Joint inversion of receiver functions and apparent incidence angles for sparse seismic data

Rakshit Joshi[1], Brigitte Knapmeyer-Endrun[2], Klaus Mosegaard[3], Heiner Igel[4], and Ulrich Christensen[5]

[1]Max Planck Institute for Solar System Research
[2]Erdbebenstation Bensberg
[3]Niels Bohr Institute, University of Copenhagen
[4]Ludwig-Maximilians University, Munich
[5]Max-Planck-Institut für Sonnensystemforschung

November 24, 2022

Abstract

The estimation of crustal structure and thickness is instrumental in understanding the formation and evolution of terrestrial planets. Initial planetary missions with seismic instrumentation on board face the additional challenge of dealing with seismic activity levels that are only poorly constrained a priori. For example, the lack of plate tectonics on Mars leads to low seismicity which could in turn hinder the application of many terrestrial data analysis techniques. Here we propose using a joint inversion of receiver functions and apparent incidence angles, which contain information on absolute S-wave velocities of the subsurface. Since receiver function inversions suffer from a velocity depth trade-off, we in addition exploit a simple relation which defines apparent S-wave velocity as a function of observed apparent P-wave incidence angles to constrain the parameter space. We then use the Neighbourhood Algorithm for the inversion of a suitable joint objective function. The resulting ensemble of models is then used to derive uncertainty estimates for each model parameter. In preparation for analysis of data from the InSight mission, we show the application of our proposed method on Mars synthetics and sparse terrestrial data sets from different geological settings using both single and multiple events. We use information theoretic statistical tests as a model selection criteria and discuss their relevance and implications in a seismological framework.

Posted on 24 Nov 2022 — CC-BY 4.0 — https://doi.org/10.1002/essoar.10506471.1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Joint inversion of receiver functions and apparent incidence angles for sparse seismic data

Rakshit Joshi[1,4], **Brigitte Knapmeyer-Endrun**[2], **Klaus Mosegaard**[3], **Heiner Igel**[4], **Ulrich R. Christensen**[1]

[1]Max-Planck-Institute for Solar System Research, Göttingen, Germany
[2]Bensberg Observatory, University of Cologne, Cologne, Germany
[3]Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark
[4]Ludwig-Maximilians-Universitt, Munich, Germany

Key Points:

- We propose the joint inversion of receiver functions and apparent S-wave velocity curves to estimate crustal thickness
- Using the Neighbourhood Algorithm, we show how a full uncertainty estimate can be computed from an ensemble solution
- The method is applied to Martian synthetics and terrestrial data sets comprising single and multiple events

Corresponding author: Rakshit Joshi, `joshir@mps.mpg.de`

Abstract

The estimation of crustal structure and thickness is instrumental in understanding the formation and evolution of terrestrial planets. Initial planetary missions with seismic instrumentation on board face the additional challenge of dealing with seismic activity levels that are only poorly constrained a priori. For example, the lack of plate tectonics on Mars leads to low seismicity which could in turn hinder the application of many terrestrial data analysis techniques. Here we propose using a joint inversion of receiver functions and apparent incidence angles, which contain information on absolute S-wave velocities of the subsurface. Since receiver function inversions suffer from a velocity depth trade-off, we in addition exploit a simple relation which defines apparent S-wave velocity as a function of observed apparent P-wave incidence angles to constrain the parameter space. We then use the Neighbourhood Algorithm for the inversion of a suitable joint objective function. The resulting ensemble of models is then used to derive uncertainty estimates for each model parameter. In preparation for analysis of data from the InSight mission, we show the application of our proposed method on Mars synthetics and sparse terrestrial data sets from different geological settings using both single and multiple events. We use information theoretic statistical tests as a model selection criteria and discuss their relevance and implications in a seismological framework.

1 Introduction

Receiver function (RF) analysis is a powerful technique to gain information about the discontinuities in the crust and upper mantle beneath a single three-component seismic station. RFs are essentially time series that are sensitive to the structure near the receiver. The basic principle behind this method is that when a seismic wave is incident upon a discontinuity, mode conversion between the compressional (P) and shear (S) waves will take place in addition to the generation of reflected and transmitted waves. The resulting converted wave (Ps or Sp) will have a time offset with respect to its parent wave, and this time offset is directly proportional to the depth of the discontinuity and the velocity of the layers above. In addition to the direct converted waves, the multiples resulting from reflections and conversions between the discontinuity and the free surface can provide further constraints on the layer thickness and help to resolve the depth-velocity trade-off. The RF can be obtained by deconvolving the vertical component from the radial component of a teleseismic event recorded on a three-component seismometer (Langston, 1979; Owens et al., 1987; Ammon, 1991). Since only a small percentage of the incident energy is converted at a discontinuity, it is difficult to observe these conversions in a single seismogram. A number of RFs can instead be used to measure the crustal thickness and average v_P/v_S ratios by H-k (crustal thickness - average v_P/v_S) stacking for individual stations (Zhu & Kanamori, 2000; Helffrich & Thompson, 2010) or imaging by CCP (Common Conversion Point) stacking of data from many stations (Dueker & Sheehan, 1997). This, however, requires assumptions on the velocity structure.

One method to obtain a detailed velocity structure is to directly invert the calculated RFs using linearised iterative procedures, but Ammon et al. (1990) showed that such inversions of RF contain an inherent trade-off between the depth to a discontinuity and the velocity above. The primary sensitivity of the RF inversion is to velocity contrasts and relative travel time, not to absolute velocity. This lack of sensitivity to absolute velocity results from the relative S - P travel time constraints along with the limited range of horizontal slowness contained in the data (Ammon et al., 1990). Thus RF data sets are generally inverted jointly with other independent data sets that provide additional constraints on absolute shear wave velocities like surface wave dispersion curves (e.g. Du & Foulger (1999); Julia et al. (2000)), or Rayleigh wave ellipticity (Chong et al., 2016). One such relation which has not been heavily exploited is between apparent S-wave velocities and P-wave polarisation. The polarisation of body waves has been traditionally used in seismology to study the anisotropy of crustal and upper mantle structures (Schulte-Pelkum et al., 2001; Fontaine et al., 2009). But the P-wave polarisation can also be used to constrain the near surface shear wave speed. Svenningsen & Jacobsen (2007) showed that the amplitudes of the vertical (Z) and radial (R) components of the P-receiver function at zero time is directly related to the polarisation of P-waves. Deconvolution removes the complex waveform of the incoming P-waves, which dominate the Z component. Hence the Z RF is an approximate zero-phase spike with arrival instant at exactly t=0, where the time is measured relative to the P-wave arrival. This can be used to estimate the apparent P-wave incidence without influences from the P-wave coda. Further, filtering at successively long periods, a frequency dependent apparent shear wave velocity profile can be obtained (Svenningsen & Jacobsen, 2007; Knapmeyer-Endrun et al., 2018) which can be used as an effective independent data set to be jointly inverted with the RFs.

Svenningsen & Jacobsen (2007) used a linearised inversion of apparent S-wave velocity curves and demonstrated its independence of the starting model. Hannemann et al. (2016) applied the method to an OBS data set and used a grid search method concluding that the method is usable for single station estimates of the local S-wave velocity structure beneath the ocean bottom. Schiffer et al. (2016) used an iterative least squares method to jointly invert apparent velocity curves and RFs utilising a minimum number of layers (6-8). Knapmeyer-Endrun et al. (2018) used a grid search over parameter space to invert the S-wave velocity curve for crustal structure at several Earth stations with varying geology and synthetic Mars data. It has also been shown that a priori S-wave velocity information deduced from P-wave polarisations can be useful when inverting RF waveforms (Peng et al., 2012). Park & Ishii (2018) further showed that the S-wave polarisation is sensitive to both the compressional and shear wave speeds, and successfully combined P- and S-wave polarisation directions measured by principal component analysis to derive the distribution of near-surface P- and S-wave speeds in Japan.

In this paper, we use a modified version of the Neighbourhood Algorithm (Sambridge, 1999a; Wathelet, 2008) for the joint inversion of receiver functions and apparent S-wave velocity profile. The Neighbourhood Algorithm (NA) is a derivative-free optimisation method which uses a pseudo-random trajectory in exploring the parameter space. Rather than making inferences on model parameters using only the lowest-misfit model, it provides the option of using the suite of all generated models for this purpose. With a well sampled parameter space, an ensemble algorithm also benefits from the possibility of a probabilistic solution with full uncertainty estimates. In contrast with earlier studies on this topic which are predominantly based on large amounts of available data, we show how this method can be used with limited data sets comprising only a few events. This becomes crucial in the context of planetary seismology where the amount of data may be limited. For example, it can be used to study the crustal structure of Mars using data from the InSight mission (Lognonné et al., 2019). Another problem associated with determining the crustal structure is the number of inter-crustal layers to be inverted for. We address this problem using a two-fold approach: we start by inverting for a model of low complexity and gradually increase it till no significant velocity contrast along with misfit reduction is observed, with major discontinuities being adequately represented by the model. We then use Akaike weights derived from AIC (Akaike Information Criterion) values (Akaike et al., 1973) for all of these models as a selection criteria. We apply this joint inversion scheme on synthetic seismograms for Mars and selected terrestrial data.

2 Datasets

2.1 Mars Synthetics

In order to demonstrate and verify our proposed method, we first use synthetic seismograms for Mars that are generated using Greens Function (GF) databases prepared for a suite of apriori 1D velocity models with varying crustal thicknesses, seismic wave speeds, densities, mantle compositions and aerotherms. These apriori models are obtained by the inversion of bulk chemistry, mineralogy and geotherm, following the approach described in Khan & Connolly (2008), Connolly (2009), and Khan et al. (2016). The GF databases are computed using a 2.5D axis-symmetrical spectral element code, AxiSEM (Nissen-Meyer et al. 2014), and are publicly available within the Marsquake Service (MQS) at ETH Zurich ((Ceylan et al., 2017), http://instaseis.ethz.ch/marssynthetics/). Synthetic broadband seismograms can be calculated from these GF databases for arbitrary moment tensors and source receiver combinations using the Instaseis package (van Driel, Krischer, et al., 2015). These simulations are based on full numerical solutions of the visco-elastic wave equation and include the effects of attenuation, are accurate down to a period of 1 s, and allow for a total simulation duration of 30 minutes.

Since a large variation in crustal thickness is expected across Mars, a thin (30 km) and thick (80 km) crust is employed to create the initial models, both with a 10 km thick upper crustal layer. Further details of these models can be found in Ceylan et al. (2017). The thin and thick crusts with different velocity contrasts at the Moho represent 1-D global end-member models, rather than what is expected beneath the InSight landing site. In this paper we have used two thin crust models (C30VH_AKSNL, C30VL_AKSNL) and one thick crust model (C80VL_AKSNL) for the purpose of demonstrating the method. For all of these models, we calculated synthetic seismograms and receiver functions at epicentral distances between between 15° and 180° in 1° increments. Assuming normal faulting, a dip slip source at an angle of 45° and at a depth of 5 km due north of the seismometer was used to generate the synthetic waveforms. Since the synthetics do not have any added noise, we assume a reasonable 25% standard deviation on mean absolute values of RFs and $V_{S,app}$ whenever appropriate for likelihood calculations. We demonstrate the results of applying our method first on a single event and then multiple events together.

2.2 Terrestrial Data

To verify how the algorithm works in a real setting, we analysed data from two stations in Central Europe - BFO in Germany (Federal Institute for Geosciences and Natural Resources, 1976) and SUW in Poland (GEOFON Data Centre, 1993). Reference values of crustal thickness for these stations were taken from the Moho depth map of the European plate (Grad et al., 2009) and Knapmeyer-Endrun et al. (2014). Because these sites have known differences in crustal structure, this gives us the opportunity to test how the method works in a range of possible scenarios and in the presence of noise. Station BFO is located on the thinned crust of the Upper Rhine Graben which is a part of the European Cenozoic Rift system (Ziegler, 1992). In contrast to this, station SUW is situated on the relatively thick East European Craton which is the core of the Baltica proto-plate and occupies the north-eastern half of Europe. It is characterized by a thick three-layer crust with an additional fast lower crustal layer (Grad et al., 2003). The East European Craton is of Precambrian origin and overlain by a young thin sedimentary cover (Bogdanova et al., 2006) which leads to strong reverberations in the P-receiver function for SUW (Wilde-Piórko et al., 2017)

3 Method

3.1 Calculation of Receiver Functions

The teleseismic P-wave receiver function represents the structural response near a recording station to the incoming teleseismic P-wave. It can be obtained by removing the source wavelet, propagation effects and the instrument response from the vertical, radial and transverse waveforms. This is generally done by deconvolving the vertical component from the radial and transverse components in a process called source equalisation (Vinnik, 1977; Phinney, 1964). Several methods have been described in the literature for this deconvolution process (e.g., see Vinnik (1977), Phinney (1964), Langston (1979), Owens et al. (1987), Kind et al. (1995)) Here we use a time-domain Wiener filter for deconvolution as described by Hannemann et al. (2017). The synthetic seismograms do not require the removal of any instrument response, but they are filtered between 1 Hz and 50 s, 1 Hz being the upper frequency limit of the synthetics. Additionally, due to the alignment of source and receiver, these data are already in the ZRT system. For the terrestrial data, we first remove the instrument response from all components and then filter the seismograms between 5 Hz and 50 s. The ZNE coordinate system is then rotated into ZRT using back-azimuths determined by polarization analysis (Jurkevics, 1988) to obtain radial and transverse components. The Wiener filter is determined such that it transforms the P-wave signal on the vertical component into a band-limited spike. This filter is then applied to all components of the signal to finally obtain the RF with the spike positioned at the centroid of the signal.

3.2 Apparent S-wave velocity

Following the relationship between true and apparent incidence angles (Wiechert, 1907), it can be shown that the apparent incidence angle is sensitive to absolute shear wave velocity

$$v_{S_{,app}} = \sin(0.5\overline{i_p})/p$$

where $\overline{i_p}$ denotes the apparent P-wave incidence angle and p denotes ray parameter. Svenningsen & Jacobsen (2007) proposed a method to directly estimate the apparent incidence angle using RFs instead of the raw waveform data which in turn emphasised the true S-wave velocity information contained in them. We follow a similar procedure and estimate the apparent P-wave incidence angle from the amplitudes of vertical and radial receiver functions at time t=0 using the relation

$$\tan\overline{i_p} = \frac{RRF(t=0)}{ZRF(t=0)}$$

Now estimating $\bar{i_p}$ as a function of low pass Butterworth filter period (T) results in a $v_S(T)$ curve which emphasises the absolute S-wave velocity variation with depth. Larger T implies more smoothening and thus more multiples at later times influence the values of the filtered receiver functions at t=0. In contrast with the squared cosine filters used by Svenningsen & Jacobsen (2007), we use a Butterworth filter which has twice the corner period as a cosine filter. For each trace we measure the dominant period of the spike in the ZRF and discard the values of filter periods smaller than that. We show cases with both single and multiple events. When multiple events are used at varying epicentral distances, we calculate the median of the apparent S-wave velocity curve at each sample period.

3.3 Inversion

For the purpose of this study, we have employed a modified version of the Neighbourhood Algorithm (NA) (Wathelet, 2008) for the joint inversions of RF and apparent S velocity curves. Being a derivative free optimisation algorithm and taking into account the low dimensionality of our problem, NA seems to be a good choice because of its simplicity (two tuning parameter scheme) and lack of dependence on starting models (Sambridge, 1999a). Moreover, an ensemble of models rather than a single model can be used to make robust statistical inferences about the model parameters. The modifications by Wathelet (2008) further implement dynamic scaling of model parameters and allows to define irregular limits to the searchable parameter space. The idea behind the NA is to start with an initial coarse sampling of the parameter space, then select the regions with lowest misfits and continue to resample these regions such that the heaviest sampled regions correspond to the models which best fit the data. In each iteration, the NA uses nearest-neighbour regions defined by Voronoi cells to partition and search the parameter space. The misfit is assumed to be constant within each of these Voronoi cells, and with each iteration, sampling is concentrated on the cells with lower misfit relative to the rest of the cells. The algorithm relies on only two control parameters : Ns - number of new samples to generate at each iteration and

Nr - number of promising models to select for further sampling. The ratio Ns/Nr controls whether the algorithm behaves exploratively or exploitatively (Sambridge, 1999a,b).

We use the L2 norm in order to measure how well a given model with a particular set of parameters can reproduce the given data quantitatively

$$\Phi(m) = \left\| \frac{g(\mathbf{m}) - d_{obs}}{\sigma_d} \right\|^2$$

where $g(\mathbf{m})$ is the estimated data and σ_d^2 is the estimated variance of the data noise. In this study, the noise has been assumed uncorrelated for simplicity and thus a simple Euclidean distance can be used. For a joint inversion of receiver function and apparent S-wave velocity, the objective function is defined by the linear combination of misfits of the weighted receiver functions Φ_{RF} and the apparent velocity curve $\Phi_{V_{app}}$, using the L2 norm, thus takes the form

$$\Phi(m) = \alpha\Phi_{RF} + \Phi_{V_{app}} \tag{1}$$

The weighting constant α is tuned manually by sample forward runs prior to the inversion process such that both the individual misfits are of the same order of magnitude. As mentioned before, the two parameters that control the NA need to be tuned depending on the problem and the style of sampling needed. For a more explorative search that is robust against local minima, we perform 1200 iterations in each inversion run with 300 models produced at each iteration (n_s) and 100 cells re-sampled at each iteration (n_r), resulting in an ensemble of $\sim$ 360000 models per run. Each inversion was repeated several times to test the stability of the results. High n_s/n_r ratio ensures faster convergence while a high number of initial models ($n_{s_0} = 3000$) ensures highly explorative behaviour.

Knapmeyer-Endrun et al. (2018) compared several algorithms used in literature for the computation of receiver functions before choosing the forward calculation implemented by Shibutani et al. (1996). The algorithm calculates the impulse response of a layer stack in the P-SV system. We then convolved the resulting synthetic Z- and RRFs with the observed ZRFs to account for the observed complexity and waveform widths. Once the RFs are obtained, we can straight away calculate the apparent S wave velocities using the procedure described in the last section. Density was not considered to be a parameter to be inverted for and was calculated using Birch's law (Birch, 1961), while the S-wave velocity and the v_P/v_S ratio were allowed to vary. Furthermore, the S-wave velocity was constrained to increase with increasing depth. The fact that a single forward calculation can be performed in a matter of seconds and the waveform complexity matches that of real data makes this algorithm suitable for the propose of this study.

3.4 Bayesian Formulation

The Bayesian formulation allows to account for prior knowledge of the parameters of our model, provided that this information can be expressed as a probability distribution $\rho(\mathbf{m})$. The prior corresponds to the knowledge that we have about our system, for example from previous studies. As new data is available, often in the form of likelihoods, this prior information can then be updated using Bayes' rule. This results in what is known as the posterior distribution for these unknowns - a distribution over the full range of these parameters.

3.4.1 Computing average Likelihoods

The likelihood $\rho(\mathbf{d_{obs}}|\mathbf{m})$ is a function of the model parameters that describes the goodness of fit of a model to the observed data. Asssuming a Gaussian error distribution for a given misfit measure, $\mathbf{\Phi(m)}$, the likelihood function is defined as :

$$\rho(\mathbf{d_{obs}}|\mathbf{m}) \propto exp\left(\frac{-\Phi(m)}{2}\right)$$

As mentioned before, the NA initially starts with a coarse sampling of the parameter space, and eventually the algorithm guides the sampling such that the best fitting regions of the parameter space are also the most heavily sampled regions. This therefore introduces a bias in the sampling of the parameter space which otherwise could be used to compute the full uncertainty from the ensemble of acceptable solutions. Sambridge (1999b) demonstrates that this could be achieved by a Gibbs re-sampling of the output ensemble which essentially concentrates on the low misfit regions and approximates the true posterior density by an approximate one. Here we show a simple alternative method to compute marginal histograms from the biased samples based on binning model parameters. In essence, each model in the ensemble has a pair-wise distance to every other model which can be calculated using multidimensional scaling. Binning model parameters within a small distance and computing average likelihoods then approximates the true posterior density as a histogram.

Consider N sample models $\mathbf{m^{(1)}}, ..., \mathbf{m^{(N)}}$ in a K-dimensional space, distributed according to an (everywhere positive) unknown distribution $\nu(\mathbf{m})$. Assume that $\nu(\mathbf{m})$ is close to the distribution, $f(\mathbf{m})$, and that we wish to compute the marginal histograms $f_k(m_k)$ from the samples.

The height $h_{[a,b]}$ of the histogram column for an interval [a,b] must (for $N \to \infty$) be proportional to the marginal probability $P_k(a < m_k < b)$. Hence,

$$h_{[a,b]} \approx \int_a^b f_k(m_k) dm_k$$

except for a normalization factor. This can be re-written as a mean value (expectation) of the ratio $\frac{f_k(m_k)}{\nu(m_k)}$ over the interval $[a, b]$ with respect to $\nu(m_k)$:

$$h_{[a,b]} \approx \int_a^b \frac{f_k(m_k)}{\nu(m_k)} \nu(m_k) dm_k$$

and since the sample models $\mathbf{m^{(1)}}, ..., \mathbf{m^{(N)}}$ are distributed according to $\nu(\mathbf{m})$, we have the approximation:

$$h_{[a,b]} \approx \frac{1}{N} \sum_{\{i|a<m_k^{(i)}<b\}} \frac{f_k(\mathbf{m}^{(i)})}{\nu(\mathbf{m^{(i)}})}$$

This expression can be used when f_k can be evaluated in the sample points, and when we can evaluate $\nu(\mathbf{m^{(i)}})$ from the density of sample points. The density at $\mathbf{m}^{(i)}$ can, e.g., be evaluated over a cube C with edge length Δm, centered at $\mathbf{m}^{(i)}$:

$$\nu(\mathbf{m^{(i)}}) = \frac{1}{(\Delta m)^K} N_c$$

where N_C is the number of sample points in C

3.4.2 Priors

We impose a minimal prior knowledge on all the parameters by using the uniform distribution as our choice of priors. The prior for each parameter takes a constant value over a defined interval. For example, if X is a model parameter which can take values over the interval $\Delta X = (X_{max} - X_{min})$, we define the prior probability density as :

$$\rho(x_i) = \begin{cases} \frac{1}{\Delta X}, & \text{if } X_{min} \leq x_i \leq X_{max} \\ 0, & \text{otherwise} \end{cases}$$

We can now apply Bayes' rule (Bayes, 1763) to combine the likelihood of observing the data with the prior distribution and to give the posterior probability density function:

$$\rho(\mathbf{m}|\mathbf{d_{obs}}) \propto \rho(\mathbf{d_{obs}}|\mathbf{m})\rho(\mathbf{m})$$

Note that the denominator in the Bayes' rule, $\rho(\mathbf{d_{obs}})$, which is a sum over all possible models has been treated as a constant in this work, leading to a proportionality sign in the equation.

3.5 Model Selection

We use Akaike's Information Criterion (AIC) (Akaike et al., 1973) as a model selection criterion, which essentially gives the Kullback-Leibler divergence between a candidate model and the true model as

$$AIC = 2k - 2ln(L)$$

where k and L denote the number of model parameters and the value of maximum likelihood of the model, assuming Gaussian errors. The first term in this equation is a measure of fit between the synthetic model and the true model representing the reality; the second term penalizes the order of complexity of this synthetic model. While raw AIC values themselves have no meaning, the quantity $exp\left(\frac{AIC_{min} - AIC_i}{2}\right)$ is an estimate of the relative likelihood of the ith model. These model likelihoods can then be normalized to obtain Akaike weights w_i(AIC)(Burnham & Anderson, 2002; Wagenmakers & Farrell, 2004),

$$w_i(AIC) = \frac{exp\{-0.5\Delta_i(AIC)\}}{\sum_{k=1}^{K} exp\{-0.5\Delta_k(AIC)\}}$$

which can be interpreted as the probability that the ith model is the best (i.e., it minimizes the estimated information loss (Anderson & Burnham, 2004)). The strength of evidence in favour of one model over the other can then also be obtained by dividing their respective Akaike weights. When the number of samples is small, a correction factor is added to the above equation giving the corrected AIC (AICc) values

$$AICc = 2k - 2ln(L) + \frac{2kn}{n-k-1}$$

Here k denoted the number of model parameters and n the number of independent samples. Since the samples of a seismogram are generally correlated, with the correlation length being proportional to sampling frequency, we instead use the product of the Nyquist rate and the signal length as a measure of the number of independent samples (van Driel, Wassermann, et al., 2015). For a band limited signal, the Nyquist rate is given by $2 * (f_{high} - f_{low})$ which gives 1.96 Hz and 9.96 Hz for synthetics and terrestrial data, respectively (f_{high} and f_{low} denote the upper and lower frequency limits). Anderson & Burnham (2004) suggest using AICc when the ratio between the sample size n and the number of model parameters k is low (< 40). We will therefore use AICc when dealing with synthetic data and AIC for terrestrial data.

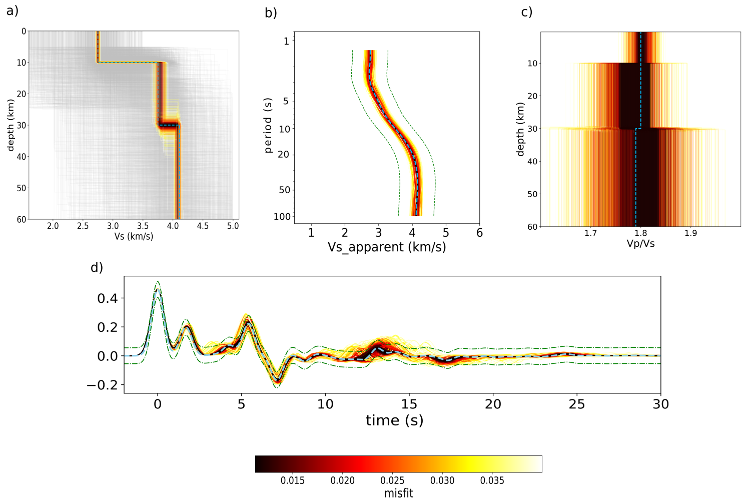

Figure 1. Result for thin crust model C30VH_AKSNL and event distance 70° (a) 1-D velocity profile. The light gray lines represent traversed models outside the maximum misfit range. The blue dashed line represents the true model. (b) Fit to $v_{S,app}$ (c) v_P/v_S ratio as a function of depth (d) Fit to receiver function waveforms. The blue dashed lines denote the observed data and the green dash-dotted lines represent the uncertainty in observations.

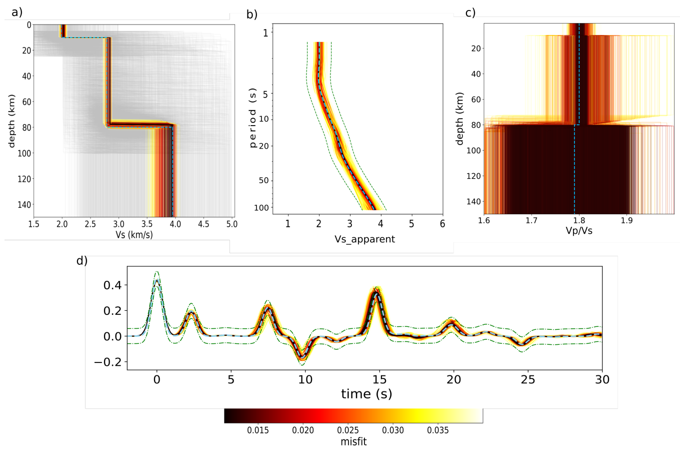

Figure 2. Same as Figure 1 for C80VL_AKSNL. Event distance is 40°.

4 Results

4.1 Mars Synthetics

Figures 1 and 2 show the result of applying the method on single events for a priori Martian velocity models with a thin fast (C30VH_AKSNL) and a thick slow (C80VL_AKSNL) crust, respectively. Since noise is not a limiting factor here, in both cases, the residual includes the misfit for the complete waveform up to 30 s and apparent S wave velocity to 117 s. Each inversion was repeated 3 times to test the stability and the results were concatenated. The plots include all models within a maximum misfit value, ranked and color coded according to misfit with black models being the best fitting solutions. This maximum misfit value is derived such that it encompasses the best 25% of all the models in the ensemble.

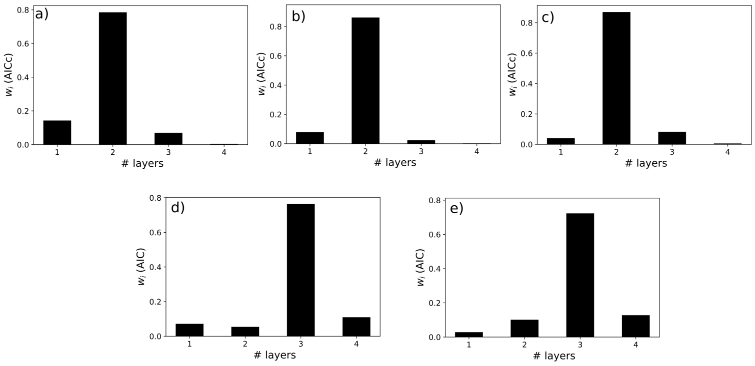

Figure 3. Model probabilities based on AICc values for (a) C30VH_AKSNL (b) C80VL_AKSNL (c) C30VL_AKSNL and AICc values for (d) BFO (e) SUW

Adding a third layer to the model parameterization did not produce any considerable changes to the result. For C30VH_AKSNL the additional third layer produced a velocity contrast of around 0.8 % against the layer adjacent to it with an insignificant misfit drop, while C30VH_AKSNL produced a similar low velocity contrast of around 0.45 %. This shows that an additional layer is not warranted by the data. This is also confirmed numerically by our model selection criteria. Figures 3(a) and (b) show the respective probabilities obtained from AICc values for 1, 2, 3 and 4 layer models with constant velocity over a half space for C30VH_AKSNL and C80VL_AKSNL respectively. For C30VH_AKSNL, there is a higher probability ($\sim 16\%$) of explaining the data with just a single layer than for C80VL_AKSNL. This is consistent with a weak Moho signal produced by the small velocity contrast. Since the 2 layer model has the highest probability (and thus minimum AIC), we conclude that it is the optimum model that explains this data set. This is also in agreement with the true models indicated by the blue dashed lines in figures 1 and 2. The apriori range for each parameter for both 2 layer and 3 layer cases are shown in Figures 4 and 5.

The top layer crustal S-wave velocity and transition depth is well resolved for both the representative end member models. For C30VH_AKSNL, there is a high uncertainty in the Moho depth which in turn escalates the uncertainty in the S-wave velocity in the lower crust. This might be explained as the direct converted phase and the multiples produced by the intra-crustal discontinuity at 10 km depth are clearly visible in the data while the Moho conversion for the thin crust model is not readily recognizable. This is in contrast to C80VL_AKSNL where the direct converted phase and the multiples produced at the Moho are clearly visible. The mantle S-wave velocities on the other hand are better constrained for C30VH_AKSNL than for C80VL_AKSNL. This is explained by the $v_{S,app}$ curves for the models. The $v_{S,app}$ curve for C80VL_AKSNL does not contain any information on the upper mantle velocity within its period range whereas in the $v_{S,app}$ curve for C30VH_AKSNL, the velocities converge to the upper mantle velocity of 4.1 km/s for periods longer than ~ 50 s. This clearly demonstrates the advantage of inverting receiver functions along with frequency dependent apparent S-wave velocities.

In both cases, the v_P/v_S ratio is also fairly well constrained for the top two layers by the method, as can be seen in the sub-figures (d). This is in agreement with Sambridge (1999a), where it was shown that the v_P/v_S ratio from the NA inversion is better resolved in the top layers than for the deeper ones. The thickness of the layers and their corresponding S-velocities are also better constrained than the v_P/v_S ratio. For C80VL_AKSNL, the v_P/v_S ratio of the half-space is not well resolved and varies across the whole model range

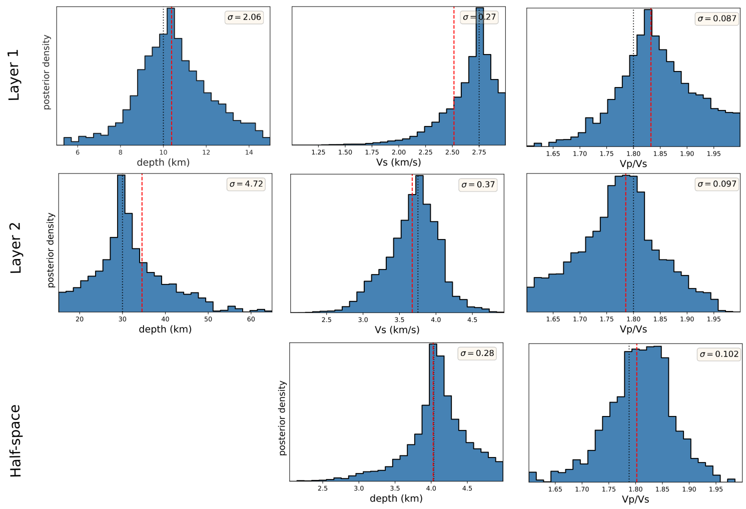

Figure 4. C30VH_AKSNL : 1D marginal posterior densities of depth, velocity and v_P/v_S ratio for each layer. The half-space has no depth parameter. The red dashed line denotes the mean value and the black dotted line represents the true parameter value.

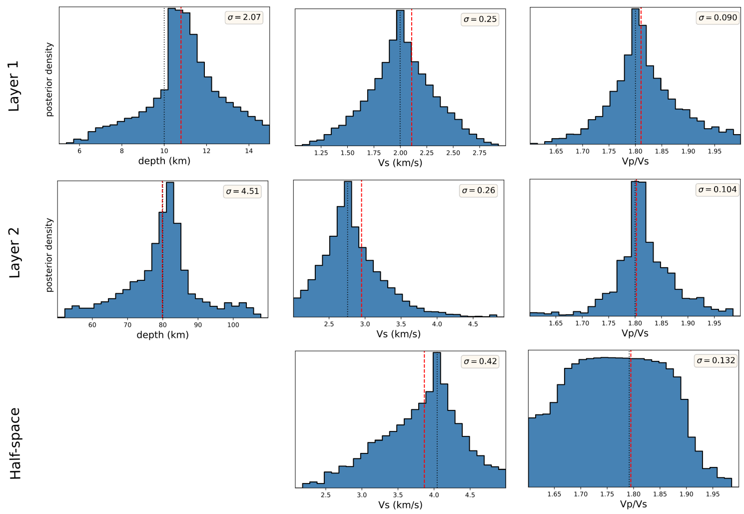

Figure 5. Same as Figure 4 for C80VL_AKSNL

investigated, whereas for C30VH_AKSNL, it is adequately resolved for all the layers even though the variance increases with depth.

To test how the method performs when multiple events are available, a median $v_{S,app}$ curve was calculated for model C30VL_AKSNL from the RFs between 40° to 90° where the the $v_{S,app}$ curves are similar for each distance (Knapmeyer-Endrun et al., 2018). This median $v_{S,app}$ curve was then jointly inverted with 6 receiver functions selected at epicentral distances of 90°, 80°, 70°, 60°, 50° and 40°. The resulting profile along with the waveform fit for each RF and $v_{S,app}$ curve is shown in Figure 6. The velocity profile lies well within the range of the uncertainty and the receiver function at each distance is also well modelled. The variance in velocity again increases with depth and is maximum for the mantle. The

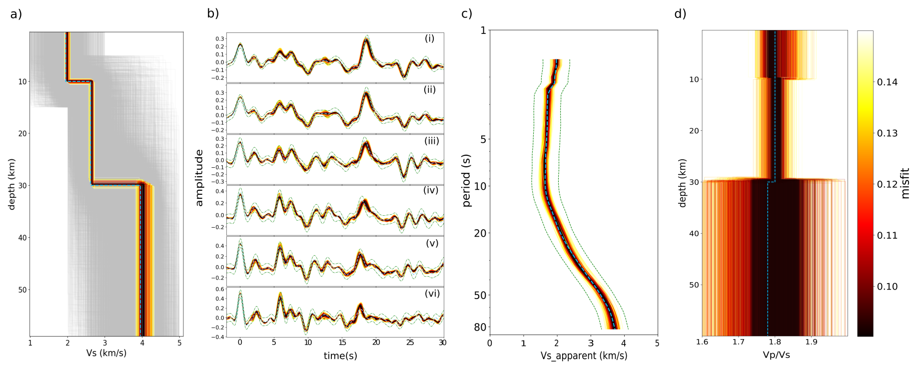

Figure 6. Example of multiple inversions for C30VL_AKSNL (a) 1-D velocity profiles. The light gray lines represent traversed models outside the maximum misfit range. (b) Fit to receiver function waveforms at epicentral distance of (i)90° (ii) 80° (iii) 70° (iv) 60° (v) 50° and (vi) 40° (c) Fit to the median $v_{S,app}$ (d)v_P/v_S ratio as a function of depth. The blue dashed curves denote the observed data and the green dash-dotted lines represent the data uncertainty.

median $v_{S,app}$ curves are also close to the observed curve, even though the kinks between 2 - 3 s and 7 s appear to be slightly sharper than in the observed curve. Unlike C30VH_AKSNL, C30VL_AKSNL has a shorter $v_{S,app}$ curve extending to 82 s. This restricts the retrieval of S-wave velocity information from longer periods and has the effect of an increased variance in the upper mantle velocity. The Moho on the other hand is well resolved due to a high impedance contrast which results in a direct phase at around 6 s for RFs at 40° and 50° , and a clear multiple at around 19 and 24 seconds for RFs at 90°, 80° and 70°. Looking at the probability densities we see that using more data has the effect of an overall decrease in uncertainty levels. From Figure 3(c), we see that the data is best explained by a 2 layer model which has the highest value for w_j(AICc). To check whether there is a decrease in the depth velocity trade off, we further compared the density plots of Moho depth and the velocity above with the results from a direct receiver function inversion which did not employ $v_{S,app}$ as an additional constraint. Here we used the best 25% models of the respective ensembles. It is evident from the Figure 8 that along with a gain in accuracy, there is a considerable reduction in trade-off between depth and velocity in the case of the joint inversion. For an application of the method to synthetic data with added noise see Drilleau et al. (2020).

4.2 Terrestrial Data

The examples above from synthetic data show that in principle the joint inversion of apparent S-wave velocity with receiver functions serves as a useful complement. This section presents inversion results for terrestrial data where the inherent data noise becomes an important consideration and has a strong influence on the resulting model parameters and their associated uncertainties. Figure 9 (a) shows the noise levels computed for stations BFO (green) and SUW (blue) using the pre-event noise of the radial component of the receiver functions since they should ideally be independent and non-correlated. For each station we calculate the mean of the pre-event noise of the radial component of each receiver function from all the events considered here for multiple inversions and bin them according to amplitude, creating a distribution from which noise parameters can be estimated. The variance in the noise level was found to be the higher for SUW with each roughly following a Gaussian distribution. Similarly Figure 9 (b) shows the noise characteristics for the $v_{S,app}$ curve for both the stations calculated by binning of residuals from the median curve.

Selection of the model complexity that best describes the data is again done using the procedure described in the previous section. Starting at a low degree, we gradually increase

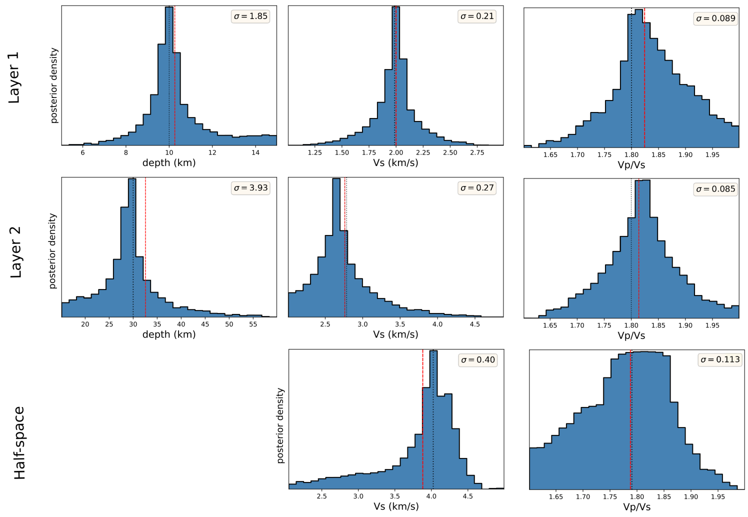

Figure 7. Same as Figure 4 for C30VL_AKSNL

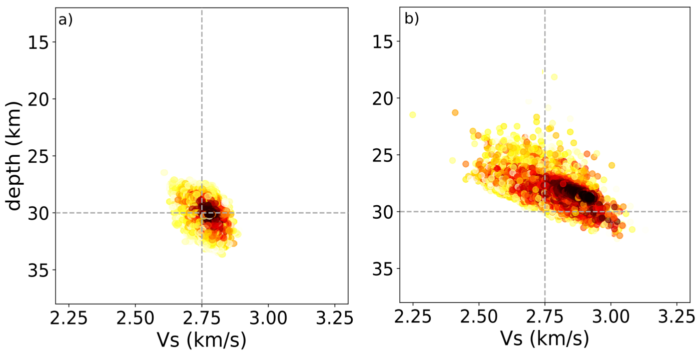

Figure 8. Comparison of depth velocity trade off for (a) Joint inversion of RF with $v_{S,app}$ (b) RF inversion without $v_{S,app}$. The grey dashed lines denote the true values of depth and velocity.

the complexity until the parameterization produces no significant deviation in profile and misfit reduction. We then compare the corresponding relative likelihood values and choose the maximum.

The results for seismic station BFO are summarized in Figure 10. From the velocity profile (subplot (a)) we can see that the data can be sufficiently described by a minimum parameterisation comprising 3 layers with constant velocity over a half-space - a low velocity top layer of sediments, an upper crustal layer extending from the base of the sediments to a depth of $\sim$ 7 km and a thick lower crust that extends from 7-8 km to the Moho at $\sim$ 25 km depth. Various studies found the Moho depth between 23.8 and 27 km for station BFO (Geissler et al., 2008; Knapmeyer-Endrun et al., 2014; Grad et al., 2009). The mantle velocities are also adequately constrained by the data showing a maximum probability for mantle v_S velocity of 4.6 km/s. The results for the S-wave velocity model also show close agreement with Svenningsen & Jacobsen (2007) (shown in blue dashed lines) and Knapmeyer-Endrun et al. (2018) (shown in green dashed lines). Since Svenningsen & Jacobsen (2007) used the apparent velocity curve up to 0.2 s in contrast to 1.3 s allowed by our data-set, the top sediment layer could be better resolved to thickness values below 1km.

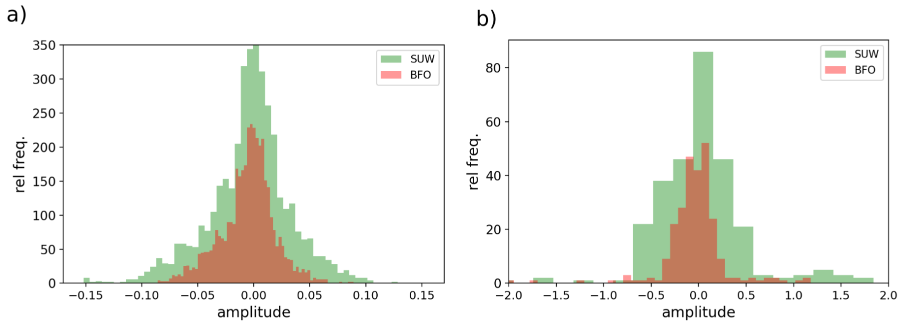

Figure 9. Noise characteristics of (a) RF shown as a frequency distribution of amplitude calculated from radial component of receiver functions for different stations (b) $v_{S,app}$ calculated as a frequency distribution of error from the median curve.

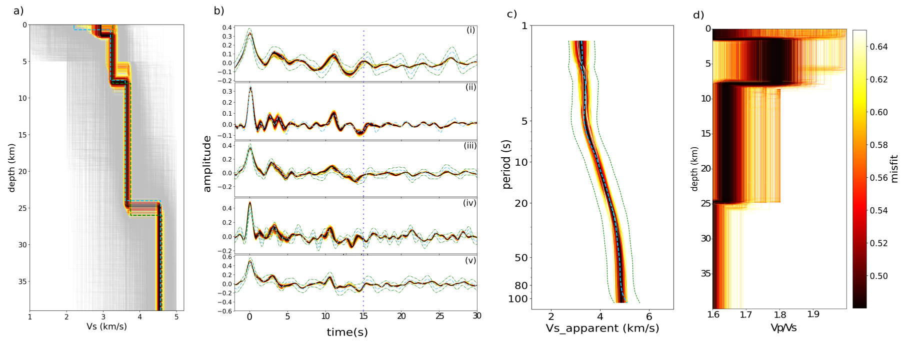

Figure 10. Example of joint inversions for terrestrial data from station BFO (a) 1-D velocity profiles. The blue and green dashed line represents the results from Svenningsen & Jacobsen (2007) and Knapmeyer-Endrun et al. (2018). The light gray lines represent traversed models outside the maximum misfit range.(b) Fit to receiver function waveforms at epicentral distance of (i) 82° (ii) 79° (iii) 70° (iv) 51° (v) 45°. The blue dashed curve denotes the observed radial RFs and green dashed lines represent the standard error. The dark blue dotted line at 15s shows the end of the misfit window. (c) Fit to the median $v_{S,app}$ (d) v_P/v_S ratio as a function of depth

Subplots (b) and (c) show the corresponding fits to the receiver function for each event and a median $v_{S,app}$ curve. Except for the RF waveform in event (i) where the phase at ~ 10 s is over-pronounced, the models fit the data from other events adequately well. The modelled $v_{S,app}$ curve also follows the data closely at all periods, including the sharp kink around ~ 2 s. At longer periods after ~ 50 s, the velocities seem to converge to ~ 4.8 km/s providing a tight constraint on the upper mantle which explains the low uncertainty seen in the the half space v_S.

Station SUW is located on the East European craton and sits on a relatively thicker crust than BFO. Using a similar parameterization as before with 3 layers including a top sedimentary layer results in a subsurface velocity profile shown in Figure 11 (a). The model predicts the Moho to be located at a depth of ~ 45 km with the highest probability density and an intra-crustal discontinuity at 15 km. Previous studies have estimated the Moho depth to lie between 41 km and 46.8km for station SUW (Geissler et al., 2008; Knapmeyer-Endrun et al., 2014; Grad et al., 2009). The thickness and v_S of the sedimentary layer, however, are not well constrained with the uncertainty for v_S being the highest amongst all all layers. This is also evident from the modelled $v_{S,app}$ curves (subplot (c)) which show a

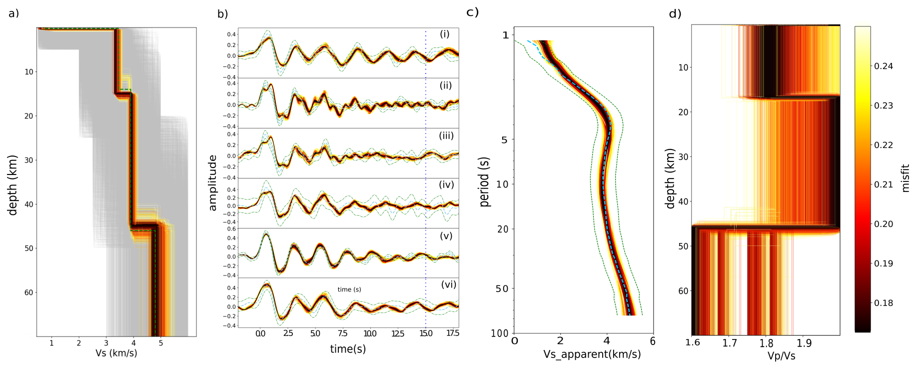

Figure 11. Same as Figure 10 for station SUW (b) shows the fit to receiver function waveforms at epicentral distances of (i) 82° (ii) 77° (iii) 72° (iv) 68° (v) 64° (vi) 60°

slight deviation from the observed curve at short periods. Such a deviation could indicate that the sedimentary layer is more complex than our parameterization which models it simply as layer with constant velocity. An increase in the model complexity (e.g., modelling the sedimentary layer with a velocity gradient) could lead to a better fit here as suggested by Knapmeyer-Endrun et al. (2018). Further, the missing v_S information at long periods in the observation leads to an increase in uncertainty in the upper mantle velocity which shows the highest probability density at a value of ~ 4.9 km/s. The modelled RFs shown in Figure 11 (b) clearly show the ringing effect with gradual decrease in amplitude with time caused by the thin sediment layer. These strong reverberations produce high amplitude oscillations in the early part of the signal and completely masks the direct Moho conversion at $\sim 6s$. This example in particular shows that caution is needed to interpret receiver functions with a sedimentary layer in terms of subsurface structures.

Figures 3(d) and 3(e) show the respective model probabilities obtained from AIC values. We see that both the data can be best explained by 3 layer models with constant velocity over a half space. However, there is still $\sim 9\%$ probability for a 4 layer model in both cases. The resulting values for v_P/v_S for each layer are also shown in suplots (d) in Figures 10 and 11. Unlike the case for synthetics, a high variation is observed here between the layers. In all the examples, the top sediment cover shows the highest uncertainty. The first and second layers are better resolved. The average v_P/v_S values estimated from RF analysis in previous studies are between 1.69 and 1.75 for BFO (Geissler et al., 2008; Knapmeyer-Endrun et al., 2014)) and between 1.81 and 1.84 for SUW. We find that the mean values from our results are broadly similar with values of 1.67 and 1.82, respectively.

5 Summary and Conclusion

In the context of the InSight mission, receiver function analysis has been envisioned as a likely method to study the crustal structure of Mars (Panning et al., 2017). In order to diminish the depth-velocity trade off inherent in travel time methods, we propose to use the information provided by apparent P-wave incidence angles derived from P-receiver functions as an additional constraint (Knapmeyer-Endrun et al., 2018). In this study, we present a method for joint inversion of receiver functions and frequency dependent apparent S-wave velocity curves using the Neighbourhood Algorithm. This results in an ensemble of model solutions along with their respective misfit values which can in turn be used to compute the full uncertainty of the model parameters. We then develop such a probabilistic solution using the resultant ensemble and apply this method to various data sets. Further, determining the sufficient number of layers for an optimal model presents another challenge in waveform inversion. We tackle this by gradually increasing the number of layers till adding

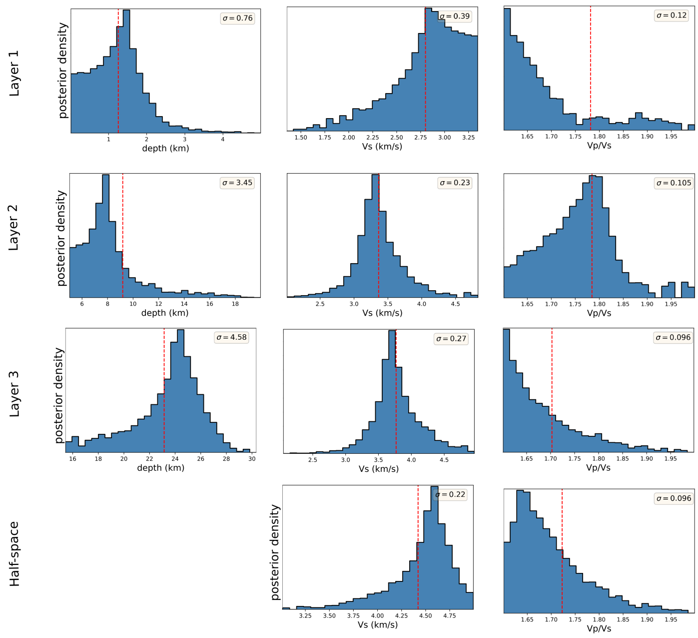

Figure 12. Same as Figure 4 for the inversion of data from station BFO

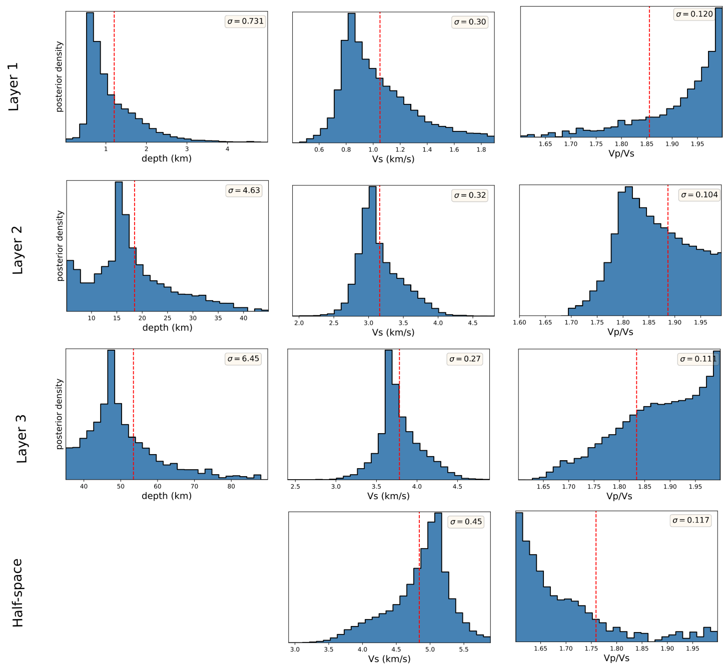

Figure 13. Same as Figure 4 for the inversion of data from station SUW

yet another produces no significant change, and then using AIC as a statistical inference test on all possible model families.

The method is successfully applied to synthetic seismograms generated for three apriori Mars subsurface models. Here we used both single and multiple events, and the uncertainty in the retrieved model parameters decreases with an increase in the size of the data set. We then applied the method on terrestrial data from three different seismic stations located in different geological settings. The resulting subsurface models were in good agreement with the results obtained in previous studies using diverse approaches which corroborated the efficacy of the method. Some aspects in applying this method to InSight data do warrant attention. The effect of location uncertainties will considerably affect the calculation of $v_{S_{,app}}$. Knapmeyer-Endrun et al. (2018) showed that the biggest affect in $v_{S_{,app}}$ can be caused by an uncertainty in distance and back-azimuth. A ±25% uncertainty in distance could yield an uncertainty of ±1 s/deg of the ray parameter for the P phase, while an erroneous back-azimuth will lead to a decrease in estimated v_S values at shorter periods. The thickness and velocity of a thin regolith layer can also be quite difficult to resolve if there is missing or erroneous information at short periods, as was the case in our study of terrestrial data. Another factor that limits the information that can be obtained from $v_{S_{,app}}$ on Mars is long period noise and effects of glitches (Scholz et al., 2020). Knapmeyer-Endrun et al. (2018) suggests that long period noise will affect longer periods while it has been observed that glitches can contaminate any part of the signal. Unlike the synthetics and terrestrial data used in this study, the $v_{S_{,app}}$ curve obtained from actual Mars data could be limited to much shorter periods. This would then increase the uncertainty in the retrieved v_S values at larger depths. A similar situation was encountered in Drilleau et al. (2020). In our previous study, Lognonné et al. (2020), we have been able to constrain the S-wave velocity and depth for the first inter-crustal layer of Mars between 1.7 to 2.1 km/s and 8 to 11 km, respectively, using such a limited $v_{S_{,app}}$ curve while further work involving the entire crust is in preparation. It is therefore important that all these factors are correctly accounted for.

Acknowledgments

R.J. acknowledges the funding provided by the IMPRS and the Emeritus group. The MPSMPG SEIS team acknowledges funding for development of the SEIS leveling system by the DLR German Space Agency. Seismic data for station BFO and SUW were obtained from the Federal Institute for Geosciences and Natural Resources and GEOFON data centre of the GFZ German Research Centre for Geosciences, respectively. This paper is InSight Contribution Number 216.

References

Akaike, H., Petrov, B. N., & Csaki, F. (1973). Second international symposium on information theory.

Ammon, C. J. (1991). The isolation of receiver effects from teleseismic P waveforms. *Bull. of the Seism. Soc. Am.*, *81*, 2504-2510.

Ammon, C. J., Randall, G. E., & Zandt, G. (1990). On the nonuniqueness of receiver function inversions. *Journal of Geophysical Research: Solid Earth*, *95*(B10), 15303–15318.

Anderson, D., & Burnham, K. (2004). Model selection and multi-model inference. *Second. NY: Springer-Verlag*, *63*(2020), 10.

Bayes, T. (1763). Lii. an essay towards solving a problem in the doctrine of chances. by the late Rev. Mr. Bayes, frs communicated by Mr. Price, in a letter to John Canton, amfr s. *Philosophical transactions of the Royal Society of London*(53), 370–418.

Birch, F. (1961). The velocities of compressional waves in rocks to 10 kilobars, Part 2. *J. Geophys. Res.*, *66*, 2199-2224.

Bogdanova, S., Gorbatschev, R., Grad, M., Janik, T., Guterch, A., Kozlovskaya, E., ... others (2006). Eurobridge: new insight into the geodynamic evolution of the East European Craton. *Geological Society, London, Memoirs*, *32*(1), 599–625.

Burnham, K. P., & Anderson, D. R. (2002). Model selection and.

Ceylan, S., van Driel, M., Euchner, F., Khan, A., Clinton, J., Krischer, L., ... Giardini, D. (2017). From initial models of seismicity, structure and noise to synthetic seismograms for Mars. *Space Science Reviews*, *211*(1-4), 595–610.

Chong, J., Ni, S., Chu, R., & Somerville, P. (2016). Joint inversion of body-wave receiver function and Rayleigh-wave ellipticity. *Bulletin of the Seismological Society of America*, *106*(2), 537–551.

Connolly, J. A. D. (2009). The geodynamic equation of state: What and how. *Geochemistry, Geophysics, Geosystems*, *10*(10). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009GC002540` doi: https://doi.org/10.1029/2009GC002540

Drilleau, M., Beucler, ., Lognonn, P., Panning, M. P., Knapmeyer-Endrun, B., Banerdt, W. B., ... Tharimena, S. (2020). MSS/1: Single-Station and Single-event Marsquake Inversion. *Earth and Space Science*, *7*(12), e2020EA001118. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020EA001118` (e2020EA001118 10.1029/2020EA001118) doi: https://doi.org/10.1029/2020EA001118

Du, Z., & Foulger, G. (1999). The crustal structure beneath the northwest fjords, Iceland, from receiver functions and surface waves. *Geophysical Journal International*, *139*(2), 419–432.

Dueker, K. G., & Sheehan, A. F. (1997). Mantle discontinuity structure from midpoint stacks of converted P to S waves across the Yellowstone hotspot track. *Journal of Geophysical Research: Solid Earth*, *102*(B4), 8313–8327.

Federal Institute for Geosciences and Natural Resources. (1976). *German regional seismic network (grsn).* Bundesanstalt fr Geowissenschaften und Rohstoffe. Retrieved from `https://www.seismologie.bgr.de/doi/grsn/` doi: 10.25928/MBX6-HR74

Fontaine, F. R., Barruol, G., Kennett, B. L., Bokelmann, G. H., & Reymond, D. (2009). Upper mantle anisotropy beneath Australia and Tahiti from P-wave polarization: Implications for real-time earthquake location. *Journal of Geophysical Research: Solid Earth*, *114*(B3).

Geissler, W. H., Kind, R., & Yuan, X. (2008). Upper mantle and lithospheric heterogeneities in central and Eastern Europe as observed by teleseismic receiver functions. *Geophysical Journal International*, *174*(1), 351–376.

GEOFON Data Centre. (1993). *Geofon seismic network.* Deutsches GeoForschungsZentrum GFZ. Retrieved from `http://geofon.gfz-potsdam.de/doi/network/GE` doi: 10.14470/TR560404

Grad, M., Jensen, S. L., Keller, G. R., Guterch, A., Thybo, H., Janik, T., ... others (2003). Crustal structure of the Trans-European suture zone region along POLONAISE'97 seismic profile P4. *Journal of Geophysical Research: Solid Earth*, *108*(B11).

Grad, M., Tiira, T., & Group, E. W. (2009). The moho depth map of the European Plate. *Geophysical Journal International*, *176*(1), 279–292.

Hannemann, K., Krüger, F., Dahm, T., & Lange, D. (2016). Oceanic lithospheric S wave velocities from the analysis of P wave polarization at the ocean floor. *Geophys. J. Int.*, *207*, 1796-1817. doi: 10.1093/gji/ggw342

Hannemann, K., Krüger, F., Dahm, T., & Lange, D. (2017). Structure of the oceanic lithosphere and upper mantle north of the Gloria fault in the eastern mid-Atlantic by receiver function analysis. *J. Geophys. Res.*, *122*, 7927-7950. doi: 10.1002/2016JB013582

Helffrich, G., & Thompson, D. (2010). A stacking approach to estimate vp/vs from receiver functions. *Geophysical Journal International*, *182*(2), 899–902.

Julia, J., Ammon, C. J., Herrmann, R., & Correig, A. M. (2000). Joint inversion of receiver function and surface wave dispersion observations. *Geophysical Journal International*, *143*(1), 99–112.

Jurkevics, A. (1988). Polarization analysis of three-component array data. *Bull. seism. Soc. Am.*, *78*(5), 1725-1743.

Khan, A., & Connolly, J. (2008). Constraining the composition and thermal state of Mars from inversion of geophysical data. *Journal of Geophysical Research: Planets*, *113*(E7).

Khan, A., van Driel, M., Böse, M., Giardini, D., Ceylan, S., Yan, J., ... others (2016). Single-station and single-event marsquake location and inversion for structure using synthetic Martian waveforms. *Physics of the Earth and Planetary Interiors*, *258*, 28–42.

Kind, R., Kosarev, G., & Petersen, N. (1995). Receiver Functions at the Stations of the German Regional Seismic Network (GRSN). *Geophys. J. Int.*, *121*, 191-202.

Knapmeyer-Endrun, B., Ceylan, S., & van Driel, M. (2018). Crustal S-wave velocity from apparent incidence angles: a case study in preparation for Insight. *Space Science Reviews*, *214*(5), 83.

Knapmeyer-Endrun, B., Krüger, F., & Group, t. P. W. (2014). Moho depth across the Trans-European Suture zone from P-and S-receiver functions. *Geophysical Journal International*, *197*(2), 1048–1075.

Langston, C. A. (1979). Structure under Mount Rainier, Washington, inferred from teleseismic body waves. *Journal of Geophysical Research: Solid Earth*, *84*(B9), 4749–4762.

Lognonné, P., Banerdt, W., Pike, W., Giardini, D., Christensen, U., Garcia, R. F., ... others (2020). Constraints on the shallow elastic and anelastic structure of mars from insight seismic data. *Nature Geoscience*, *13*(3), 213–220.

Lognonné, P., Banerdt, W. B., Giardini, D., Pike, W., Christensen, U., Laudet, P., ... others (2019). Seis: Insights seismic experiment for internal structure of Mars. *Space Science Reviews*, *215*(1), 12.

Owens, T. J., Taylor, S. R., & Zandt, G. (1987). Crustal structure at regional seismic test network stations determined from inversion of broadband teleseismic P waveforms. *Bulletin of the Seismological Society of America*, *77*(2), 631–662.

Panning, M. P., Lognonné, P., Bruce Banerdt, W., Garcia, R., Golombek, M., Kedar, S., ... Wookey, J. (2017, Oct 01). Planned Products of the Mars Structure Service for the Insight Mission to Mars. *Space Science Reviews*, *211*(1), 611–650. doi: 10.1007/s11214-016-0317-5

Park, S., & Ishii, M. (2018). Near-surface compressional and shear wave speeds constrained by body-wave polarization analysis. *Geophysical Journal International*, *213*(3), 1559–1571.

Phinney, R. A. (1964). Structure of the Earth's crust from spectral behavior of long-period body waves. *Journal of Geophysical Research*, *69*(14), 2997–3017.

Sambridge, M. (1999a). Geophysical inversion with a neighbourhood algorithm – I. Searching a parameter space. *Geophys. J. Int.*, *138*, 479-494. doi: 10.1046/j.1365-246X.1999.00876.x

Sambridge, M. (1999b). Geophysical inversion with a neighbourhood algorithmii. appraising the ensemble. *Geophysical Journal International*, *138*(3), 727–746.

Schiffer, C., Stephenson, R., Oakey, G. N., & Jacobsen, B. H. (2016). The crustal structure of Ellesmere Island, Arctic Canadateleseismic mapping across a remote intraplate orogenic belt. *Geophysical Journal International*, *204*(3), 1579–1600.

Scholz, J.-R., Widmer-Schnidrig, R., Davis, P., Lognonné, P., Pinot, B., Garcia, R. F., ... others (2020). Detection, analysis, and removal of glitches from insight's seismic data from mars. *Earth and Space Science*, *7*(11).

Schulte-Pelkum, V., Masters, G., & Shearer, P. M. (2001). Upper mantle anisotropy from long-period P-wave polarization. *Journal of Geophysical Research: Solid Earth*, *106*(B10), 21917–21934.

Shibutani, T., Sambridge, M., & Kennett, B. (1996). Genetic algorithm inversion for receiver functions with application to crust and uppermost mantle structure beneath eastern Australia. *Geophysical Research Letters*, *23*(14), 1829–1832.

Svenningsen, L., & Jacobsen, B. (2007). Absolute S-velocity estimation from receiver functions. *Geophysical Journal International*, *170*(3), 1089–1094.

van Driel, M., Krischer, L., Stähler, S. C., Hosseini, K., & Nissen-Meyer, T. (2015). Instaseis: Instant global seismograms based on a broadband waveform database. *Solid Earth*, *6*(2), 701–717.

van Driel, M., Wassermann, J., Pelties, C., Schiemenz, A., & Igel, H. (2015, 07). Tilt effects on moment tensor inversion in the near field of active volcanoes. *Geophysical Journal International*, *202*(3), 1711-1721. Retrieved from `https://doi.org/10.1093/gji/ggv209` doi: 10.1093/gji/ggv209

Vinnik, L. (1977). Detection of waves converted from P to SV in the mantle. *Physics of the Earth and planetary interiors*, *15*(1), 39–45.

Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic bulletin & review*, *11*(1), 192–196.

Wathelet, M. (2008). An improved neighborhood algorithm: Parameter conditions and dynamic scaling. *Geophys. Res. Lett.*, *35*, L09301. doi: 10.1029/2008GL033256

Wiechert, E. (1907). Über Erdbebenwellen. I. Theoretisches über die Ausbreitung der Erdbebenwellen. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 415-429.

Wilde-Piórko, M., Grycuk, M., Polkowski, M., & Grad, M. (2017). On the rotation of teleseismic seismograms based on the receiver function technique. *Journal of seismology*, *21*(4), 857–868.

Zhu, L., & Kanamori, H. (2000). Moho Depth variations in southern California from teleseismic receiver functions. *J. geophys. Res.*, *105*(B2), 2969-2980.

Ziegler, P. A. (1992). European Cenozoic rift system. *Geodynamics of rifting*, *1*, 91–111.