signal to noise ratio. Ultimately, the total number of 99 events within 30-100° epicentral distance range was taken into consideration. ~~Seismograms~~ Their seismograms were cut 300 s before and 300 s after the theoretical P-onset calculated for the *iasp91* model (Kennett and Engdahl, 1991)~~for earthquakes with epicentral distance of 30-100° and viewed to leave for further analysis only that ones with~~ . Those with no visible energy on vertical components were rejected. One-pass low-pass filtering with Butterworth filter of corner frequency 5 Hz was applied before resampling seismograms to 20 Hz. Then, seismograms were cut in time window, 100 s before and 100 s after the onset of direct P-wave calculated due to *iasp91* model.

Calculation of backazimuth ~~and polarization angle were performed in two steps: (1)~~ angle was performed in the following steps: seismograms were filtered with two-pass band-pass Butterworth filter with corner periods of 2 and 10 s for searching the backazimuth angle for teleseismic P-wave; N and E components of seismograms were rotated with backazimuth angle

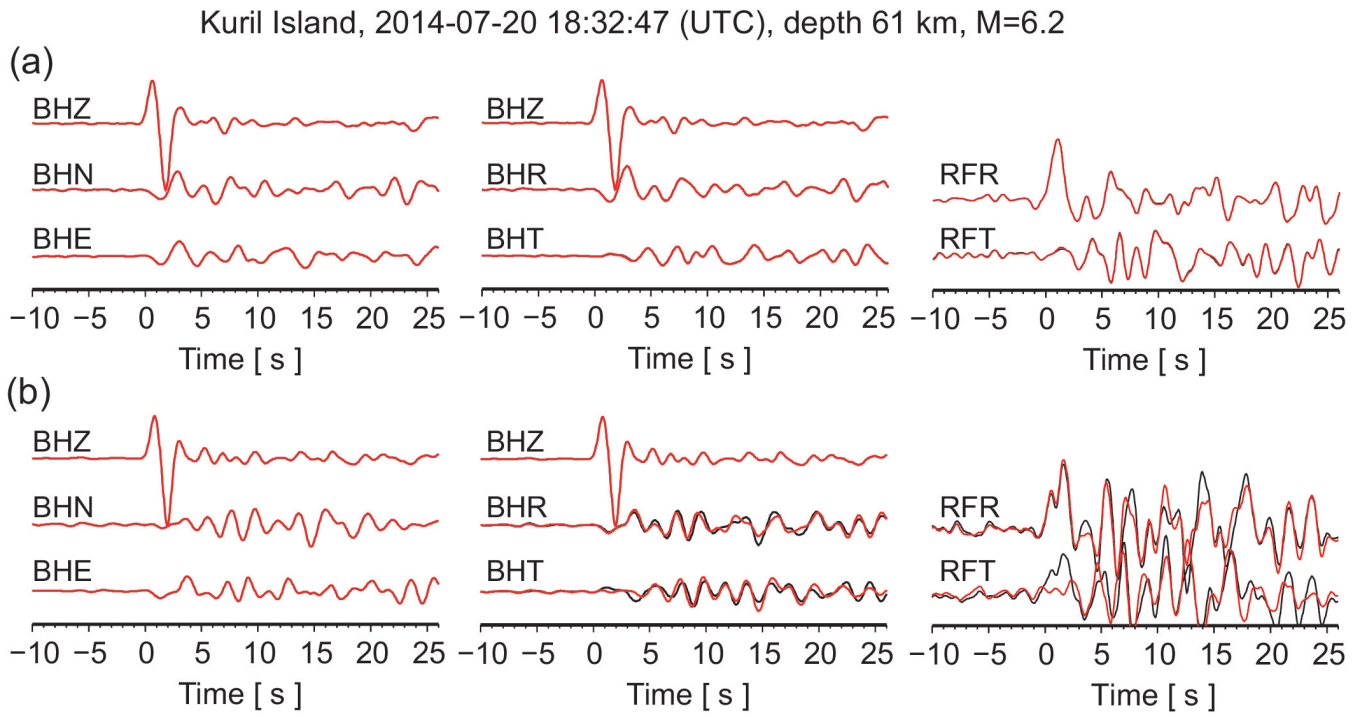

Figure 3. Example of the rotation of teleseismic P-waves and RFs for~~the~~ : (a) C6; and (b) B6 seismic stations of "13 BB Star" array calculated ~~by~~ using the RF-rotation procedure (red line) and theoretical backazimuth ~~and polarization~~ angles (black line). Seismograms and ~~RF~~ RFs are filtered with band-pass Butterworth filter of corner frequencies 0.01 and 0.8 Hz. Time zero refers to the direct P-wave.

from 0 to 360° every 1° and radial receiver function (RFR) were calculated by time-domain Wiener deconvolution; next RFR with maximal sum of amplitudes between time 0 and 1 s were selected (equivalent of rotation ~~ZN-E components to ZRT); (2) seismograms were filtered with two-pass band-pass Butterworth filter with corner periods of 2 and 10 s for searching the polarization angle for teleseismic P-wave; seismograms were rotated from ZNE to ZRT components by backazimuth angle found in step (1) and then rotated from ZRT to LQT with polarization angle from 0~~ ZNE components to ~~45° every 1° and Q receiver function (RFQ) were calculated; next RFQ were cut 5 s before and 5 s after direct P-wave, a mean value and trend were removed and root square mean (rms) values were calculated for each RFQ for time window between -2 and 0 s; then as best were chosen RFQ for which the rms value was lower than rms calculated for RFQ of next polarization angle (equivalent of rotation ZRTcomponents to LQT) . The final RF were filtered with Gaussian filter with parameter 4.~~ ZRT) as final RFR.

An example of presented procedure is shown in Fig. 3 together with RFs calculated in ~~traditional way – ZNE components were rotated to LQT ones for theoretical backazimuth and polarization angles~~a traditional way. The final RFR used in ~~linearised~~ linearized joint inversion were filtered with Gaussian filter with parameter 4, move-out corrected for mean slowness of each station and stacked (Fig. 4a).

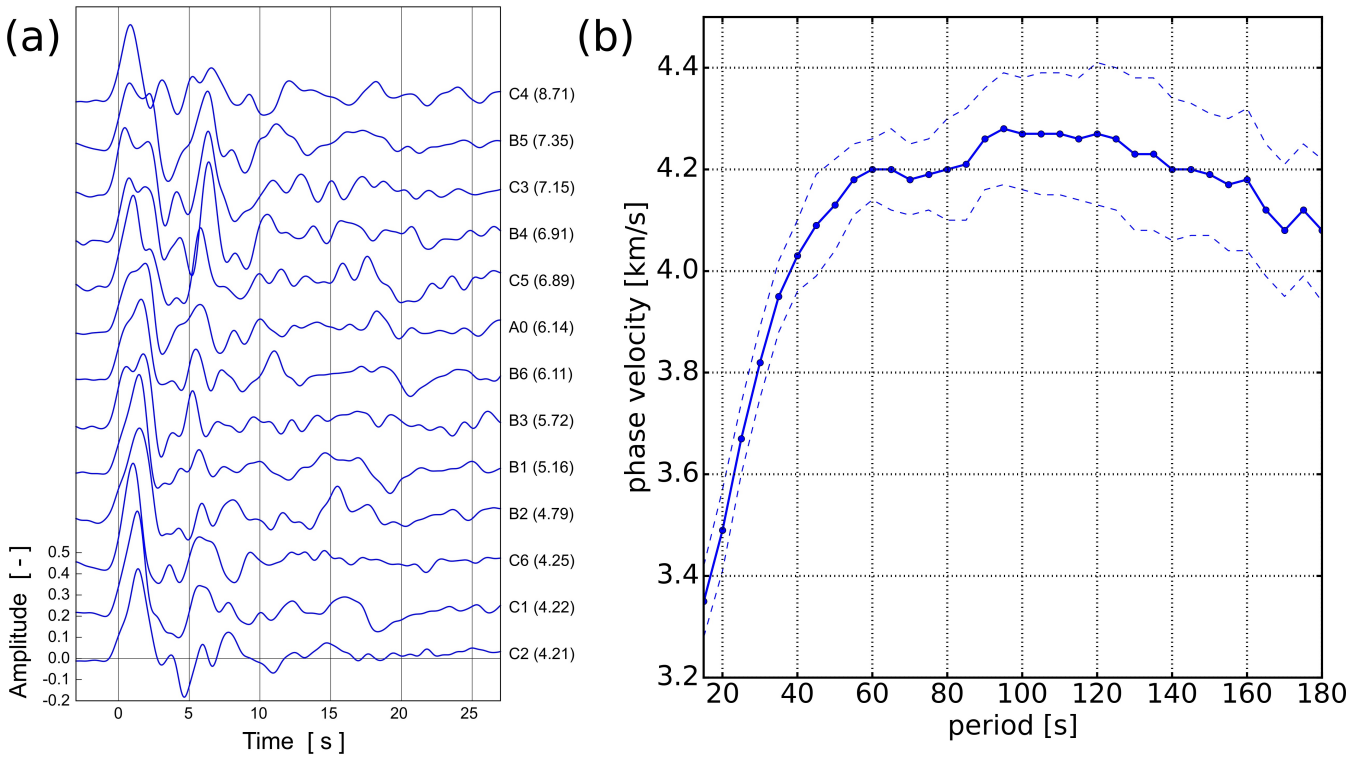

Figure 4. ~~Stacked RFR for all~~ Data gathered in the "13BB Star" experiment: (a) stacked RFR for all broadband seismic stations ~~. RFs were calculated based on~~ ordered by thickness [km] of the ~~RF-rotation procedure. Time~~ sedimentary cover (in parantheses after the stations codes); time zero refers to the direct P-wave; (b) fundamental mode Rayleigh wave dispersion curve for "13BB Star" array, the mean (solid blue line with points) and its standard deviation (dashed blue lines).

2.2 Fundamental mode Rayleigh wave phase velocity dispersion

To calculate dispersion curves ~~from surface waves the ASWMS~~ of the surface waves recorded in the experiment, the ASWMS (Automated Surface Wave Phase Velocity Measuring System) package was used (Jin and Gaherty, 2015). It allows measuring phase and amplitude of surface waves from raw seismic waveforms and ~~use~~ uses these measurements to calculate phase velocity maps for different wave periods using Eikonal and Helmholtz equation.

The ASWMS ~~package bases calculation on~~ calculation is based on the surface waves incoming to an array of stations ~~form~~ from natural teleseismic events. Before performing phase velocity calculation, data obtained during the "13 BB Star" experiment required preparation. Firstly, a list of all recorded events with magnitude higher than 4 was created and manually checked for quality. ~~Total~~ No prior constraints on hypocentral depth were involved. A total of 706 events ~~from~~ recorded in the period from 2013-07-10 to 2016-09-23 were selected for the further study. All ~~706 events~~ of them were manually verified on a station basis. For each event a list of stations with good quality recording was prepared. ~~Fundamental mode Rayleigh wave dispersion curve for "13BB Star" array, average with $\pm 3\sigma$ deviation.~~ ~~Stations~~ The stations with no data, partial data or bad

data quality were omitted. ~~135 events were selected as good quality~~ The number of 135 events was selected as good-quality on all 13 stations, 478 events were recorded with good quality on over 50 % of ~~stations~~ the stations, and 103 events were rejected on all stations. From all 9178 station-event pairs, 5697 were selected as ~~good quality~~ good-quality (over 62 %). ~~In average~~On average, 8 stations were selected as ~~good quality~~ good-quality per event. The depth of the source was lower than 100 km for most earthquakes. For each station-event pair raw seismic data was converted ~~form~~ from daily mini-seed files to single SAC files with embedded event parameters (latitude, longitude, depth and magnitude). SAC data files were then imported by ASWMS package and used for phase velocity calculations. The primary result of the calculations are phase velocity maps for different wave periods. In this case maps were calculated for periods from 10 to 300 s every ~~5 s~~ 1 s without smoothing. Results were then prepared as dispersion curves, with values taken from corresponding grid cells on all maps. ~~Final~~ Each cell was $0.05° \times 0.05°$ big. The final dispersion curve was obtained by averaging all dispersion curves for each period and is shown ~~as green dots and green line~~ in Fig. ~~5, with $\pm 3\sigma$ deviation.~~ 4b.

3 Inversion

We used a ~~linearised~~ linearized damped least-squares inversion scheme implemented in ~~CPS~~ the "Computer Programs in Seismology" (CPS) package (Julià et al., 2000; Herrmann, 2013), appropriate for over-determined problems such as ~~inversion of RF~~an inversion of a receiver function. The bottom line of every linearized inversion is to iteratively minimize some scalar-valued function (e.g. expressing the misfit between the data) starting from a certain initial model. The main step of ~~the procedure is to find~~ this procedure consists in finding a general inverse of the forward operator matrix (~~forward function linearised in~~ full forward function linearized in the current iteration), which is usually done with a singular value decomposition. The ultimate goal is to find all models explaining the data and remaining consistent with the prior knowledge of the problem. S-wave velocity was the only parameter inverted here. P-wave velocities and densities were updated based on respectively v_P/v_S and ρ/v_S ratios determined in the starting models.

3.1 Theoretical prerequisites

~~The~~ A joint data vector is ~~a simple concatenation of~~ concatenation of components corresponding to each data type, here RF and SWD~~components, : therefore~~

$$\boldsymbol{d} = \left(\boldsymbol{d}^{RF}, \boldsymbol{d}^{SWD}\right)^T = \begin{pmatrix} \bar{r}_1 \pm \sigma_{r_1} \\ \bar{r}_2 \pm \sigma_{r_2} \\ \vdots \\ \bar{r}_R \pm \sigma_{r_R} \\ \bar{s}_1 \pm \sigma_{s_1} \\ \bar{s}_2 \pm \sigma_{s_2} \\ \vdots \\ \bar{s}_S \pm \sigma_{s_S} \end{pmatrix} \tag{1}$$

where r_i stands for one of R number of RF samples and s_j for one of S phase velocity values. In this respect, joint inversion of two or more data sets doesn't differ much from the inversion of a single data type. For such a data vector, one can propose a simple misfit function (Julià et al., 2000):

$$|\Delta d|^2 = \frac{1-p}{R}\sum_{i=1}^{R}\left(\frac{\Delta r_i}{\sigma_{r_i}}\right)^2 + \frac{p}{S}\sum_{i=1}^{S}\left(\frac{\Delta s_i}{\sigma_{s_i}}\right)^2, \tag{2}$$

where ~~r_i stands for~~ Δr_i denotes the difference between observed and synthetic data for the i^{th} ~~of R number of RF samples, s_j is a~~ RF sample, Δs_i for the j^{th} ~~of S number of phase velocity values, and~~ point of the dispersion curve, and $|\Delta d|^2$ expresses a modified Mahalanobis distance (Mahalanobis, 1936) between random vectors in case of no correlation (we assume a diagonal covariance matrix ~~, that~~ which in general is not true, especially for RF).

The parameter p is added to allow weighting the relative influence of each data set. For $p = 0$ we invert only RF and for $p = 1$ only SWD. The thing worth noticing is that the uncertainties (denoted with σ) are squared, and thus contribute to the weighting of each datum more than p.

In order to stabilize the solution, it is desirable to add an extra term to the misfit function resulting in a trade-off between final data fit and smoothness of the model:

$$\Phi(\boldsymbol{m}) = |\Delta \boldsymbol{d}|^2 + \theta^2 \cdot |\mathbf{T}(\Delta \boldsymbol{m})|^2, \tag{3}$$

where θ^2 is a positive parameter called damping, $\Delta \boldsymbol{m}$ is a difference between the model in current and previous iteration, and $\mathbf{T}$ is a Toeplitz matrix of the following form:

$$\mathbf{T}(\Delta \boldsymbol{m}) = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \Delta m_1 \\ \Delta m_2 \\ \Delta m_3 \\ \vdots \\ \Delta m_M \end{pmatrix} = \begin{pmatrix} \Delta m_1 - \Delta m_2 \\ \Delta m_2 - \Delta m_3 \\ \Delta m_3 - \Delta m_4 \\ \vdots \\ \Delta m_M \end{pmatrix}. \tag{4}$$

3.2 ~~Exploitation of a priori information~~Tuning initial parameters through synthetic tests

As the problem of calculating RF is highly non-linear (e.g. Bodin et al., 2012) , synthetic tests of RF inversion are generally not fully plausible. Even for quite similar structures there may exist a difference among the sets of inversion parameters illuminating them optimally.

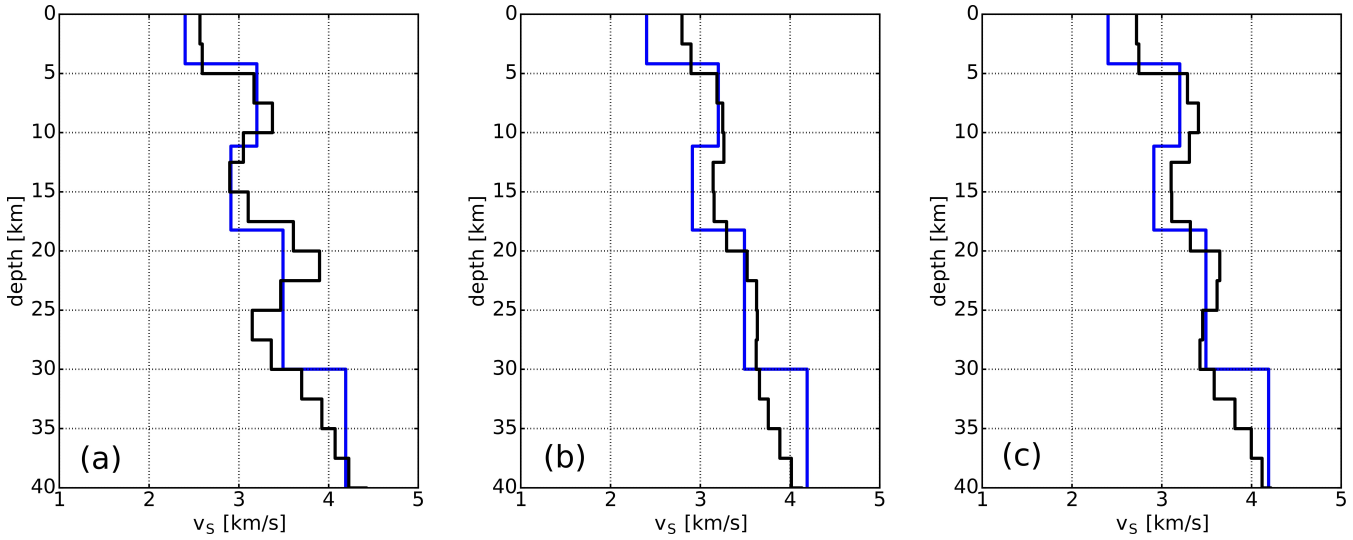

Figure 5. ~~A priori information exploited in the inversion: (~~Results of a ~~) crustal structure beneath all stations~~ joint inversion of ~~"13 BB Star" array computed~~ synthetic data calculated from ~~3D~~ a true model ~~(Grad et al., 2016)~~ marked in blue, θ^2 equal to (a) 0.1; (b) ~~ensemble of~~ 10.0; (c) 1.0; $p = 0.1$; 20 iterations; 10 homogeneous starting models ~~with crust from central station (A0),~~ equally distributed between 2 and ~~different homogeneous mantle velocities within physically plausible 4-~~5 km/s~~S-velocity~~ , each 100 km deep; RF window: -3 to 6.5 s after direct P-wave; SWD period range: $20 - 40$ s.

~~The "13 BB Star" experiment was conducted in the area covered by high-resolution 3D model of the crust (Grad et al., 2016) . To use this prior knowledge~~

Here synthetic analysis was performed in two stages. The first incorporated a simple model (blue line in Fig. 5) and served to get the rough idea of the values of the parameters suitable for the considered data types. An example of finding the reasonable value of the damping parameter θ^2 is shown in Fig. 5. Too low ($\theta^2 = 0.1$) value overshoots the true structure (Fig. ~~??~~a)~~in analogy to to Bayesian inference (Sambridge and Mosegaard, 2002) , we include it in all starting models~~ 5a), too high ($\theta^2 = 10.0$) makes the model too smooth (Fig. ~~??~~5b). ~~We assign alower weight to ("freeze")those layers, preventing the inversion from modifying them too much. We test two options of freezing the crust: the first (referred to as "option 1") involves the layers characterized by velocities and thicknesses exactly from 3D model(Grad et al., 2016) . In the second case ("option 2")we consider thinner layers: 0.1-0.2 and 2 km for upper-middle and lower crust~~ The value $\theta^2 = 1.0$ regarded here as optimal (Fig. 5c) has worse resolving power than the first case (Fig. 5a), but doesn't create any artificial structures which had a higher priority for us.

In a similar way we estimated the value of the influence parameter ($p \approx 0.1$), as well as a suitable number of iterations (after twenty no significant rise of the fit was observed). Note that higher values of p (favouring the dispersion data more) might smoothen the results similarly to the increase of θ^2. Therefore we checked the optimal value of θ^2 through the inversions of a single data type as well.

All these parameters can be assumed independent of the model size and data length, however we revisited their values during the resolution tests for real-size models and data sets – the second stage of the synthetic analysis described in the next section. Testing initial guesses on inversion parameters on full-size structures may turn out too time-consuming for deep, many-layer models and larger number of inversion parameters, therefore we advocate splitting synthetic analysis into two steps.

3.3 Model depth

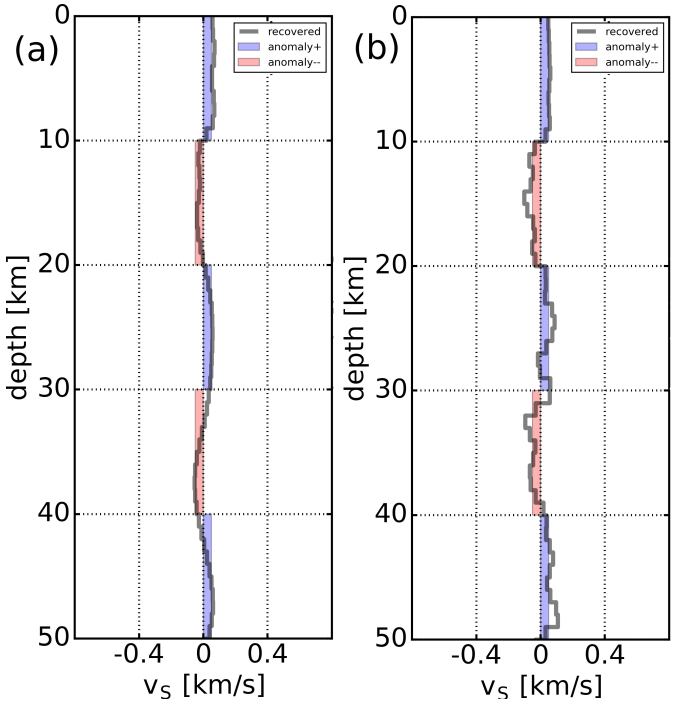

Figure 6. Checkerboard study showing the importance of choosing the correct model depth; synthetic data (RF) to invert was generated from homogeneous background model ($v_S = 3.5$ km/s) with checkerboard (anomaly size: 10 km size, amplitude 5 % of the background model) added across the full depth range of the background model: (a) 50 km, (b) 500 km; depth and velocity of the starting models: 50 km and 3.5 km/s in both cases; RF time window: 6.5 s (a), and 30 s (b).

The range of the inverted data determines the minimum depth of the model. The rationale for this is that medium below the last layer of the starting model is usually assumed to be a homogeneous half-space. As a consequence, too shallow models may introduce some artefacts likely to appear e.g. while trying to fit the RF tail that represents structure below the bottom of

the model. Such distortion can be observed in Fig. 6b where it is juxtaposed with the results for a correct model depth (Fig. 6a). In both cases a model of 50 km depth was inverted for, however a different model to generate synthetic data (RF) was used: 50 km and 500 km deep for Fig. 6a and 6b respectively. The ~~velocity-depth dependence remains the same as in option 1. Densities are calculated with combined formulas from (Berteussen, 1977) and (Gardner et al., 1974) . We assume 1.8-1.9 km /s in sediments, 1.67, 1.73, 1.77 km /s in upper, middle, lower crust respectively, i. e. values similar to those reported by Środa et al. (1999) . All mantle~~ corresponding time window of RF used in synthetic calculation and subsequent inversion was: 6.5 s (6a) and 30 s (6b). For the description of the checkerboard test used to obtain this figure see Sect. 4.

For dispersion curves the reasoning behind the choice of the model depth with respect to the maximum period is slightly different due to the substantial widths of their sensitivity kernels for longer periods (e.g. Romanowicz, 2002) . However, it leads to similar conclusions – too shallow models may be affected by sampling the structure below the model depth. Using models considerably deeper than the part constrained by the data seems to avoid these problems. Models 100 km deep were used in this study to invert SWD curve in $15-30$ s period range, and RF of time window from -3 to 6.5 s after direct P-wave~~velocities are calculated assuming $v_P/v_S = 1.73$. The thickness of mantle layers is 5 km in every model.~~ . We extended the model's depth to 900 km when using $15-180$ s period range of SWD.

3.4 Relative weighting of data sets of different uncertainty

It may seem that the influence parameter p (Julià et al., 2000, 2003) in Eq. 2 is solely responsible for the relative weighting of different data sets with $p = 0.5$ implying their equal contribution to the solution (Horspool et al., 2006) . However, even more important weighting factor is represented by a squared uncertainty of a given data point (Eq. 2). Stringent assessment of seismic data errors is in general highly challenging (Julià et al., 2000, 2003; Bodin et al., 2012) , which makes the problem of finding the optimum value of p essential. Furthermore, in the CPS package only the SWD uncertainty is allowed to be predefined by the user (we used the values of 1σ presented in Fig. 4b). The uncertainty of each RF sample is defined as a standard deviation of the misfit of the full time series, and thus it changes from iteration to iteration. As a result, SWD data may be unintentionally favoured at the beginning of the inversion. Overall, the proper weighting may require value of p other than 0.5, and this can be found only through synthetic tests.

3.5 Ensemble inference

Among downsides of the ~~linearised~~ linearized approach, trapping by local minima is listed as one of the most profound (Bodin et al., 2012), nonetheless Monte Carlo methods are not completely free of it either (Minato et al., 2008; Wathelet, 2008). There were several solutions to this issue ~~proposed~~suggested, e.g. ~~in FWI~~for a full waveform inversion (FWI) Bharadwaj et al. (2016) proposed a functional that pulls the model out of the local minimum~~(Bharadwaj et al., 2016)~~ . As the RF and SWD inversion is not as computationally expensive as FWI ~~(Virieux and Operto, 2009; Warner et al., 2013)~~ (Warner et al., 2013) , we take a different approach based on an ensemble of starting models, which densely ~~cover~~ covers the space of physically plausible solutions. In our case it ~~consists~~ consisted of 20 homogeneous mantle structures evenly distributed between 4 and 5 km/s~~(Fig. ??b)~~, within which falls the great majority of S-wave velocities reported for cratonic mantle in other studies ~~(Fischer et al., 2010; Jones et al., 2010; Vinn~~

All mantle P-wave velocities were calculated assuming $v_P/v_S = 1.73$. Densities are calculated with combined formulas from Gardner et al. (1974) and Berteussen (1977) . The thickness of mantle layers was 5 km down to 300 km, and 10 km for 300 – 900 km depth for all models. Contrary to the prior distribution used in Bayesian analysis, the solution of ~~linearised~~ linearized inversion can converge to the values beyond this range, which is a slightly weaker initial assumption. Homogeneity ensures that we do not make any guess about the mantle which would drive the inversion towards the most "reasonable" structure. We rather let the data do it for us. Approaches with inhomogeneous models can be found e.g. in (Wilde-Piórko et al., 2005; Graw et al., We emphasize the fact that a high density of distribution of the starting models within given interval (here every 0.05 km/s) enables to cover the whole space of physically plausible solutions.

3.6 ~~Data range~~Exploitation of a priori information

~~We used a 30 s wide RF window (3 s before – 27 s after the direct P-wave) finding longer times not reliable enough. By analysing the synthetic RFs calculated from 2-layer models, we conclude that this would encompass converted phases down to 260~~The "13 BB Star" experiment was conducted in the area covered by high-resolution 3D v_P model of the crust (Grad et al., 2016) . Similarly to the Bayesian inference (Sambridge and Mosegaard, 2002) , we wanted to include this prior knowledge in the starting models, which required a calculation of S-wave velocities. To this end we used v_P/v_S values of 1.8~~-280 km depth. We made the inverted model considerably deeper (300 km) to avoid artefacts likely to appear while trying to fit into RF tail that represents structure below the bottom of the model. For dispersion curves the reasoning behind choice of optimal maximum period T_{max} is slightly different due to substantial widths of their sensitivity kernels for longer periods (e.g., Darbyshire et al., 2013; Romanowicz, 2002; Tsuboi and Saito, 1983) . However, it leads to similar conclusions – too shallow models may be affected by sampling the structure below the model depth.~~

3.7 ~~Relative weighting of data sets of different uncertainty~~

~~It may seem that the influence parameter p (Julià et al., 2000, 2003) is solely responsible for the relative weighting of different data sets with $p = 0.5$ implying their equal contribution to the solution (Horspool et al., 2006) . However, even more important weighting factor is represented by a squared uncertainty of a given data point (Eq. 2), and~~ 1.9 in sediments and: 1.67, 1.73, 1.77 in upper, middle, lower crust respectively, roughly the same as reported by Środa et al. (1999) . These assumptions along with the fact that the 3D model was derived from the measurements of different sensitivity (mainly P-wave tomography) surely contributed to the large misfit between observed and synthetic RFs computed from this model. As we didn't want this discrepancy to drive the inversion, we performed an additional inversion for the crustal structure only, starting with the values from the 3D model (see Sect. 5). Then, we adopted the crust to the ~~proper weighting may require value of p other than 0.5. Stringent assessment of seismic data errors in generalis regarded as highly challenging (Julià et al., 2000, 2003; Bodin et al., 2012) , which makes the problem of finding the optimum value of p essential. We address it by performing resolution tests for different values of p, which we describe in Sect. 4.~~ full-depth models and assigned lower weight to partly "freeze" it in the actual inversion.

3.7 ~~Tuning the initial parameters through synthetic tests~~

~~As the problem of calculating RF is highly non-linear (Bodin et al., 2012) , synthetic tests of the inversion considered are not fully plausible. Even for quite similar structures there may exist a difference among the sets of inversion parameters illuminating themoptimally. Here synthetic tests served to get~~

4 Resolution analysis

Resolution analysis was the integral part of the inversion parameter tuning in the case of real-size models and full data range. Apart from the resolution matrix analysis we proposed a 1-D checkerboard test to provide a more intituitive display of expected resolution in case of smoothing and different weighting of model layers.

4.1 Resolution matrix

The model resolution matrix **R** is defined by the relation:

$$\boldsymbol{m} = \mathbf{R}\boldsymbol{m}_{true}, \tag{5}$$

and in the case of damping is expressed by:

$$\mathbf{R} = (\mathbf{A}^T\mathbf{C}_e^{-1}\mathbf{A} + \theta^2\mathbf{T}\mathbf{A}^T\mathbf{C}_e^{-1}\mathbf{A}), \tag{6}$$

where $\theta^2\mathbf{T}$ has the same meaning as in Eq. 3, A is a forward operator mapping the model vector into the data space, and C_e is the error covariance matrix (Gubbins, 2004) . In general, the ~~rough idea of the values of parameters suitable for the considered data types. In such a way we estimated the value of damping ($\theta^2 \approx 1.0$) and influence parameter ($p \approx 0.1$), as well as a number of iterations (after 20, no significant increase in fit was observed)~~presence of non-zero off-diagonal elements is caused by the smoothing and the correlation of the data. In Fig. 7 we juxtapose the resolution matrices for the models of different number and weights of crustal layers. The reason for poor resolution of the shallow structure in the model in Fig. 7a is the usage of the coarser layers with lower weights assigned to them. Although Gubbins (2004) argues that **R** matrix should be used wherever possible since it carries the maximum information one can get about the resolution of the model, Fig. 7 shows that it may fail to provide an intuitive picture of what structure can be resolve in the given inversion. Furthermore it involves additional calculations which may in some cases turn out too expensive.

5 ~~Resolution analysis~~

4.1 Checkerboard test for 1D problems

~~We adopted a checkerboard test , (e.g., Janutyte et al., 2015) to the~~ To mitigate some inconvenience inherent in the resolution matrix analysis, we adopted the checkerboard test known from seismic tomography (e.g. Janutyte et al., 2015) to a 1D problem. The algorithm consists of the following steps:

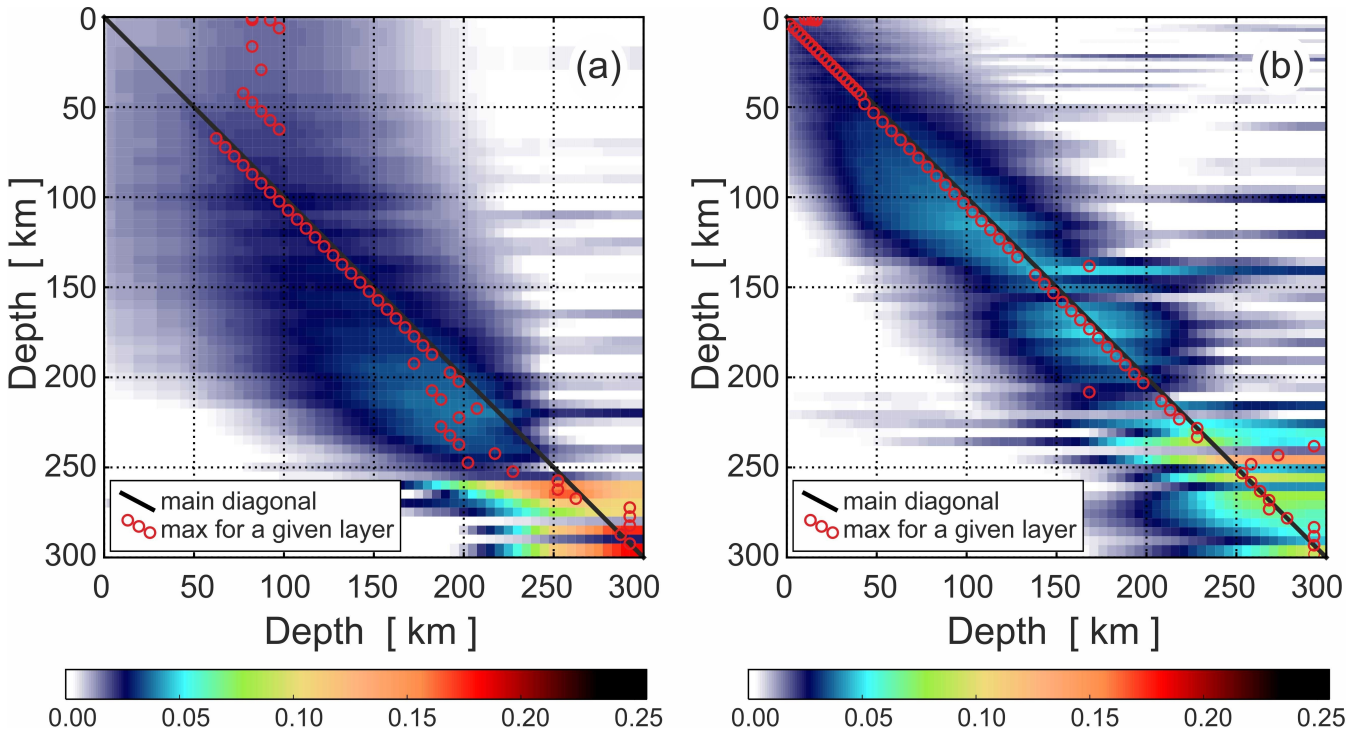

Figure 7. Resolution matrix for a model with: (a) 9 coarse crustal layers, partly frozen (weight 0.5); (b) 31 fine crustal layers, no freezing (weight 1.0); inversion parameters in each case: $\theta^2 = 1.0$, $p = 0.1$, RF window: -3 to 27 s after direct P-wave; SWD period range $15 - 180$ s.

1. Add the pattern of alternating positive/negative anomalies of constant size (thickness) and amplitude (percentage of model velocity at given depth) to the inversion result (arithmetic mean of the ensemble of final models);
2. Compute synthetic data from the model perturbed in above way;
3. Invert the synthetics in the exactly same way as the unperturbed model was obtained;
4. Subtract the original model from the result of previous point to recover the anomaly;
5. Compare the recovered and original anomaly.

A failure to recover the anomalies is indicative of no resolving power of the method regardless of the nature of the wave phenomena behind it.

4.1.1 Size, amplitude and polarity of the anomalies

Despite the simplicity of its idea, checkerboard test may be misleading as has been shown for ~~2D tomography problem (Leveque et al., 1993) , therefore some care has to be taken also in 1D, non-tomographic case.~~ tomography problem (e.g. Leveque et al., 199

Although resolving power of joint inversion of RF and SWD depends on different factors, the same, or even more care should be taken in analyzing its results.

For instance, usually only one of the two possible polarities of anomaly-pattern enables its correct recovery~~(Fig. 7a and 7b). Moreover,~~ , different for different ~~sizes, different polarities are favourable (compare Fig. 7a and 7c). The~~ size of the anomaly (Fig. 8a-c). When using RF data higher amplitudes of anomalies tend to be worse recovered (Fig. ~~7~~8d), which ~~may result from~~ is the effect of the excessive distortion of the synthetic RF time series. The dependence on the polarity has been explained in more detail in Fig. 9. In 9a-b the starting models surround the true model symmetrically (4.53 km/s of the true model falls roughly in the middle of the starting models ensemble), and the recovery is identical for both polarities. For 9c-d in turn, the true model (4.84 km/s) is closer to the right boundary of the starting ensemble (5 km/s), thus the first positive anomaly is worse recovered. As the samples of RF time series are correlated, this affects the stability of the whole solution (visible especially in $150-250$ km depth range). In this way the checkerboard test is able to detect non-optimal choice of the starting models. It also indicates a necessity of weighted (instead of arithmetic) average in obtaining the final model whenever spread of the ensemble results is significant.

4.1.2 Tuning of the inversion parameters

~~Each parameter required in inversion was adjusted~~ A value of each inversion parameter estimated in the first stage of synthetic tests was verified by performing a dedicated checkerboard test. The optimum set has been shown in Fig. ~~7. Figure 8~~8. Even in this case, anomaly of one of the polarities is poorly recovered 8b.

Figure 10 in turn presents the effect of choosing the wrong values, such as too short dispersion curve (Fig. ~~8~~10a), or too low damping (Fig. ~~8~~10b) and influence parameter (Fig. ~~8~~10c). Crust ~~frozen with option 2~~ which is not frozen at all (Fig. ~~8~~10d) exhibits the presence of a fictitious bump and slightly worse resolving power at about 150 km depth.

4.2 ~~Resolution matrix~~

Regardless of the size and amplitude of anomalies used, SWD data (at least for Rayleigh wave fundamental mode) is unable to successfully recover the pattern at greater depths of the upper mantle (Fig. 11). This limitation resulting from the long waves used in the analysis suggests that all the results derived with this data type should be interpreted with special care.

~~The model resolution matrix **R** is defined by the relation:~~

$$m = \mathbf{R}m_{true},$$

~~and in case of damping is expressed by (Gubbins, 2004) :~~

$$\mathbf{R} = (\mathbf{A}^T\mathbf{C}_e^{-1}\mathbf{A} + \mathbf{C}_m^{-1}(\mathbf{A}^T\mathbf{C}_e^{-1}\mathbf{A})).$$

~~The presence of non-zero off-diagonal elements is caused by the smoothing and the correlation of RF data . The poor crustal resolution of model for option 1~~

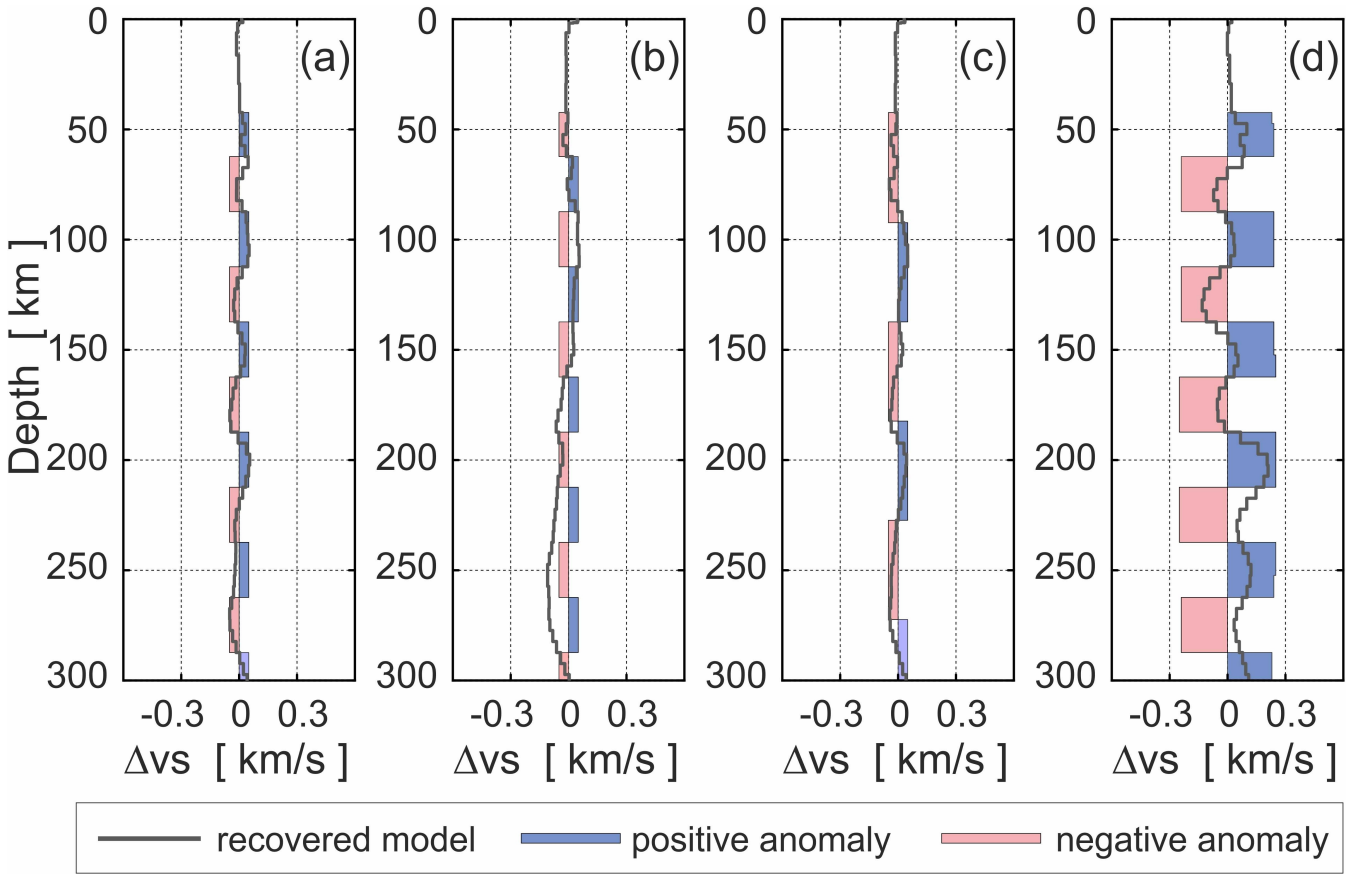

Figure 8. Dependence of the checkerboard test results on a: (a) positive, (b) negative sign (polarity); ~~size:~~ (c) size (40 km); and ~~amplitude:~~ (d) amplitude of the anomaly: 5 % of the ~~anomaly pattern~~background model. ~~Inversion~~ Size of the anomaly in (a), (b), (d): 20 km; amplitude of the anomaly in (a)-(c): 1 % of the background model; inversion parameters in each case: ~~$T_{max} = 180$ s,~~ $\theta^2 = 1.0$~~,~~; $p = 0.1$~~, crust freezing~~; 20 iterations; RF window: ~~option 1.~~ – 3 to 27 s after direct P-wave; SWD period range: 15 – 180 s; crust partly frozen (weight 0.5); average of the models from the Fig. 14b.

5 Results

As described in Sect. 3, a two-stage inversion was used for the observed data in order to take advantage of a priori information in a proper way.

Starting models for the precondition, first-step inversion (Fig. ~~??a) is the effect of freezing. For option 2 there is an apparent dependence of the model at about 140 km on the deeper structure (Fig. ??b). It seems to agree with the results of checkerboard test, which was unable to recover the anomaly at similar depth~~ 12a) consisted of the crust from the 3D model (Grad et al., 2016) and 20 homogeneous mantle structures evenly distributed between 4 and 5 km/s down to about 100 km depth. Since inverting for the crustal structure was not our goal per se, the data misfit was in this case a sufficient criterion of

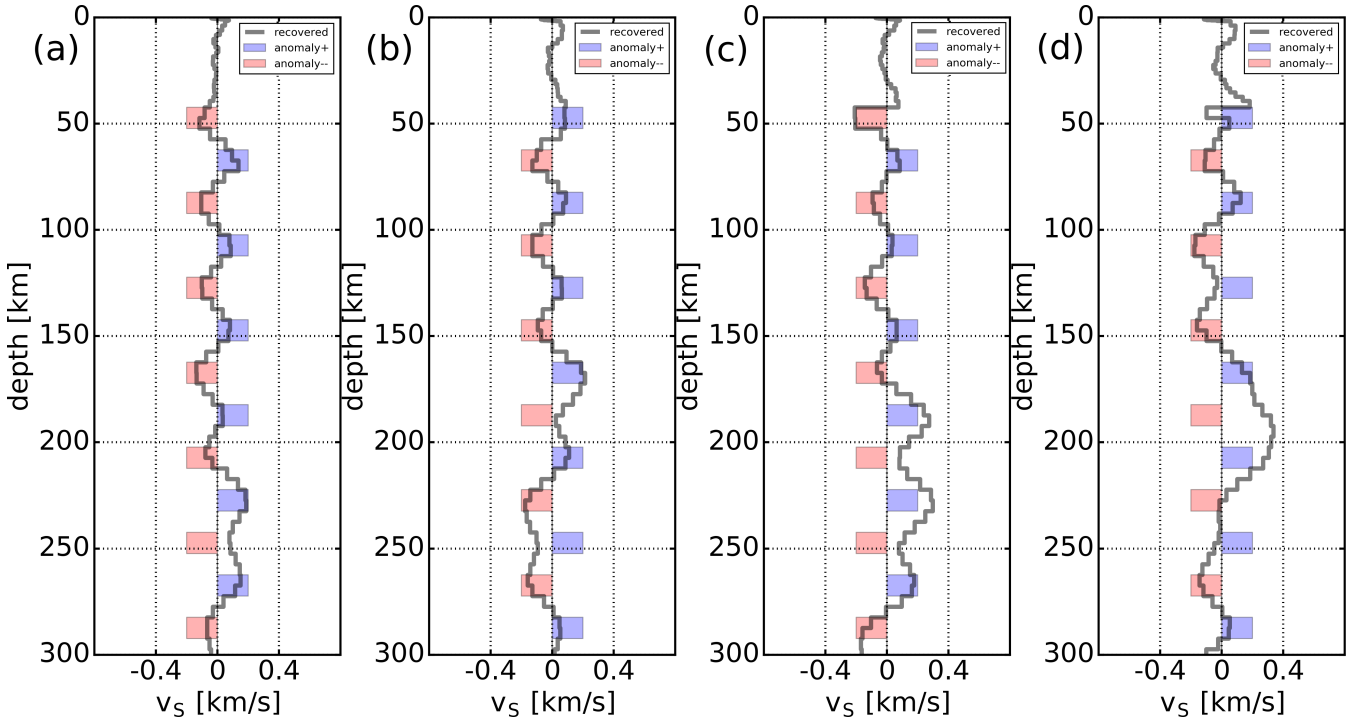

Figure 9. Dependence of the checkerboard test results on the relative position of the background and starting models; in all cases starting models as in Fig. 12a (mantle velocities $4-5$ km/s), but extended to 300 km depth; background model: one of the starting models of the velocity: (a-b) 4.53 km/s (a centre of the ensemble); (c-d) 4.84 km/s; anomaly size: 10 km, amplitude: 4 % of the background model; inversion parameters in each case: $\theta^2 = 0.1$; $p = 0.1$; 10 iterations; RF window: -3 to 30 s after direct P-wave; SWD period range: $15-150$ s; crust partly frozen (weight 0.5); note that in order to highlight the phenomenon we used a lower value of $\theta^2 = 0.1$ along with a sparser checkerboard pattern.

the results quality (the smaller the misfit, the better the model, which is not necessarily true in general). That's why a lower value of damping $\theta^2 = 0.1$, and larger number of iterations (40) was used. The resulting model (Fig. ~~8d) . The final loss of the resolution begins slightly lower~~12b) is characterized by a prominent low velocity zone (LVZ) at the depth of about 25 km, and the additional, thin layer within the sedimentary crust. Two major discontinuities: sedimentary/~~higher than 200 km depth for option 1 and 2 respectively~~crystalline crust and Moho are much smoother compared to the starting models. Worth noticing is a small spread of the final models, especially for the initial diversity of the Moho contrasts. The data fit is satisfactory both for RF (13a) and SWD (13b), so the main goal of the first stage of the inversion workflow has been met.

6 ~~Results~~

~~Results of the ensemble-inversion of: (a) SWD only; (b) RF only; and (c) both data sets. RF fit for all final models in the inversion of (a) both data sets; and (b) RF only. Below we present the results of the inversion (Fig.10) including the data~~

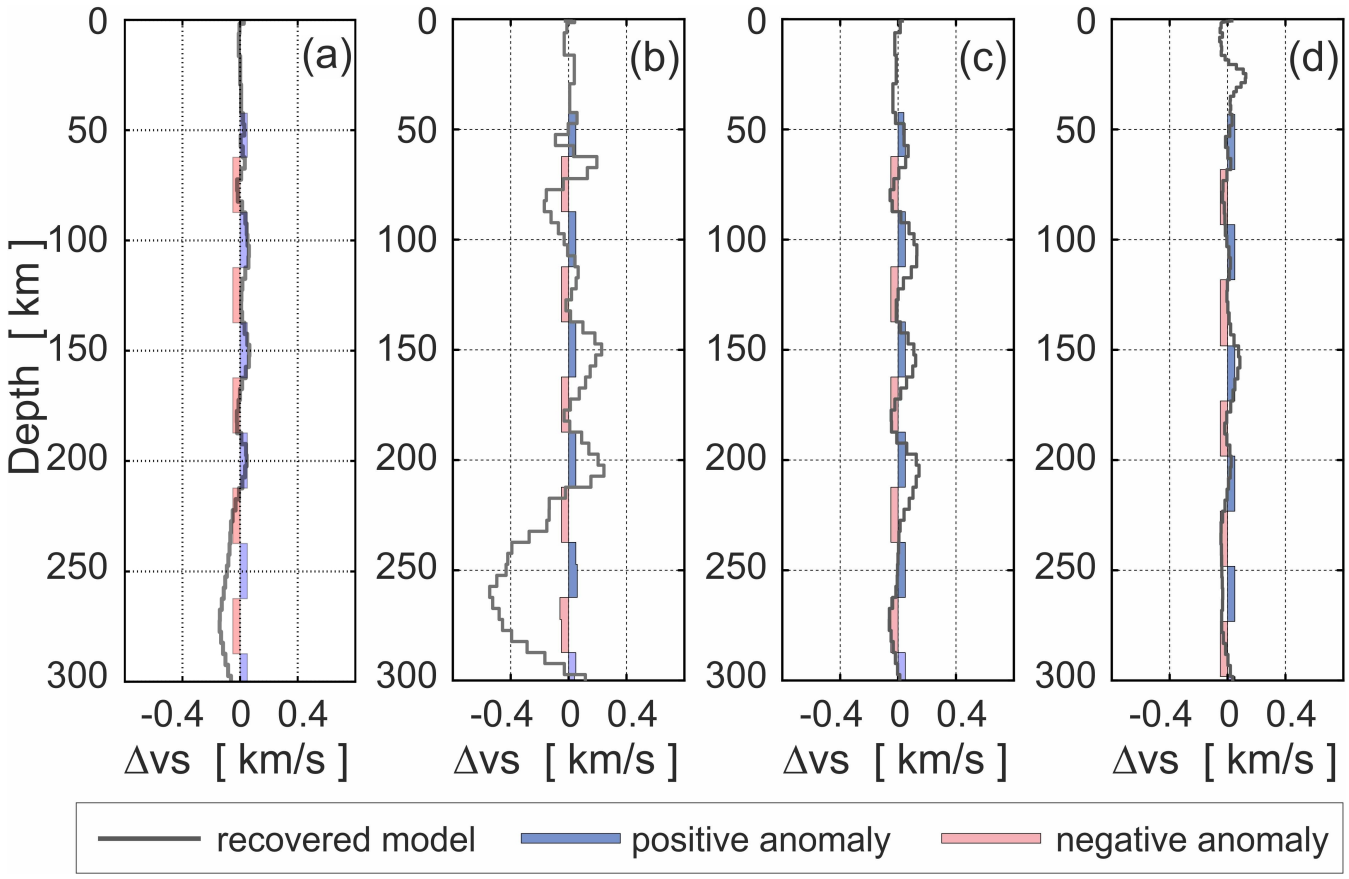

Figure 10. Examples of poorly tuned initial parameters: (a) ~~$T_{max} = 150$~~ maximum period of SWD 100 s; (b) $\theta^2 = 0.1$; (c) $p = 0.01$; (d) no crust freezing ~~: option 2. The~~ (weight 1.0); the remainder of the parameters are the same as in Fig. ~~7~~8.

~~fit~~ After the first step, we incorporated the crust of its final models into the full-size (900 km, i.e. reaching well below the resolving power of both data types) starting models of the second-stage inversion (Fig. ~~11 and 12)for the central station of the "13 BB Star" array (~~14a). The weight 0.5 meant that the crust was partially frozen, allowing only for minor updates. The maximum period of SWD was extended to 180 s, but the time window of RF remained the same. The true amplitudes of reflected phases had been lost due to the move-out correction and stacking, and the time series after direct Moho conversion wouldn't be reliable to invert. Because our RF data set wasn't large enough, we couldn't stack RFs over different slowness ranges and use them as separate inputs in the inversion. For this reason, RF didn't constrained most of the mantle and only smooth mantle structures were expected in the second-stage inversion. The results are presented for A0 ~~). The following values of inversion parameters were used: $p = 0.1$ (for joint inversion), $\theta^2 = 1.0$, $T_{max} = 180$ s, option 1. As shown in the previous section, the interpretation should not be made below 200 km depth. Introducing the SWD data helps to diminish the inherent non-uniqueness of the RF inversion(Fig. 10). The result for option 2 exhibits distinctive crust variations compared to option 1~~

~~Resolution matrix for : (a) option 1; (b) option 2 of crust freezing. Inversion parameters in each case: $T_{max} = 180$ s, $\theta^2 = 1.0$, $p = 0.1$.~~

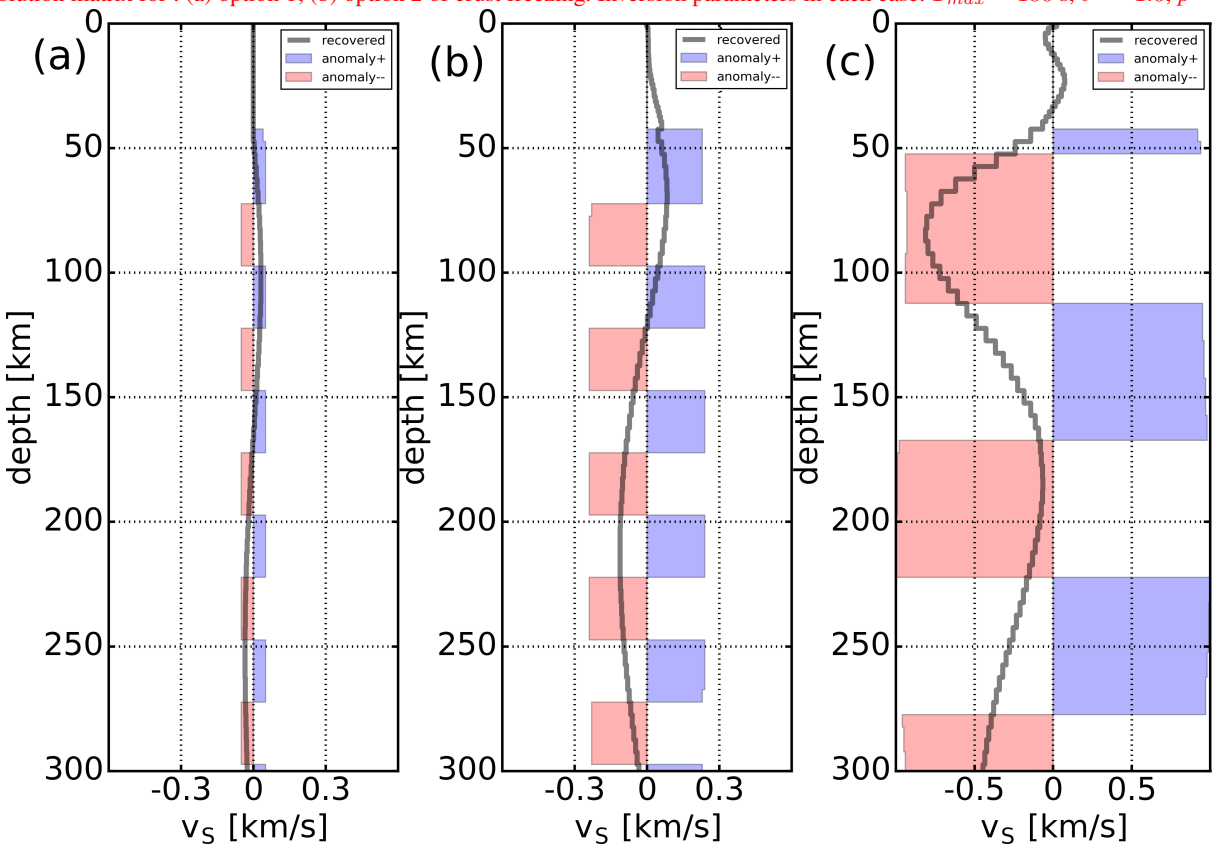

Figure 11. Checkerboard anomalies recovery in the case where most of the mantle is not constrained by the RF. Anomaly: (a) size: 20 km; amplitude 1 % of the background model; (b) size: 20 km; amplitude 5 % of the background model (compare with Fig. 7d); (c) 50 km; amplitude 20 % of the background model; inversion parameters in each case: $\theta^2 = 1.0$; $p = 0.1$; 20 iterations; RF window: -3 to 6.5 s after direct P-wave; SWD period range: $15 - 180$ s; crust partly frozen (weight 0.5); background model: average of the models from the Fig. 14b.

~~(fig. 13a) , which , according to the checkerboard test~~ station's RF only, because the crust wasn't the main target of this study. The resulting models (Fig. 14b) are very close to each other which indicates that we succeeded to mitigate the non-uniqueness of the method.

6 Discussion

To obtain the results presented in Fig. 14b we developed a multi-step workflow for linearized inversion, which aimed at obtaining credible models of S-wave velocity down to 300 km with RF and SWD data. Its key features include: two-stage synthetic analysis of the resolution which e.g. allows a correct tuning of the parameter p weighting the influence of each data type (Fig. ~~8d), may not be fully plausible. Results for all stations are presented for option 1 only~~ 10c); two-stage adaptation

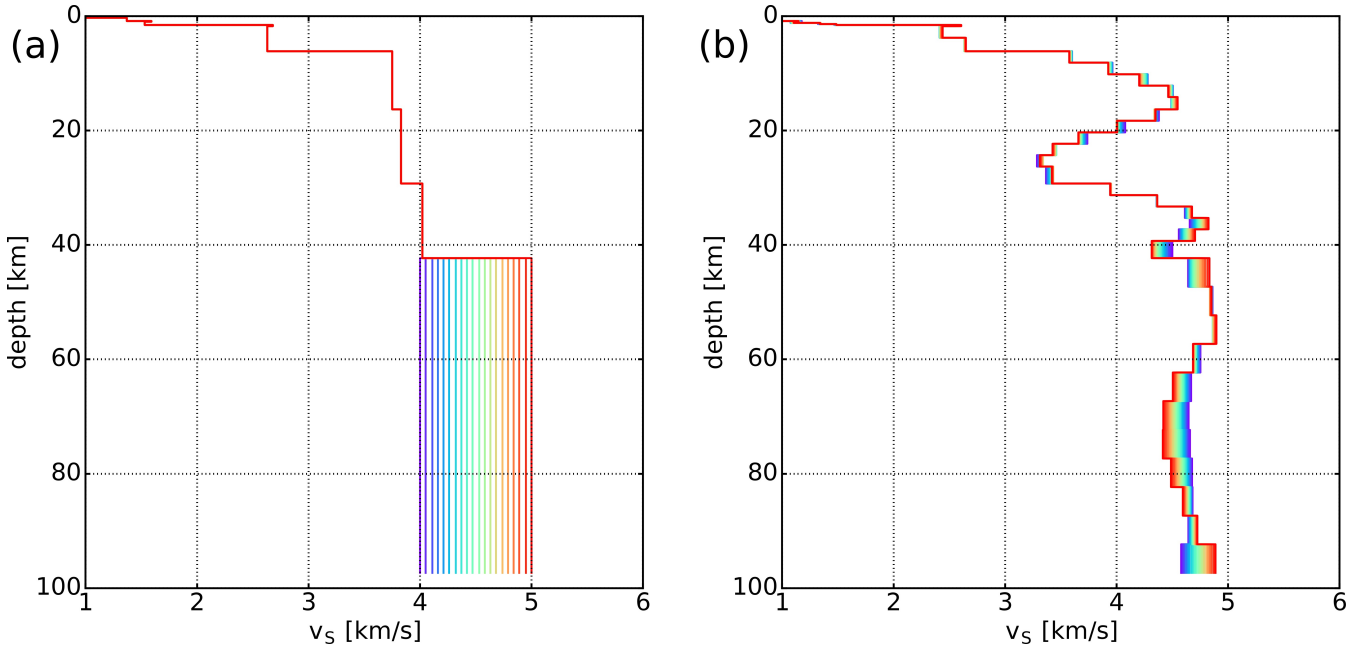

Figure 12. ~~SWD fit for all final models in the~~ First-step inversion – adaptation of ~~:~~ the crust; (a) ~~SWD only~~starting; and (b) ~~both data sets~~final models; RF for the central station (A0) of the "13 BB Star" array; inversion parameters: $\theta^2 = 0.1$, $p = 0.1$, 40 iterations, model depth: 100 km, RF window: -3 to 6.5 s after direct P-wave; SWD period range: $15 - 30$ s. Each final model in (b) corresponds to the starting model in (a) of the same color.

of the a priori knowledge on the crustal part of the model obtained with different methods (Fig. 12 and 14); careful choice of the model's depth relative to the range of observed data in order to prevent synthetic data from sampling the bottom half-space instead of real structure and thus distorting the inversion's results; mitigating the inherent non-uniqueness of the inverse problem by using the ensemble of starting models permitting to find all of its physical solutions (Fig. ~~13b~~14a).

7 ~~Discussion~~

~~The common feature of mantle S-wave velocity models (see Fig. 10 and 13) is lowering velocity in the 80-120 km depth range~~ Following this workflow provided us with the models that exhibit very low spread in the mantle at the depth $70 - 300$ km (Fig. ~~However, we do not interpret this zone as EEC asthenosphere, but rather, because of the neighbouring TESZ, a wedge of younger Paleozoic asthenosphere (Hoernle et al., 1995) . The asthenosphere could be identified with smooth and deeper~~ 14b) and that fit the observed data well, in most cases falling inside the range of one standard deviation of the mean (Fig. 15). The blue models responsible for the RF misfit exceeding the accepted error do not differ from the others below the 70 km depth. The inability of the synthetic curves to fit the notch present in the observed SWD curve between $60 - 90$ s is likely related to the oversimplified forward calculations assuming e.g. isotropy of the model.

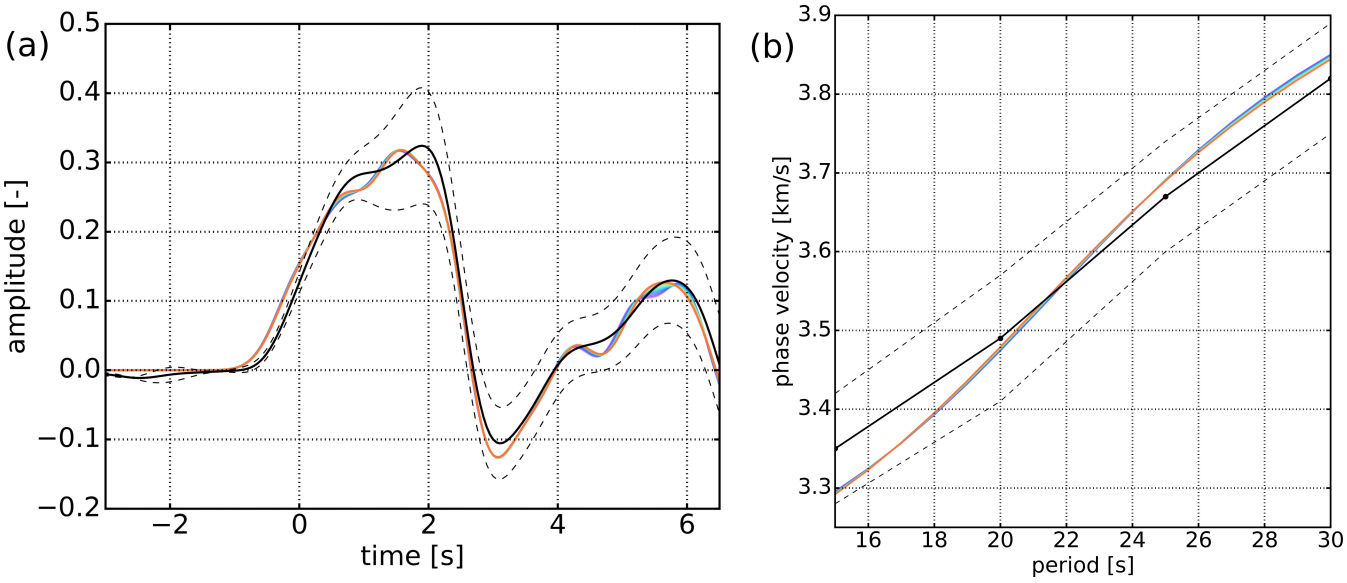

Figure 13. ~~Mean~~ Data fit for all final models ~~from joint~~ in the first-step inversion~~for~~: (a) ~~central~~ RF (A0 station~~, different crust freezing options, with $\pm\sigma$ bounds~~); (b) ~~all stations, crust freezing option 1~~.SWD; black solid lines: observed data; black dashed lines $\pm\sigma$ of observed data; each synthetic data set was computed from a corresponding final model in Fig. 12b of the same color.

One of the main targets of cratonic upper-mantle investigations is the lithosphere-asthenosphere boundary/transition zone (LAB) (Eaton et al., 2009) . Numerous studies have shown that LAB beneath Precambrian cratons is not easily detected by seismic methods, e.g. receiver function, because converted phases from smooth transition zones are not well pronounced (Kind et al., 2012) . Geissler et al. (2010) , using S receiver functions, show that beneath the stations located in the Precambrian platform of Eastern Europe LAB deepens to approximately 200 km. This value is consistent with LAB depths of other old cratons of the Earth (Eaton et al., 2009; Jones et al., 2010) .

S-wave velocity lowering starting ~~from the depth 180-200 km . Similar depth of the~~ at about 180 – 200 km depth visible in Fig. 14b might be identified as a top of lithosphere-asthenosphere transition/boundary (LAB). Similar depth of the LAB has been obtained along profile P4 located about 200 km south-east of "13 BB Star" array. It was derived from a study of relative P-wave residuals by Świeczak (2007), using the method of Babuška and Plomerová (1992). The average depth of the LAB for the EEC was estimated at 193 km, significantly in contrast to TESZ, where it lies at 119 km depth (Wilde-Piórko et al., 2010). ~~Numerous seismic studies have shown that the LAB beneath Precambrian cratons are not easily detected by seismic methods, e. g. receiver function, because converted phases from smooth transition zones are not well pronounced (Kind et al., 2012) . Geissler et al. (2010) , using S receiver functions, show that beneath the stations located in the Precambrian platform of Eastern Europe LAB deepens to approximately 200 km. This value is consistent with LAB depths of other old cratons of the Earth (Eaton et al., 2009; Jones et al., 2010) .~~ However, cutting our RF just after the direct Moho conversions, caused the deeper structure to be constrained almost exclusively by the SWD data. As shown with the checkerboard tests (Fig. 11), the estimation

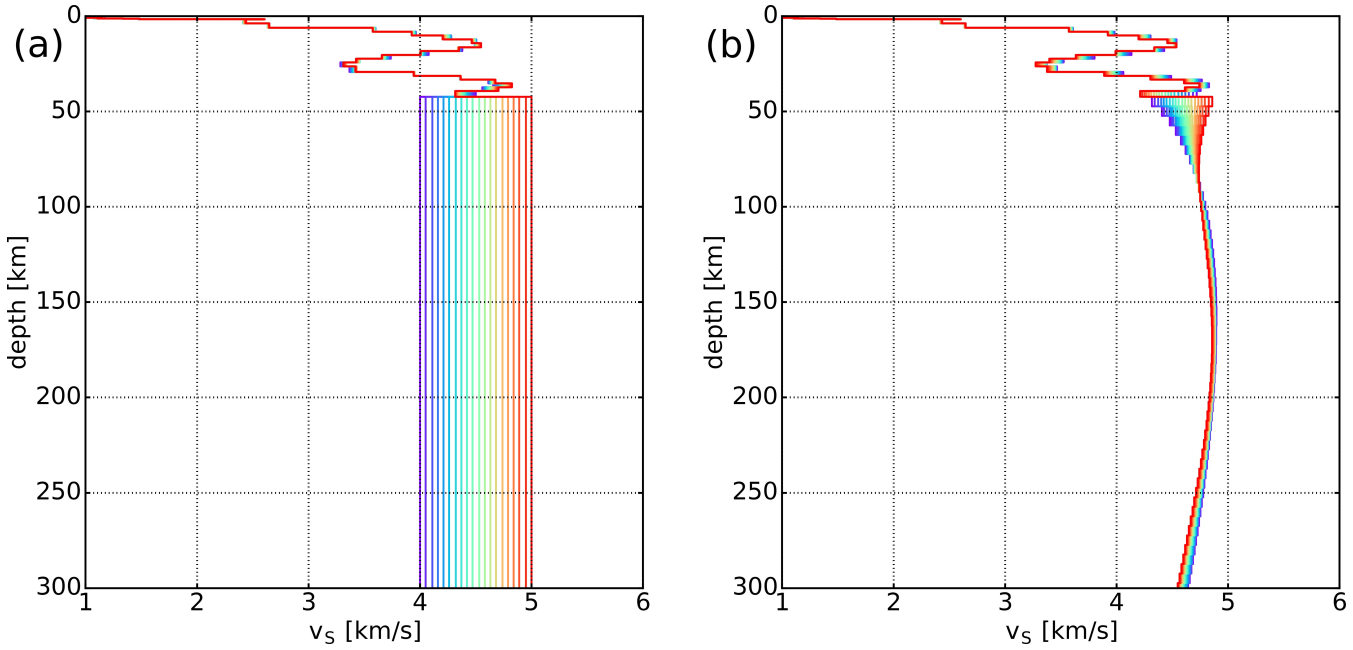

Figure 14. Second-step inversion – full depth; (a) starting; and (b) final models; RF for the central station (A0) of the "13 BB Star" array; inversion parameters: $\theta^2 = 1.0$, $p = 0.1$, 20 iterations, model depth: 900 km, RF window: -3 to 6.5 s after direct P-wave; SWD period range: $15 - 180$ s. Each final model in (b) corresponds to the starting model in (a) of the same color.

based on SWD data only, like in this case, is not fully credible in terms of assessing depth and character of the transition. Consequently, we don't draw any conclusions about them. Presumably the only parameter not affected by this limitation is a maximum velocity in the considered depth range. The value $v_S \approx 4.85$ found at about 180 km depth (Fig. 14b) is quite high even for mantle velocities beneath old Precambrian cratons (Fischer et al., 2010; Vinnik et al., 2015; Meier et al., 2016) . Results reported Meier et al. (2016) concern a different part of the EEC, and were obtained from much larger data sets. The big difference between the dispersion curves used there compared to this study may results from the automated procedure of selecting curves, which might have rejected the data illuminating local structure in finer scale. For a detailed discussion of the geophysical properties of the lithosphere in this area, see (Grad et al., 2017) .

7 Conclusions

~~Using the data from recent "13 BB Star" experiment, we studied the limitations of a linearised~~ We investigated the caveats of a linearized joint inversion of receiver function and surface-wave dispersion ~~in terms of its ability to resolve upper mantle structure~~curves. We argue for ~~the utility of linearised inversion whenever it is handled properly. Taking into account the ensemble of homogeneous starting models and having carefully studied~~ its utility when handled properly, e.g. studing the

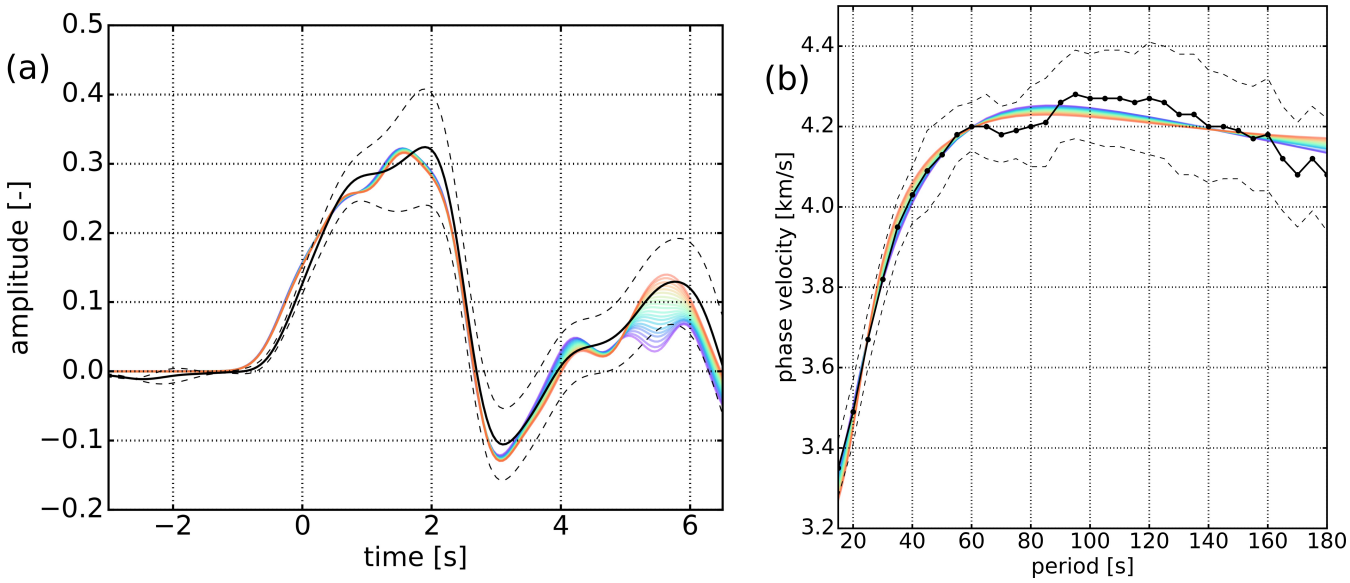

Figure 15. Data fit for all final models in the second-step inversion: (a) RF (A0 station); (b) SWD; black solid lines: observed data; black dashed lines $\pm\sigma$ of observed data; each synthetic data set was computed from a corresponding final model in Fig. 14b of the same color.

influence of inversion parameters on ~~resolution of the final models, this type of modelling can give insights on the credibility of the solution, including approximate estimation of uncertainty~~the resolution of final models.

We developed a full workflow, whose main steps are summarized below. It is versatile and can be applied to both linearized and Monte Carlo inversion scheme, as well as to other data types.

1. Choice of the reliable data range;
2. Determination of the model depth so that it distinctively exceeds the scope of a non-zero sensitivity of each data type;
3. Finding appropriate values of the inversion parameters, in particular damping and relative weights between different data types through the synthetic tests using simple, small-size, and thus computationally cheap model;
4. Verification of the adequacy of the chosen parameters for full-size models approximating the expected structure via 1D checkerboard tests and resolution matrix analysis (if possible). We advocate that ~~a resolution analysis, e.g. a novel 1-D checkerboard test proposed here,~~ it ought to be an integral part of the parameter-tuning workflow, not ~~the afterthought. This approach is versatile and can be applied to both linearised and Monte Carlo inversion scheme. Here, it~~ an afterthought.
5. First-stage of the inversion for part of the model (e.g. crust) recognized a priori from other studies, regarding the minimization of the misfit as the only criterion of the result quality;
6. Incorporating the result in the second-stage inversion using a full range of the data and full depth of the model;

For every inversion we advocate using an ensemble of homogeneous starting models covering the entire space of physically plausible solution to address the non-uniqueness of the problem without driving the inversion towards any arbitrary minimum.

Taking this approach allowed us to obtain ~~reliable~~ a credible S-wave velocity ~~models down to 200 km depth~~ model of Earth's structure beneath the "13 BB Star" ~~array. A feature that all final models have in common is a presumable beginning of low-velocity zone at 180-200 km depth, which~~ network down to 300 km depth. Its crustal part adds new information on the number and a character of discontinuities below the central station of the array. The mantle structure exhibits similar, but rather high v_S values compared to those reported for other cratons. Smooth LVZ starting at about $180 - 200$ km depth. may be interpreted as ~~an upper part of the cratonic asthenosphere~~the lithosphere-asthenosphere boundary.

8 Code availability

For bash scripts adapting CPS package (Herrmann, 2013) ~~for~~ to the ensemble inference and checkerboard tests, as well as ~~small Python tools~~ Python libraries and programs for data preparation and plotting, please contact the corresponding author.

Acknowledgements. We are thankful to R. B. Herrmann for making his CPS package (Herrmann, 2013) freely available. Some figures in Sect. 2 were prepared using GMT package (Wessel et al., 2013). Calculation of RF was performed by Seismic Handler package (Stammler, 1993). National Science Centre Poland provided financial support for this work by NCN grant DEC-2011/02/A/ST10/00284. We thank two anonymous reviewers for numerous insightful remarks that helped to improve this manuscript.

References

Ammon, C. J., Randall, G. E., and Zandt, G.: On the Nonuniqueness of Receiver Function Inversions, Journal of Geophysical Research, 95, 15,303–15,318, 1990.

Babuška, V. and Plomerová, J.: The lithosphere in central Europe-seismological and petrological aspects, Tectonophysics, 207, 141–163, 1992.

Bao, X., Sun, X., Xu, M., Eaton, D. W., Song, X., Wang, L., Ding, Z., Mi, N., Li, H., Yu, D., Huang, Z., and Wang, P.: Two crustal low-velocity channels beneath SE Tibet revealed by joint inversion of Rayleigh wave dispersion and receiver functions, Earth and Planetary Science Letters, 415, 16–24, 2015.

Berteussen, K.-A.: Moho depth determinations based on spectral-ratio analysis of NORSAR long-period P waves, Physics of the Earth and Planetary Interiors, 15, 13–27, 1977.

Bharadwaj, P., Mulder, W., and Drijkoningen, G.: Full waveform inversion with an auxiliary bump functional, Geophysical Journal International, 206, 1076–1092, 2016.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., and Rawlinson, N.: Transdimensional inversion of receiver functions and surface wave dispersion, Journal of Geophysical Research: Solid Earth, 117, 1–24, 2012.

Bodin, T., Yuan, H., and Romanowicz, B.: Inversion of receiver functions without deconvolution–application to the Indian craton, Geophysical Journal International, 196, 1025–1033, 2014.

Darbyshire, F. A., Eaton, D. W., and Bastow, I. D.: Seismic imaging of the lithosphere beneath Hudson Bay: Episodic growth of the Laurentian mantle keel, Earth and Planetary Science Letters, 373, 179–193, 2013.

Deng, Y., Shen, W., Xu, T., and Ritzwoller, M. H.: Crustal layering in northeastern Tibet: A case study based on joint inversion of receiver functions and surface wave dispersion, Geophysical Journal International, 203, 692–706, 2015.

Du, Z. J. and Foulger, G. R.: The crustal structure beneath the northwest fjords, Iceland, from receiver functions and surface waves, Geophysical Journal International, 139, 419–432, 1999.

Eaton, D. W., Darbyshire, F., Evans, R. L., Gr??tter, H., Jones, A. G., and Yuan, X.: The elusive lithosphere-asthenosphere boundary (LAB) beneath cratons, Lithos, 109, 1–22, 2009.

Fischer, K. M., Ford, H. A., Abt, D. L., and Rychert, C. A.: The Lithosphere-Asthenosphere Boundary, Annual Review of Earth and Planetary Sciences, 38, 551–575, 2010.

Fontaine, F. R., Barruol, G., Tkalčić, H., Wölbern, I., Rümpker, G., Bodin, T., and Haugmard, M.: Crustal and uppermost mantle structure variation beneath La Reunion hotspot track, Geophysical Journal International, 203, 107–126, 2015.

Gardner, G. H. F., Gardner, L. W., and Gregory, A. R.: Formation velocity and density; the diagnostic basics for stratigraphic traps, Geophysics, 39, 1974.

Geissler, W. H., Sodoudi, F., and Kind, R.: Thickness of the central and eastern European lithosphere as seen by S receiver functions, Geophysical Journal International, 181, 604–634, 2010.

Grad, M., Tiira, T., Behm, M., Belinsky, A. A., Booth, D. C., Brückl, E., Cassinis, R., Chadwick, R. A., Czuba, W., Egorkin, A. V., England, R. W., Erinchek, Y. M., Fougler, G. R., Gaczyński, E., Gosar, A., Grad, M., Guterch, A., Hegedüs, E., Hrubcová, P., Janik, T., Jokat, W., Karagianni, E. E., Keller, G. R., Kelly, A., Komminaho, K., Korja, T., Kortström, J., Kostyuchenko, S. L., Kozlovskaya, E., Laske, G., Lenkey, L., Luosto, U., Maguire, P. K. H., Majdański, M., Malinowski, M., Marone, F., Mechie, J., Milshtein, E. D., Motuza, G., Nikolova, S., Olsson, S., Pasyanos, M., Petrov, O. V., Rakitov, V. E., Raykova, R., Ritzmann, O., Roberts, R., Sachpazi, M., Sanina, I. A.,

Schmidt-Aursch, M. C., Serrano, I., Špičák, A., Środa, P., Šumanovac, F., Taylor, B., Tiira, T., Vedrentsev, A. G., Vozár, J., Weber, Z., Wilde-Piórko, M., Yegorova, T. P., Yliniemi, J., Zelt, B., and Zolotov, E. E.: The Moho depth map of the European Plate, Geophysical Journal International, 176, 279–292, 2009.

Grad, M., Polkowski, M., Wilde-Piórko, M., Suchcicki, J., and Arant, T.: Passive Seismic Experiment "13 BB Star" in the Margin of the East European Craton, Northern Poland, Acta Geophysica, 63, 352–373, 2015.

Grad, M., Polkowski, M., and Ostaficzuk, S. R.: High-resolution 3D seismic model of the crustal and uppermost mantle structure in Poland, Tectonophysics, 666, 188–210, 2016.

Grad, M., Puziewicz, J., Majorowicz, J., Chrapkiewicz, K., Lepore, S., Polkowski, M., and Wilde-Piórko, M.: Geophysical characteristic of the lower lithosphere and asthenosphere in the marginal zone of the East European Craton, Submitted, pp. 58–73, 2017.

Graw, J. H., Hansen, S. E., Langston, C. A., Young, B. A., Mostafanejad, A., and Park, Y.: An assessment of crustal and upper-mantle velocity structure by removing the effect of an ice layer on the P-wave response: An application to antarctic seismic studies, Bulletin of the Seismological Society of America, 107, 639–651, 2017.

Green, P. J. and Hastie, D. I.: Reversible jump MCMC, Genetics, 155, 1391–1403, 2009.

Gubbins, D.: Time Series Analysis and Inverse Theory for Geophysicists, Cambridge University Press, 2004.

Herrmann, R. B.: Computer programs in seismology: An evolving tool for instruction and research, Seismological Research Letters, 84, 1081–1088, 2013.

Hoernle, K., Zhang, Y.-S., and Graham, D.: Seismic and geochemical evidence for large-scale mantle upwelling beneath the eastern Atlantic and western and central Europe, Nature, 374, 34–39, 1995.

Horspool, N. A., Savage, M. K., and Bannister, S.: Implications for intraplate volcanism and back-arc deformation in northwestern New Zealand, from joint inversion of receiver functions and surface waves, Geophysical Journal International, 166, 1466–1483, 2006.

Janutyte, I., Majdanski, M., Voss, P. H., Kozlovskaya, E., Wilde-Piórko, M., Geissler, W. H., Plomerova, J., Grad, M., Babuška, V., Bruckl, E., Cyziene, J., Czuba, W., England, R., Gaczyński, E., Gazdova, R., Gregersen, S., Guterch, A., Hanka, W., Hegedus, E., Heuer, B., Jedlička, P., Lazauskiene, J., Keller, G. R., Kind, R., Klinge, K., Kolinsky, P., Komminaho, K., Kruger, F., Larsen, T., Majdański, M., Malek, J., Motuza, G., Novotny, O., Pietrasiak, R., Plenefisch, T., Růžek, B., Sliaupa, S., Środa, P., Świeczak, M., Tiira, T., Voss, P., and Wiejacz, P.: Upper mantle structure around the Trans-European Suture Zone obtained by teleseismic tomography, Solid Earth, 6, 73–91, 2015.

Jin, G. and Gaherty, J. B.: Surface wave phase-velocity tomography based on multichannel cross-correlation, Geophysical Journal International, 201, 1383–1398, 2015.

Jones, A. G., Plomerova, J., Korja, T., Sodoudi, F., and Spakman, W.: Europe from the bottom up: A statistical examination of the central and northern European lithosphere–asthenosphere boundary from comparing seismological and electromagnetic observations, Lithos, 120, 14–29, 2010.

Julià, J., Ammon, C. J., Herrmann, R. B., and Correig, A. M.: Joint inversion of receiver function and surface wave dispersion observations, Geophysical Journal International, 143, 99–112, 2000.

Julià, J., Ammon, C. J., and Herrmann, R. B.: Lithospheric structure of the Arabian Shield from the joint inversion of receiver functions and surface-wave group velocities, Tectonophysics, 371, 1–21, 2003.

Kennett, B. L. N. and Engdahl, E. R.: Traveltimes for global earthquake location and phase identification, Geophysical Journal International, 105, 429–465, 1991.

Kind, R., Yuan, X., and Kumar, P.: Seismic receiver functions and the lithosphere–asthenosphere boundary, Tectonophysics, 536, 25–43, 2012.

Kind, R., Yuan, X., Mechie, J., and Sodoudi, F.: Structure of the upper mantle in the north-western and central United States from USArray S-receiver functions, Solid Earth, 6, 957–970, 2015.

Langston, C. A.: The effect of planar dipping structure on source and receiver responses for constant ray parameter, Bulletin of the Seismological Society of America, 67, 1029–1050, 1977.

Leveque, J.-J., Rivera, L., and Wittlinger, G.: On the use of the checker-board test to assess the resolution of tomographic inversions, Geophysical Journal International, 115, 313–318, 1993.

Li, M., Zhang, S., Wang, F., Wu, T., and Qin, W.: Crustal and upper-mantle structure of the southeastern Tibetan Plateau from joint analysis of surface wave dispersion and receiver functions, Journal of Asian Earth Sciences, 117, 52–63, 2016.

Mahalanobis, P. C.: On the generalised distance in statistics, Proc. Natl. Acad. Sci. India, 12, 49–55, 1936.

Malinverno, A.: Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem, Geophysical Journal International, 151, 675–688, 2002.

Meier, T., Soomro, R., Viereck, L., Lebedev, S., Behrmann, J., Weidle, C., Cristiano, L., and Hanemann, R.: Mesozoic and Cenozoic evolution of the Central European lithosphere, Tectonophysics, 692, 58–73, 2016.

Minato, S., Tsuji, T., Matsuoka, T., Nishizaka, N., and Ikeda, M.: Global optimisation by simulated annealing for common reflection surface stacking and its application to low-fold marine data in southwest Japan, 2008.

Özalaybey, S., Savage, M. K., Sheehan, A. F., Louie, J. N., and Brune, J. N.: Shear-Wave Velocity Structure in the Northern Basin and Range Province from the Combined Analysis of Receiver Functions and Surface Waves, Bulletin of the Seismological Society of America, 87, 183–199, 1997.

Pharaoh, T. C.: Palaeozoic terranes and their lithospheric boundaries within the Trans-European Suture Zone (TESZ): A review, Tectonophysics, 314, 17–41, 1999.

Polkowski, M. and Grad, M.: Seismic Wave Velocities in Deep Sediments in Poland: Borehole and Refraction Data Compilation, Acta Geophysica, 63, 698–714, 2015.

Roberts, G. O., Gelman, A., and Gilks, W. R.: Efficient Metropolis Jumping Rules, Bayesian Statistics, 5, 599–607, 1996.

Roberts, G. O., Gelman, A., and Gilks, W. R.: Weak convergence and optimal scaling of random walk Metropolis algorithms, Annals of Applied Probability, 7, 110–120, 1997.

Romanowicz, B.: Inversion of surface waves: A review, International Handbook of earthquake and engineering seismology, 81, 149–173, 2002.

Ryka, W.: Prekambryjska ewolucja platformy wschodnioeuropejskiej w Polsce, Kwartalnik Geologiczny, 26, 257–272, 1982.

Sambridge, M.: Geophysical inversion with a neighbourhood algorithm - II. Appraising the ensemble, Geophysical Journal International, 138, 727–746, 1999.

Sambridge, M. and Mosegaard, K.: Monte carlo methods in geophysical inverse problems, Reviews of Geophysics, 40, 2002.

Shen, W., Ritzwoller, M. H., Schulte-Pelkum, V., and Lin, F.-c.: Joint inversion of surface wave dispersion and receiver functions: a Bayesian Monte-Carlo approach, Geophysical Journal International, pp. 807–836, 2013.

Sosa, A., Thompson, L., Velasco, A. A., Romero, R., and Herrmann, R. B.: 3-D structure of the Rio Grande Rift from 1-D constrained joint inversion of receiver functions and surface wave dispersion, Earth and Planetary Science Letters, 402, 127–137, 2014.

Środa, P., Czuba, W., Guterch, A., Grad, M., Thybo, H., Keller, G. R., Miller, K. C., Tiira, T., Luosto, U., Yliniemi, J., Motuza, G., and Nasedkin, V.: P- and S-wave velocity model of the southwestern margin of the Precambrian East European Craton; POLONAISE'97, profile P3, Tectonophysics, 314, 175–192, 1999.

Stammler, K.: SeismicHandler – programmable multichannel data handler for interactive and automatic processing of seismological analyses, Computers and Geosciences, 19.2, 135–140, 1993.

Świeczak, M.: System litosfera-astenosfera w strefie TESZ w Polsce na podstawie modelowań sejsmicznych i grawimetrycznych, Ph.D. thesis, 2007.

Tsuboi, S. and Saito, M.: Partial derivatives of Rayleigh wave particle motion, Journal of Physical Earth, 31, 103–113, 1983.

Vinnik, L., Kozlovskaya, E., Oreshin, S., Kosarev, G., Piiponen, K., and Silvennoinen, H.: The lithosphere, LAB, LVZ and Lehmann discontinuity under central Fennoscandia from receiver functions, Tectonophysics, 667, 189–198, 2015.

Vinnik, L. P.: Detection of waves converted from P to SV in the mantle, Physics of the Earth and Planetary Interiors, 15, 39–45, 1977.

Virieux, J. and Operto, S.: An overview of full-waveform inversion in exploration geophysics, GEOPHYSICS, 74, WCC1–WCC26, 2009.

Wang, W., Wu, J., Fang, L., and Lai, G.: S wave velocity structure in southwest China from surface wave tomography and receiver functions, Journal of Geophysical Research: Solid Earth, 119, 1061–1078, 2014.

Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Štekl, I., Guasch, L., Win, C., Conroy, G., and Bertrand, A.: Anisotropic 3D full-waveform inversion, Geophysics, 78, R59–R80, 2013.

Wathelet, M.: An improved neighborhood algorithm: Parameter conditions and dynamic scaling, Geophysical Research Letters, 35, L09 301, 2008.

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J. F., and F., W.: Generic Mapping Tools: Improved version released, EOS Trans. AGU, 94, 409–410, 2013.

Wilde-Piórko, M.: Crustal and upper mantle seismic structure of the Svalbard Archipelago from the receiver function analysis, Polish Polar Research, 36, 145–161, 2015.

Wilde-Piórko, M., Saul, J., and Grad, M.: Differences in the crustal and uppermost mantle structure of the Bohemian Massif from teleseismic receiver functions, Stud. Geophys. Geod., 49, 85–107, 2005.

Wilde-Piórko, M., Świeczak, M., Grad, M., and Majdański, M.: Integrated seismic model of the crust and upper mantle of the Trans-European Suture zone between the Precambrian craton and Phanerozoic terranes in Central Europe, Tectonophysics, 481, 108–115, 2010.

Wilde-Piórko, M., Grycuk, M., Polkowski, M., and Grad, M.: On the rotation of teleseismic seismograms based on the receiver function technique, Journal of Seismology, pp. 1–12, 2017.